



UNIVERSITÀ DEGLI STUDI DI PALERMO

Corso di Dottorato in Scienze Economiche, Statistiche, Psicologiche e Sociali
Dipartimento di Scienze Economiche, Aziendali e Statistiche
SECS-S/05

VARIANCE ESTIMATION FOR THE HORVITZ-THOMPSON TOTAL ESTIMATOR IN UNEQUAL PROBABILITY SAMPLING DESIGNS

IL DOTTORE
Roberto Sichera

IL COORDINATORE
Chiar.mo Prof. Vito M. R. Muggeo

IL TUTOR
Chiar.ma Prof.ssa Ornella Giambalvo

CICLO XXXI
ANNO CONSEGUIMENTO TITOLO 2019

CONTENTS

Acknowledgements	11
Introduction	13
1 INFERENCE FOR FINITE POPULATIONS AND UNEQUAL PROBABILITY SAMPLING	15
1.1 Finite Population inference: General Framework	15
1.2 Unequal Probability Sampling	17
1.2.1 Optimality conditions	17
1.2.2 Computation of first-order inclusion probabilities	18
1.2.3 Horvitz–Thompson estimation	18
1.2.4 Limitations	20
2 SOME SOLUTIONS FOR ESTIMATING VARIANCE	21
2.1 Random Groups	21
2.1.1 Selection of Random Groups	22
2.2 Jackknife	22
2.3 Approximate π ps designs	24
2.3.1 Wright sampling	24
2.3.2 Sunter sampling	24
2.3.3 Poisson Sampling	26
2.4 List–Sequential sampling designs	27
2.4.1 Chao sampling	28
2.4.2 Tillé sampling	30
2.5 Bootstrap	31
2.5.1 Bootstrap for finite populations	32
2.6 Approximate Variance Estimators	36
2.6.1 Entropy	36
2.6.2 Conditional Poisson Sampling	37
2.6.3 Approximate estimators	38
2.6.4 Performances	43
2.6.5 Conclusions	45
2.7 Monte Carlo approach	46
3 VARIANCE APPROXIMATIONS: A COMPARISON THROUGH A SIMULATION STUDY	49
3.1 Methodology	50
3.2 Results	52
3.3 Conclusions	66
4 INFLUENCE AND ROBUST ESTIMATION	67
4.1 Influence measures	67
4.2 Robust Horvitz–Thompson estimator	68

4.3	Estimating the MSE of the Robust Horvitz–Thompson estimator	69
4.3.1	Analytic estimators	69
4.3.2	Bootstrap	74
4.4	Conclusions	76
5	AN APPLICATION TO A REAL STUDY: ESTIMATING THE NUMBER OF ROAMING DOGS IN AN URBAN AREA	79
5.1	Background	79
5.2	Features of the survey and data collection	80
5.2.1	Data collection settings and procedure	80
5.2.2	Data collection form	81
5.3	Sampling procedure	82
5.4	Estimation	85
5.5	Conclusions	87
	Conclusions	89
A	SAMPLING PROCEDURES	91
A.1	Simple Random Sampling Without replacement	91
A.2	Stratified Random Sampling	91
A.3	Brewer method	91
A.3.1	Derivation of working probabilities	92
A.3.2	Sampling procedure	93
A.3.3	Joint inclusion probabilities	94
A.3.4	Proofs	94
A.4	Rao–Sampford sampling	95
A.5	Randomised Systematic Sampling	96
B	RESULTS FROM THE SIMULATION STUDY ON VARIANCE APPROXIMATIONS	97
C	R PACKAGES	157
C.1	jipApprox	157
C.2	UPSvarApprox	158
C.3	bootstrapFP	159
C.4	Other Packages	160
	References	166

LIST OF TABLES

Table 2.1	Generation of weight adjustments for the generalised bootstrap.	36
Table 3.1	Summary of population characteristics.	50
Table 3.2	Correlation between X and Y for the Sukhatme population space.	51
Table 3.3	Correlation between X and Y for the MU273 population space.	51
Table 3.4	Characteristics of the simulation study.	52
Table 3.5	Relative Bias summary statistics, MU273 population.	57
Table 3.6	Relative Bias summary statistics, Sukhatme population.	57
Table 3.7	Relative Bias summary statistics by sampling design, MU273 population.	58
Table 3.8	Relative Bias summary statistics by sampling design, Sukhatme population.	59
Table 3.9	Relative Stability summary statistics, MU273 population.	60
Table 3.10	Relative Stability summary statistics, Sukhatme population.	60
Table 3.11	Relative Stability summary statistics by sampling design, MU273 population.	61
Table 3.12	Relative Stability summary statistics by sampling design, Sukhatme population.	62
Table 3.13	MSE summary statistics, MU273 population.	63
Table 3.14	MSE summary statistics, Sukhatme population.	63
Table 3.15	MSE summary statistics by sampling design, MU273 population.	64
Table 3.16	MSE summary statistics by sampling design, Sukhatme population.	65
Table 4.1	Relative Bias of the MSE estimators of $\hat{\vartheta}_{RHT}$	73
Table 4.2	Relative Bias of bootstrap variance estimators.	76
Table 5.1	Sample data for the roaming dogs study.	86
Table 5.2	Variance estimation for the HT and RHT estimators using different approaches.	87
Table B.1	Relative Bias summary statistics, MU273 population.	98
Table B.2	Relative Bias summary statistics, Sukhatme population.	98

Table B.3	Relative Stability summary statistics, MU273 population.	99
Table B.4	Relative Stability summary statistics, Sukhatme population.	99
Table B.5	MSE summary statistics, MU273 population.	100
Table B.6	MSE summary statistics, Sukhatme population.	100
Table B.7	Relative Bias summary statistics by sampling design, MU273 population.	101
Table B.8	Relative Bias summary statistics by sampling design, Sukhatme population.	102
Table B.9	Relative Stability summary statistics by sampling design, MU273 population.	103
Table B.10	Relative Stability summary statistics by sampling design, Sukhatme population.	104
Table B.11	MSE summary statistics by sampling design, MU273 population.	105
Table B.12	MSE summary statistics by sampling design, Sukhatme population.	106
Table B.13	Simulation results for MU273 population: $N = 273$, $CV(Y) = 1.29$, $CV(X) = 1.27$	107
Table B.14	Simulation results for MU273 population: $N = 273$, $CV(Y) = 1.29$, $CV(X) = 1$	108
Table B.15	Simulation results for MU273 population: $N = 273$, $CV(Y) = 1.29$, $CV(X) = 0.81$	109
Table B.16	Simulation results for MU273 population: $N = 273$, $CV(Y) = 1.29$, $CV(X) = 0.67$	110
Table B.17	Simulation results for MU273 population: $N = 273$, $CV(Y) = 1.29$, $CV(X) = 0.5$	111
Table B.18	Simulation results for MU273 population: $N = 273$, $CV(Y) = 1$, $CV(X) = 1.27$	112
Table B.19	Simulation results for MU273 population: $N = 273$, $CV(Y) = 1$, $CV(X) = 1$	113
Table B.20	Simulation results for MU273 population: $N = 273$, $CV(Y) = 1$, $CV(X) = 0.81$	114
Table B.21	Simulation results for MU273 population: $N = 273$, $CV(Y) = 1$, $CV(X) = 0.67$	115
Table B.22	Simulation results for MU273 population: $N = 273$, $CV(Y) = 1$, $CV(X) = 0.5$	116
Table B.23	Simulation results for MU273 population: $N = 273$, $CV(Y) = 0.82$, $CV(X) = 1.27$	117
Table B.24	Simulation results for MU273 population: $N = 273$, $CV(Y) = 0.82$, $CV(X) = 1$	118
Table B.25	Simulation results for MU273 population: $N = 273$, $CV(Y) = 0.82$, $CV(X) = 0.81$	119

Table B.26	Simulation results for MU273 population: N = 273, CV(Y) = 0.82, CV(X) = 0.67	120
Table B.27	Simulation results for MU273 population: N = 273, CV(Y) = 0.82, CV(X) = 0.5	121
Table B.28	Simulation results for MU273 population: N = 273, CV(Y) = 0.67, CV(X) = 1.27	122
Table B.29	Simulation results for MU273 population: N = 273, CV(Y) = 0.67, CV(X) = 1	123
Table B.30	Simulation results for MU273 population: N = 273, CV(Y) = 0.67, CV(X) = 0.81	124
Table B.31	Simulation results for MU273 population: N = 273, CV(Y) = 0.67, CV(X) = 0.67	125
Table B.32	Simulation results for MU273 population: N = 273, CV(Y) = 0.67, CV(X) = 0.5	126
Table B.33	Simulation results for MU273 population: N = 273, CV(Y) = 0.5, CV(X) = 1.27	127
Table B.34	Simulation results for MU273 population: N = 273, CV(Y) = 0.5, CV(X) = 1	128
Table B.35	Simulation results for MU273 population: N = 273, CV(Y) = 0.5, CV(X) = 0.81	129
Table B.36	Simulation results for MU273 population: N = 273, CV(Y) = 0.5, CV(X) = 0.67	130
Table B.37	Simulation results for MU273 population: N = 273, CV(Y) = 0.5, CV(X) = 0.5	131
Table B.38	Simulation results for Sukhatme population: N = 34, CV(Y) = 1.05, CV(X) = 1.1	132
Table B.39	Simulation results for Sukhatme population: N = 34, CV(Y) = 1.05, CV(X) = 0.96	133
Table B.40	Simulation results for Sukhatme population: N = 34, CV(Y) = 1.05, CV(X) = 0.77	134
Table B.41	Simulation results for Sukhatme population: N = 34, CV(Y) = 1.05, CV(X) = 0.67	135
Table B.42	Simulation results for Sukhatme population: N = 34, CV(Y) = 1.05, CV(X) = 0.5	136
Table B.43	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.95, CV(X) = 1.1	137
Table B.44	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.95, CV(X) = 0.96	138
Table B.45	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.95, CV(X) = 0.77	139
Table B.46	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.95, CV(X) = 0.67	140
Table B.47	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.95, CV(X) = 0.5	141
Table B.48	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.76, CV(X) = 1.1	142

Table B.49	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.76, CV(X) = 0.96	143
Table B.50	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.76, CV(X) = 0.77	144
Table B.51	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.76, CV(X) = 0.67	145
Table B.52	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.76, CV(X) = 0.5	146
Table B.53	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.67, CV(X) = 1.1	147
Table B.54	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.67, CV(X) = 0.96	148
Table B.55	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.67, CV(X) = 0.77	149
Table B.56	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.67, CV(X) = 0.67	150
Table B.57	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.67, CV(X) = 0.5	151
Table B.58	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.5, CV(X) = 1.1	152
Table B.59	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.5, CV(X) = 0.96	153
Table B.60	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.5, CV(X) = 0.77	154
Table B.61	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.5, CV(X) = 0.67	155
Table B.62	Simulation results for Sukhatme population: N = 34, CV(Y) = 0.5, CV(X) = 0.5	156

LIST OF FIGURES

Figure 3.1	Relative Bias distribution of variance estimators	54
Figure 3.2	MSE distribution of variance estimators	55
Figure 5.1	Data collection form (2014) – front	83
Figure 5.2	Data collection form (2014) – back	84

ACKNOWLEDGEMENTS

Over the course of my PhD I met many people who enriched these years with their ideas, opinions, help, advices, making it an unique experience. Although acknowledging everyone would be impossible, I am grateful to each of them for their contribution to this journey.

In particular, I wish to thank my parents for their support, and Nathalie for being by my side the whole time.

I would like to express my appreciation to the staff and professors of the department of Economics, Business and Statistics of the University of Palermo, as well as to the coordinator of the PhD program, for their work.

I also wish to mention some of the colleagues with whom I shared this experience. Ieva Mikaliūnaitė, who started the PhD course with me. Andrea Marletta and Federico Torretta, who warmly welcomed me in my first year. Marianna Siino, one of the nicest people I have met. Ola Hajjaj, who restlessly trained my patience. Marco Ciziceno and Pietro Pizzuto, source of information and new perspectives, who provided me with a few challenges. Fabio Parla, for the stimulating conversations. Simona Diliberto, for her inspiring enthusiasm.

A special thanks to prof. Linda Gattuso, for being my guide in Montréal, Canada, during my research period there.

I am particularly grateful to prof. David Haziza, for contributing to part of the project with his precious suggestions.

Above all, I wish to thank prof. Ornella Giambalvo, who introduced me to survey sampling and made this PhD project possible, for her patient and thorough supervision along these years, and for teaching me more than I could expect.

INTRODUCTION

This thesis focuses on variance estimation of the Horvitz–Thompson estimator in Unequal Probability Sampling, with particular interest to the case of small populations.

The motivation for this PhD project comes from a study carried out in Palermo, Italy, for estimating the total of roaming dogs in the first city district. The study is described in details in Chapter 5. In the aforementioned study, a sample of $n = 12$ out of $N = 76$ areas was drawn according to a FPDUST spatial sampling with Probabilities Proportional to Size (PPS) (Barabesi et al., 1997) and estimation was performed through the Horvitz–Thompson estimator. The study presented two issues. First, variance could not be directly estimated because exact inclusion probabilities are not available for FPDUST sampling. Second, data were affected by the presence of outliers.

Compared to equal probability designs, Unequal Probability Sampling approach is attractive as it provides better efficiency, under certain conditions (Cochran, 1977). A large number of sampling designs with unequal probabilities have been proposed (Brewer and Hanif, 1982; Tillé, 2006), however, estimation of the Horvitz–Thompson variance may be unfeasible due to the high computational complexity of joint inclusion probabilities, which quickly increases with population and sample size. A number of solutions has been proposed in literature, such as sampling procedures that produce joint inclusion probabilities that can be expressed in a simple closed–form equation, or the more recent approximate variance estimators, which perform well under high entropy sampling and when population and sample are large. The more flexible Monte Carlo simulation approach will also be taken into consideration.

The aim of this thesis is to evaluate and compare the main available solutions, in order to determine which are the scenarios where the use of any of them may be convenient and even preferable. In particular, the main interest is to find an optimal strategy for variance estimation in small samples. Also, the case in which outliers are present is considered; a summary of the main influence measures for finite–population inference is presented, and then an attempt to estimate the MSE of the robust Horvitz–Thomson estimator (Beaumont, Haziza, et al., 2013) is made .

In Chapter 1, the general framework of finite–population inference is pictured, the Unequal Probability Sampling approach is described and the difficulties of estimating the variance under this settings are discussed. The second chapter illustrates the main solutions available in literature for estimating variance under an Unequal Probability

Sampling design and discusses their strengths and weaknesses. The third chapter compares two approaches to approximate the variance: approximate variance estimators and Monte Carlo approximation of inclusion probabilities. A simulation study that compares such techniques under a large number of scenarios is presented, in order to highlight when each method is preferable. Chapter 4 summarises the measures of influence for finite-population inference and the robust Horvitz–Thompson estimator, and presents some proposals for estimating the MSE of such estimator. Finally, Chapter 5 shows an application to the roaming–dog survey. Furthermore, appendix A describes some of the sampling algorithms mentioned in this thesis and appendix B presents the results from the simulation study in chapter 3, while appendix C briefly illustrates the R packages created to support the analyses for this thesis.

INFERENCE FOR FINITE POPULATIONS AND UNEQUAL PROBABILITY SAMPLING

1.1 FINITE POPULATION INFERENCE: GENERAL FRAMEWORK

In finite–population inference, we consider a *population* $U = \{1, 2, \dots, N\}$ consisting of a finite number N of *identifiable* units, and a *variable of interest* Y which is unknown, but deterministic and measurable for each population unit. Given a *sample* s of U , the aim is to estimate a parameter ϑ of the population by means of a function of the observed values $y_i, i \in s$:

$$\hat{\vartheta} = f(y_i, i \in s)$$

Generally speaking, a sample is a subset of the population; in its broadest sense, it is one of all possible subsets that can be generated from U . However, usually only a particular subset of such partition is of interest. For instance, one may be interested only in the samples with a given number of units, which are referred to as *fixed–size* samples, or even on those generated through a selection without replacement of the units. Throughout this thesis, if not otherwise specified, we will be interested in fixed–size samples without replacement, with the exception of Poisson sampling, which produces samples with variable size, and we will denote the sample space with S .

Selection of sample units is performed by means of some kind of sampling procedure, or *sampling design*, which may be probabilistic or non–probabilistic. Probability sampling, which is the subject of this thesis, includes those designs that satisfy the following properties (Cochran, 1977):

- It is possible to define the set S of samples s and precisely know which units belong to any sample $s \in S$;
- A *strictly* positive selection probability can be assigned to each sample: $p(s) > 0, s \in S$. One sample s is then selected with probability $p(s)$;
- Given an estimator $\hat{\vartheta}$ of a population characteristic Y , each sample must lead to a unique estimate.

A combination of a sampling design and an estimator is defined *sampling strategy*.

In practice, listing all possible samples and computing their probability $p(s)$ would usually be a laborious task, which is especially true

for large populations. Thus, rather than attaching selection probabilities to samples, it is usually preferred to assign *inclusion probabilities* to population units, and draw units until a sample of the desired size is obtained (Cochran, 1977).

Inclusion probabilities are defined as the probability that one or more units are included in the sample. We define as *inclusion probability of the k -th order* the probability that a group of $k = 1, \dots, n$ units is included in the sample. In practice, most times only the first and second-order (or joint) inclusion probabilities are needed for estimation.

Let us define the random variables δ_i and δ_{ij} , which represent the inclusion in the sample of unit i and couple (i, j) , respectively, as

$$\delta_i = \begin{cases} 1, & \text{if } i \in s \\ 0, & \text{otherwise} \end{cases} \quad \delta_{ij} = \begin{cases} 1, & \text{if } (i, j) \in s \\ 0, & \text{otherwise} \end{cases}$$

The δ_i variables satisfy the following properties (Hájek, 1981, p. 22):

$$E[\delta_i] = \sum_{s \in S} \delta_i(s) p(s) = \sum_{s \ni i} p(s) = \pi_i \quad (1.1a)$$

$$E[\delta_i \delta_j] = \sum_{s \in S} \delta_i(s) \delta_j(s) p(s) = \sum_{s \ni \{i, j\}} p(s) = \pi_{ij} \quad (1.1b)$$

$$\text{Var}[\delta_i] = E[\delta_i^2] - E[\delta_i]^2 = E[\delta_i] - E[\delta_i]^2 = \pi_i - \pi_i^2 = \pi_i(1 - \pi_i) \quad (1.1c)$$

$$\text{Cov}[\delta_i, \delta_j] = E[\delta_i \delta_j] - E[\delta_i] E[\delta_j] = \pi_{ij} - \pi_i \pi_j, \quad (1.1d)$$

where $E[\cdot]$, $\text{Var}[\cdot]$ and $\text{Cov}[\cdot]$ represent design expectation, design variance and design covariance, respectively. The term *design* indicates that we consider the expectation, variance or covariance over all possible samples of the sample space S .

Moreover, let us denote first-order inclusion probabilities by π_i , and joint-inclusion probabilities by π_{ij} . Inclusion probabilities are then defined as:

$$\pi_i = P(i \in s) = P(\delta_i = 1) = \sum_{s \ni i} p(s) \quad (1.2a)$$

$$\pi_{ij} = P(i, j \in s) = P(\delta_{ij} = 1) = \sum_{s \ni \{i, j\}} p(s) \quad (1.2b)$$

The approach described above is referred to as design-based. In this approach, the sampling process has a central role, as the estimates of both the parameter of interest and the variance strictly depend on its choice. This may be considered as the main difference between design-based inference and infinite-population inference, where estimates depend on the probability distribution of Y . The other approach to finite-population inference, called model-based

(or model-assisted, or prediction-based), will not be discussed in this thesis. For details about the model-based approach, see Brewer (2002), Estevao et al. (1995), Särndal (1996), and Särndal et al. (1992)

1.2 UNEQUAL PROBABILITY SAMPLING

When inclusion probabilities are equal for each population unit, the design is said to be an *equal probability sampling* design, otherwise, it is called an Unequal Probability Sampling (UPS) design. In this thesis, we are interested in the latter group, and especially to designs with Probability Proportional to Size (PPS), where the probability of selecting an unit from U is proportional to an auxiliary variable X , or size variable. When, in addition, inclusion probabilities π_i are proportional to the size variable, we indicate the design by πps . Variable X should be positively correlated to the variable of interest Y , and its value must be known for all population units prior to sampling.

UPS designs are highly attractive because, taking advantage of the auxiliary information provided by the size variable X , they usually produce more efficient estimates than equal probability sampling designs, under certain conditions that are listed in the next subsection.

1.2.1 Optimality conditions

Let us denote by x_i the value of the size variable for the i -th unit and by $n(s)$ the number of distinct units in sample s . Hanurav (1967) defined the following desirable properties for the inclusion probabilities:

$$\pi_i = n \frac{x_i}{\sum_{i \in U} x_i} \quad (1.3a)$$

$$n(s) = n \quad \forall s \text{ such that } p(s) > 0 \quad (1.3b)$$

$$\pi_{ij} > 0 \quad (1.3c)$$

$$\pi_{ij} \leq \pi_i \pi_j \quad (1.3d)$$

$$\frac{\pi_{ij}}{\pi_i \pi_j} > \beta \quad \text{with } \beta \text{ not too close to } 0 \quad (1.3e)$$

A sampling design satisfying such conditions provides optimal estimates, in the sense that the total estimator $\hat{\vartheta}$ would produce estimates that are close to the true total $\vartheta = \sum_{i \in U} y_i$, and that the variance would be unbiased, stable and nonnegative.

If property (1.3b) is satisfied, then the following properties for inclusion probabilities hold:

$$\sum_{i \in U} \pi_i = n \quad (1.4a)$$

$$\sum_{\substack{j \in U \\ j \neq i}} \pi_{ij} = (n-1)\pi_i \quad (1.4b)$$

$$\sum_{\substack{j \in U \\ j \neq i}} \pi_i \pi_j = \pi_i(n - \pi_i) \quad (1.4c)$$

$$\sum_{i \in U} \sum_{j \neq i} \pi_{ij} = n(n-1) \quad (1.4d)$$

1.2.2 Computation of first-order inclusion probabilities

Given Hanurav (1967) conditions for optimum sampling strategies, the common procedure to compute first-order inclusion probabilities for most UPS designs is:

1. For all $i = 1, \dots, k$, compute inclusion probabilities as

$$\pi_i = n \frac{x_i}{\sum_{i \in U} x_i};$$

2. If for some units $\pi_i \geq 1$, denote by $U^s = \{i \mid \pi_i \geq 1\}$ the set of such units and with $k = \mathbf{card}(U^s)$ its cardinality, define also the reduced population $U^* = U \setminus U^s$. Then, $\forall j \in U^*$, set

$$\begin{cases} \pi_j = (n-k) \frac{x_j}{\sum_{j \in U^*} x_j} & \text{if } j \in U^* \\ \pi_j = 1 & \text{if } j \in U^s \end{cases}$$

3. Repeat *step 2* until $\pi_i \leq 1, \forall i \in U$.

Naturally, if any $\pi_i = 1$, these will be included in the sample with certainty. Such units are called *self-selecting* units.

1.2.3 Horvitz–Thompson estimation

In this context, estimates are most commonly obtained by the Horvitz–Thompson total estimator (Horvitz and Thompson, 1952), which weights sample observations by the reciprocal of the inclusion probabilities π_i . Given a sample $s = \{1, \dots, n\}$ of population $U = \{1, \dots, N\}$ and observations $y_i, i \in s$, the Horvitz–Thompson estimator is defined as

$$\hat{\vartheta}_{HT} = \sum_{i \in s} \frac{y_i}{\pi_i} = \sum_{i \in U} \frac{y_i \delta_i}{\pi_i}, \quad (1.5)$$

Noting that in (1.5) the only random variable is δ_i , and using identity (1.1a) one finds that

$$E[\hat{\vartheta}_{HT}] = E\left[\sum_{i \in U} \frac{y_i \delta_i}{\pi_i}\right] = \sum_{i \in U} \frac{y_i}{\pi_i} E[\delta_i] = \sum_{i \in U} \frac{y_i}{\pi_i} \pi_i = \sum_{i \in U} y_i = \vartheta$$

Therefore $\hat{\vartheta}_{HT}$ is unbiased and, by identities (1.1c) and (1.1d), its variance is given by

$$\begin{aligned} \text{Var}_{HT}[\hat{\vartheta}_{HT}] &= V\left[\sum_{i \in U} \frac{y_i \delta_i}{\pi_i}\right] = \\ &= \sum_{i \in U} \left(\frac{y_i}{\pi_i}\right)^2 V[\delta_i] + 2 \sum_{i \in U} \sum_{j > i} \frac{y_i y_j}{\pi_i \pi_j} \text{Cov}[\delta_i, \delta_j] = \\ &= \sum_{i \in U} \left(\frac{y_i}{\pi_i}\right)^2 \pi_i (1 - \pi_i) + 2 \sum_{i \in U} \sum_{j > i} \frac{y_i y_j}{\pi_i \pi_j} (\pi_{ij} - \pi_i \pi_j) = \\ &= \sum_{i \in U} \frac{1 - \pi_i}{\pi_i} y_i^2 + 2 \sum_{i \in U} \sum_{j > i} (\pi_{ij} - \pi_i \pi_j) \frac{y_i y_j}{\pi_i \pi_j} \end{aligned} \quad (1.6)$$

For fixed n , Sen (1953) and Yates and Grundy (1953) derived an alternative form of (1.6), which expression is

$$\text{Var}_{SYG}[\hat{\vartheta}_{HT}] = \sum_{i \in U} \sum_{j > i} (\pi_i \pi_j - \pi_{ij}) \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2 \quad (1.7)$$

Equation (1.7) is obtained by considering relations (1.4a) and (1.4b), which lead to the identity

$$\begin{aligned} \sum_{\substack{j \in U \\ j \neq i}} (\pi_{ij} - \pi_i \pi_j) &= \sum_{\substack{j \in U \\ j \neq i}} \pi_{ij} - \sum_{j \neq i} \pi_i \pi_j \\ &= (n - 1)\pi_i - \pi_i \sum_{\substack{j \in U \\ j \neq i}} \pi_j \\ &= (n - 1)\pi_i - \pi_i (n - \pi_i) \\ &= -\pi_i (1 - \pi_i) \end{aligned}$$

which gives

$$1 - \pi_i = \frac{\sum_{j \neq i} (\pi_{ij} - \pi_i \pi_j)}{-\pi_i} = \frac{\sum_{j \neq i} (\pi_i \pi_j - \pi_{ij})}{\pi_i} \quad (1.8)$$

By substituting the quantity $1 - \pi_i$ in the Horvitz–Thompson variance with the right-hand term of equation (1.8), the first term of (1.6) becomes (Cochran, 1977, p. 260)

$$\sum_{i \in U} \frac{(1 - \pi_i)}{\pi_i} y_i^2 = \sum_{i \in U} \sum_{j \neq i} (\pi_i \pi_j - \pi_{ij}) \left(\frac{y_i}{\pi_i}\right)^2 = \sum_{i \in U} \sum_{j > i} (\pi_i \pi_j - \pi_{ij}) \left[\left(\frac{y_i}{\pi_i}\right)^2 + \left(\frac{y_j}{\pi_j}\right)^2 \right]$$

Thus,

$$\begin{aligned} \text{Var}_{\text{SYG}}[\hat{\vartheta}_{\text{HT}}] &= \sum_{i \in \mathcal{U}} \sum_{j > i} (\pi_i \pi_j - \pi_{ij}) \left[\left(\frac{y_i}{\pi_i} \right)^2 + \left(\frac{y_j}{\pi_j} \right)^2 - 2 \frac{y_i y_j}{\pi_i \pi_j} \right] \\ &= \sum_{i \in \mathcal{U}} \sum_{j > i} (\pi_i \pi_j - \pi_{ij}) \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j} \right)^2 \end{aligned}$$

Unbiased sample estimators of (1.6) and (1.7) are respectively given by

$$v_{\text{HT}}[\hat{\vartheta}_{\text{HT}}] = \sum_{i \in s} \frac{1 - \pi_i}{\pi_i^2} y_i^2 + 2 \sum_{i \in s} \sum_{j > i} \frac{(\pi_{ij} - \pi_i \pi_j)}{\pi_i \pi_j \pi_{ij}} y_i y_j \quad (1.9)$$

and

$$v_{\text{SYG}}[\hat{\vartheta}_{\text{HT}}] = \sum_{i \in s} \sum_{j > i} \frac{(\pi_i \pi_j - \pi_{ij})}{\pi_{ij}} \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j} \right)^2 \quad (1.10)$$

1.2.4 Limitations

Variance estimators (1.9) and (1.10) are unbiased;+ however, they can assume negative values, and be unstable when inclusion probabilities are small. If conditions (1.3a) to (1.3d) hold, the Sen–Yates–Grundy (SYG) estimator is always nonnegative and it is often more efficient than the Horvitz–Thompson (HT) variance estimator.

Moreover, both estimators depend on joint–inclusion probabilities π_{ij} , which computation may be unfeasible for some sampling designs. In fact, considering equation (1.2b) it can be seen that, when n and N are relatively large, the number of possible samples is so high that estimating π_{ij} becomes unfeasible. As an example, let us consider the sample design proposed by Brewer (1975), for which joint–inclusion probabilities are expressed through a recursive formula (for details about Brewer’s sampling, see Appendix A):

$$\pi_{IJ}(n) = P_I(n) \frac{\pi_J(n-1)}{1 - P_I} + P_J(n) \frac{\pi_I(n-1)}{1 - P_J} + \sum_{K \neq I, J}^N P_K(n) \pi_{IJ}^{(K)}(n-1)$$

which clearly shows how the computation of π_{ij} quickly becomes overwhelming as N increases. Furthermore, many traditional sampling procedures do not even have a close–form for the π_{ij} formula for the case $n > 2$ (see Brewer and Hanif (1982) for a list of more than fifty sampling designs).

In the next chapter, a review of the main solutions to the problem of variance estimation in UPS sampling is presented.

As mentioned in the previous chapter, estimating the variance of the Horvitz–Thompson estimator may be unfeasible, either because joint-inclusion probabilities are too computationally demanding, especially for large sample and population sizes, or because they are not available in closed form.

In this chapter, a review of the literature is performed, and a summary of the main solutions proposed for variance estimation in unequal probability sampling is presented.

2.1 RANDOM GROUPS

Random Groups method is a simple and flexible way to estimate the variance in complex surveys.

The main idea is to split the sample in k subsamples (*groups*) of size m , such that $k = n/m$, and computing an estimate of the parameter of interest $\hat{\vartheta}_g$ for each group g , $g = 1, \dots, k$ (Wolter, 1985, chapter 2). The general estimate is then found as an average of the k group estimators $\hat{\vartheta}_g$:

$$\hat{\vartheta}_{\text{RG}} = \frac{1}{k} \sum_{g=1}^k \hat{\vartheta}_g \quad (2.1)$$

Assuming independence among the groups, the variance of (2.1) is given by

$$\text{Var}[\hat{\vartheta}_{\text{RG}}] = \text{Var}\left[\frac{1}{k} \sum_{g=1}^k \hat{\vartheta}_g\right] = \frac{1}{k^2} \sum_{g=1}^k \text{Var}[\hat{\vartheta}_g] = \frac{1}{k} \text{Var}[\hat{\vartheta}_g] \quad (2.2)$$

A natural estimator of (2.2) is

$$v_1(\hat{\vartheta}_{\text{RG}}) = \frac{1}{k} \left\{ \frac{1}{k-1} \sum_{g=1}^k (\hat{\vartheta}_g - \hat{\vartheta}_{\text{RG}})^2 \right\}. \quad (2.3)$$

Alternatively, the variance of $\hat{\vartheta}_{\text{RG}}$ can be estimated by replacing $\hat{\vartheta}_{\text{RG}}$ with $\hat{\vartheta}$, the estimate computed on the whole sample:

$$v_2(\hat{\vartheta}_{\text{RG}}) = \frac{1}{k} \left\{ \frac{1}{k-1} \sum_{g=1}^k (\hat{\vartheta}_g - \hat{\vartheta})^2 \right\}. \quad (2.4)$$

When the estimator of ϑ is linear, $\hat{\vartheta}_{\text{RG}} = \hat{\vartheta}$. For nonlinear estimators (2.4) is biased and we have $v_1(\hat{\vartheta}_{\text{RG}}) \leq v_2(\hat{\vartheta}_{\text{RG}})$, so (2.4) is preferable when a conservative variance estimate is desired (Wolter, 1985, chapter 2).

The advantage of this method is that the computation of joint-inclusion probabilities is avoided entirely.

2.1.1 Selection of Random Groups

Groups can be independent or correlated among each other. In the former case, $k \geq 2$ subsamples of size $m = n/k$ are drawn from the target population one by one, replacing them into the population at each step.

In the latter situation, the groups are formed after the overall sample is drawn by randomly selecting k subsamples of size m from the original sample, without replacement and according to the original sampling design; if n/k is not an integer, that is $n = km + q$, with $0 < q < k$, the q excess units may either be left out of the k random groups, or they may be added unit by unit to the first q random groups.

Estimator (2.3) is unbiased when random groups are selected independently of each other. For correlated groups, however, such estimator is biased; for large populations and small sampling fractions, this bias will tend to be negligible, but it could be large with small populations.

2.2 JACKKNIFE

Jackknife estimation is another subsample replication technique originally proposed by Quenouille (1949), who introduced it in the infinite population setting for bias reduction, and was successively suggested by Tukey (1958) as a variance estimation technique.

Given the independent and identically distributed (i.i.d) random variables Y_1, \dots, Y_n and an estimator $\hat{\vartheta}$ of the population parameter ϑ computed on the whole sample, let us randomly partition the sample in k groups of m units each, where $km = n$ and k, m and n are all integers. The jackknife estimator is then given by

$$\hat{\vartheta}_{JK} = \frac{1}{k} \sum_{g=1}^k \hat{\vartheta}_g$$

where $\hat{\vartheta}_g$ are called *pseudovalues* and are obtained as

$$\hat{\vartheta}_g = k\hat{\vartheta} - (k-1)\hat{\vartheta}_{(-g)}$$

and the $\hat{\vartheta}_{(-g)}$ are estimators of the same form as $\hat{\vartheta}$, but computed from the reduced sample of size $m(k-1)$ obtained by omitting the g -th group.

In finite population inference, a sample is splitted in k subsamples as for the case of correlated random groups, illustrated in the previous subsection, then the jackknife estimator is applied as showed

above (Wolter, 1985). For unequal probability sampling without replacement, using the Horvitz–Thompson estimator the pseudovalues are

$$\hat{\vartheta}_g = k \hat{\vartheta}_{\text{HT}} - (k-1) \hat{\vartheta}_{\text{HT}(-g)},$$

where

$$\hat{\vartheta}_{\text{HT}(-g)} = \sum_{i=1}^{m(k-1)} \frac{y_i}{\pi_i m(k-1)/n}$$

is the Horvitz–Thompson estimator obtained from the reduced sample where units of the group g are removed.

The variance of the jackknife estimator can be estimated by (Tukey, 1958):

$$v(\hat{\vartheta}_{\text{JK}}) = \frac{1}{k(k-1)} \sum_{g=1}^k (\hat{\vartheta}_g - \hat{\vartheta}_{\text{JK}})^2. \quad (2.5)$$

Berger and Skinner (2005) proposed a jackknife estimator for the variance of a generic estimator of a function of means $\hat{\vartheta} = f(\hat{\mu}_1, \dots, \hat{\mu}_Q)$ in unequal probability sampling, where $\hat{\mu}_q = \sum_{i \in s} w_i y_{qi}$, with $w_i = \frac{1}{N \pi_i}$ and $\hat{N} = \sum_{i \in s} \frac{1}{\pi_i}$. The jackknife variance estimator is defined as

$$v_{\text{JK}}(\hat{\vartheta}) = \sum_{i \in s} \sum_{j \in s} \frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}} \varepsilon_{(i)} \varepsilon_{(j)}, \quad (2.6)$$

with

$$\begin{aligned} \varepsilon_{(i)} &= (1 - w_i)(\hat{\vartheta} - \hat{\vartheta}_{(i)}); \\ \hat{\vartheta}_{(i)} &= f(\hat{\mu}_{1(i)}, \dots, \hat{\mu}_{Q(i)}); \\ \hat{\mu}_{q(i)} &= \sum_{j \in s_{-i}} \frac{y_{qj}}{\hat{N}_{(i)} \pi_j}; \\ \hat{N}_{(i)} &= \sum_{j \in s_{-i}} \frac{1}{\pi_j}, \end{aligned}$$

where s_{-i} is the reduced sample with the i -th unit removed.

To use variance (2.6), joint-inclusion probabilities π_{ij} need to be known, if they are unknown, these can be approximated by one of the methods proposed in subsection 2.6. The authors suggest the use of Hájek (1964) approximation (see section 2.6):

$$\pi_{ij} \approx \pi_i \pi_j [1 - d^{-1}(1 - \pi_i)(1 - \pi_j)], \quad \text{with } i \neq j \quad (2.7)$$

where $d = \sum_{i \in U} \pi_i (1 - \pi_i)$.

The authors proved that (2.6) is design-consistent and showed through a simulation study that its relative bias and relative root-mean-square-error is lower than those of estimator (2.5).

Jackknife variance estimators have also been proposed for stratified sampling (Berger, 2007) and with adjustment for imputed data (Berger and Rao, 2006).

2.3 APPROXIMATE π PS DESIGNS

The properties listed in section 1.2.1 define a π ps design, namely, a sampling design with selection probabilities proportional to the inclusion probabilities π_i . A few sampling procedures have been formulated which manage to provide easy-to-compute joint-inclusion probabilities by relaxing one or more of such properties.

2.3.1 *Wright sampling*

For instance, Wright (1983, 1989) proposed a stratified simple random sampling which included approximate π PS features. In this case, proportionality to the auxiliary variable of each unit's inclusion probability (property (1.3a)) is relaxed, and units are selected with probability proportional to mean size in the stratum instead. Given a population U and a size measure with values x_i , Wright sampling is as follows:

1. Order the population units by increasing order of the x_i values;
2. Create H strata in such a way that the sum of the size variable in each stratum U_h , $h = 1, \dots, H$, is one H -th of the total in the population U (as closely as possible). That is, $\sum_{i \in U_h} x_i \approx X/H$, with $X = \sum_{i \in U} x_i$. To do so, take the first N_1 units such that $\sum_{i=1}^{N_1} x_i \approx X/H$, the second N_2 units such that $\sum_{i=(N_1+1)}^{N_2} x_i \approx X/H$, and so on.
3. Select $n_h \approx n/H$ units from each stratum with equal probabilities, where $n = \sum_{h=1}^H n_h$ is the desired total sample size. Inside each stratum U_h all units have inclusion probability $\pi_h = n_h/N_h$.

This procedure selects units with probability proportional to mean size in the stratum, in fact,

$$\sum_{i \in U_h} x_i = N_h \bar{x}_{U_h} = N \frac{\bar{x}_U}{H}$$

then

$$N_h = \frac{N \bar{x}_U}{H \bar{x}_{U_h}}$$

hence,

$$\pi_h = \frac{n_h}{N_h} = \frac{H^{-1} n}{H^{-1} N (\bar{x}_U / \bar{x}_{U_h})} = \frac{n \bar{x}_{U_h}}{N \bar{x}_U}.$$

2.3.2 *Sunter sampling*

Sunter (1977, 1986) proposed a list-sequential design (see section 2.4) where fixed size (property (1.3b)) is not respected and inclusion probabilities are only approximately proportional to the size variable.

Let us assume that, if any self-selecting unit is present, it has already been selected and removed from population. Then, Sunter (1977) sampling procedure is as follows:

1. Sort the units in decreasing (or increasing) order with respect to the size measure X ;
2. Define the set of self-selective units: $I = \{i : nx_i/X_i^* > 1; i = 1, \dots, N\}$, with $X_i^* = \sum_{j=i}^N x_j$ and set $i^* = \min\{\inf(I), N - n + 1\}$;
3. Select first unit with probability nx_i/X , and successive units with probability $(n - n_i)x_i/X_i^*$, where $n_i = 0, \dots, n$ is the number of units that are in the sample at the end of step $i - 1$ and X_i^* is defined as in the previous step. Stop when $i = i^*$ or $n_i = n$, whichever occurs first;
4. If, at this point, there are still units to be selected, that is $n_i < n$, draw a sample of size $n - n_i$ from the remaining units with index $i = i^*, \dots, N$ by simple random sampling.

Second-order inclusion probabilities can be expressed as (Särndal et al., 1992):

$$\pi_{ij} = \begin{cases} \frac{n(n-1)}{X} g_i x_i x_j & \text{for } i < j < i^* \\ \frac{n(n-1)}{X} g_{i^*-1} \frac{X_i^* - x_{i^*-1}}{X_{i^*} - \bar{x}_{i^*}} (\bar{x}_{i^*})^2 & \text{for } i^* \leq i \leq j \leq N \\ \frac{n(n-1)}{X} g_i x_i \bar{x}_{i^*} & \text{for } i < i^*, i^* \leq j \leq N \end{cases}$$

where

$$g_i = \left(1 - \frac{x_1}{X_2^*}\right) \left(1 - \frac{x_2}{X_3^*}\right) \dots \left(1 - \frac{x_{i-1}}{X_i^*}\right) / X_{i+1}^* = g_{i-1} \frac{X_i^* - x_{i-1}}{X_{i+1}^*}$$

and $\bar{x}_{i^*} = X_{i^*}^*/(N - i^* + 1)$ is the average size of the units in the equal probability subset.

This procedure satisfies the properties:

- $\pi_i = nx_i/X, \quad 1 \leq i < i^*$;
- $\pi_i = n\bar{x}/X, \quad i^* \leq i \leq N$;
- $\pi_{ij} > 0, \quad i < j < N$;
- $\pi_i \pi_j - \pi_{ij} > 0, \quad i < j \leq N$,

where $\bar{x} = (N - i^* + 1)^{-1} \sum_{i=i^*}^N x_i$ is the average size of the units in the equal probability subset, if any, as of step 4.

This sampling gives exact joint-inclusion probabilities which can be computed and stored simultaneously with the drawing of the sample. However, inclusion probabilities depend on the ordering of the list and the procedure is not an actual unequal probability sampling, as some of the units may be selected by a simple random sampling.

Sunter (1986) provided new sampling procedures, including one with fixed sample size that generalises the previous one and is an actual UPS designs. However, computing joint-inclusion probabilities under this new procedure is a hard task. This is because it is necessary to enumerate and sum the probabilities of all the appropriate event sequences, which is impractical for large sample and population sizes.

2.3.3 Poisson Sampling

Another sampling design with variable sample size is Poisson sampling (Hájek, 1981), which has an important role in official statistics, where it is largely used for its simplicity and attractive properties. It is quick and simple to perform and produces easy-to-compute joint-inclusion probabilities.

The sampling procedure consists in performing independent bernoullian trials for each population unit, where unit i is selected with probability π_i , given by equation (1.3a). Due to independence among units, joint-inclusion probabilities are simply obtained as $\pi_{ij} = \pi_i \pi_j$, $i \neq j$, this reduces the Horvitz-Thompson variance to

$$v_{HT}[\hat{\vartheta}_{HT}] = \sum_{i \in s} \frac{1 - \pi_i}{\pi_i^2} y_i^2,$$

which requires only a single sum over sample observations, so its estimation is extremely quick. Note that the Sen-Yates-Grundy variance estimator (see equation (1.10)) cannot be employed in this case as the sample size is not fixed.

Poisson sampling has random sample size $n(s)$, given by the sum of bernoullian trials with different probabilities, which generates a Poisson-binomial distribution (Tillé, 2006). From the mean of a Poisson-binomial distribution follows that the expected size of a Poisson sample is n :

$$E[n(s)] = \sum_{i \in U} \pi_i = \sum_{i \in U} n \frac{x_i}{\sum_{i \in U} x_i} = n.$$

Hájek (1981) shows that, by conditioning this design on a fixed-sample size n , one obtains a new sampling design, called Conditional Poisson Sampling (CPS) or Maximum Entropy sampling, for which Aires (1999), Chen et al. (1994), and Deville (2000) developed algorithms for the exact computation of second-order inclusion probabilities. Although Conditional Poisson Sampling is a strict π ps sampling, it is more complex than Poisson sampling and does not provide the same attractive properties. More details about this design are provided in section 2.6.2.

2.4 LIST-SEQUENTIAL SAMPLING DESIGNS

List-sequential sampling procedures are designs that skim through the population list unit by unit, and sequentially apply an inclusion rule to the units, until either a sample of size n is obtained or the list ends.

One of the first list-sequential procedures to be proposed was Chao sampling (Chao, 1982), which was inspired by Reservoir sampling (Knuth, 1969), employed in computer science and particularly convenient for sampling large files because it does not require population size. It could thus be performed in just one scan of the file. Besides this computational advantage, the statistical attractiveness of Chao sampling is due to the availability of a closed form equation for joint-inclusion probabilities with low computational complexity. In short, Chao sampling is composed by exactly $N - n$ steps. The algorithm initialises by filling the sample with the first n units of the list and, successively, at each step k , $k = n + 1, \dots, N$, unit k is selected with probability $\pi(k; k)$, where $\pi(k; k)$ is the inclusion probability of unit k at step k . If unit k is selected, it randomly replaces a sample unit, otherwise the procedure continues with the next step. Alongside with its simplicity, the easy computation of joint-inclusion probabilities and the possibility to sample lists of unknown size, this sampling procedure comes with some drawbacks. Inclusion probabilities depend on the order of population units, and most importantly second-order inclusion probabilities can be null. Bethlehem and Schuerhoff (1984) provide a necessary and sufficient condition for the inclusion probabilities to be strictly positive, namely that at most $n - 2$ self-selecting units are included in the sample at each step k , $k = n + 1, \dots, N$.

Although the joint-inclusion probabilities are easy to compute, they are memory-intensive due to the large number of probabilities that are to be computed. Berger (1998b) provided a variance estimator to be used under this design, which needs only the computation of N values.

Other list-sequential sampling designs are proposed in Sunter (1977, 1986). Despite not being a list-sequential sampling, it is worth to mention here Tillé's elimination procedure (Tillé, 1996), as *the calculation of the elimination probabilities was inspired by Chao (1982)*. It shares with Chao sampling the simplicity in its application and in the computation of second-order inclusion probabilities. However, while the sampling design previously described sequentially skims the population list to evaluate the selection of the units, Tillé sampling starts from a population of size N , and sequentially evaluates the exclusion of the units, until only n units remain. Contrary to Chao sampling, inclusion probabilities in Tillé's elimination procedure do not depend on the order of the list but the population size must be known. The algorithm produces easy-to-compute joint-inclusion probabilities, which

nonetheless may sometimes be null. The author provides a necessary and sufficient condition for strictly positive joint-inclusion probabilities.

Chao and Tillé sampling designs are described in details in the next subsections.

2.4.1 Chao sampling

The main idea of Chao sampling is to keep sample size fixed throughout the selection procedure and to consider one population unit at the time, letting it in the sample with probability proportional to size. Briefly, it works by selecting the first n units of the population to constitute an initial sample; then, one by one, the remaining units are either selected to be part of the sample or not, through a random scheme. If selected, the unit replaces at random one which is already in the sample.

Given an auxiliary variable X with values x_i , the selection procedure is as follows:

1. Form the initial sample by selecting the first n units in the list. Inclusion probabilities at this step will always be equal to one.
2. For $k = n + 1, \dots, N$:
 - a) For all $i = 1, \dots, k$, compute inclusion probabilities

$$\pi(k; i) = n \frac{x_i}{\sum_{i=1}^k x_i},$$

these probabilities are such that

- $\sum_{i=1}^k \pi(k; i) = n$,
- $\pi(k; i) \propto x_i$.

It may occur that $\pi(k; i) \geq 1$ for some i , in this case, we can proceed by setting these probabilities to 1, and recompute inclusion probabilities for the remaining units, as explained in section 1.2.2.

- b) Generate a random variable $u \sim U[0, 1]$.
- c) If $u < \pi(k, k)$, where $\pi(k; k)$ is the inclusion probability of the population unit taken into consideration at step k , then unit k is selected.
- d) If the unit is not selected, we ignore it and move on to next unit, otherwise, we remove one unit from the sample at random with probability R_{ki} (given below) and replace it with unit k .

First-order inclusion probabilities

It is possible to write inclusion probabilities as follows (Chao, 1982, Lemma 1):

$$\begin{cases} \pi(k; i) = (1 - W_k R_{ki}) \pi(k; i) & 1 \leq i \leq k; k > n. \\ \pi(k; k) = W_k \end{cases} \quad (2.8)$$

W_k is the probability of inclusion of unit X_{k+1} at time $k + 1$, while R_{ki} represents the conditional probability that unit i is removed and it is defined as:

$$R_{ki} = \begin{cases} 0 & (j \in A_k), \\ T_{ki} & (j \in B_k), \\ (1 - T_k)/(n - L_k) & (j \in C_k), \end{cases} \quad (2.9)$$

where A_k , B_k and C_k are sets defined as

$$\begin{aligned} A_k &= \{i : \pi(k-1; i) = \pi(k; i) = 1; i \leq k\}, \\ B_k &= \{i : \pi(k-1; i) = 1, \pi(k; i) < 1; i \leq k\}, \\ C_k &= \{i : \pi(k-1; i) < 1, \pi(k; i) < 1; i \leq k\}. \end{aligned}$$

Moreover, T_{ki} is the conditional probability of unit i being removed from the sample, given that unit $k + 1$ is selected:

$$T_{ki} = [1 - \pi(k; i)] / W_k$$

and $T_k = \sum_{i \in B_k} T_{ki}$ is the conditional probability of removing a unit which is in B_k . Finally, L_k is the number of elements which are either in set A_k or in set B_k .

Second-order inclusion probabilities

Chao gives the following lemma: for all $m \leq n$ and all $1 \leq i_1 < \dots < i_m \leq k$, $k \geq n$,

$$\pi(k+1; i_1, \dots, i_m) = \left(1 - W_k \sum_{j=1}^m R_{ki_j}\right) \pi(k; i_1, \dots, i_m)$$

and for $1 \leq i_1 < \dots < i_{m-1} \leq k$,

$$\pi(k+1; i_1, \dots, i_{m-1}, k+1) = W_k \left(1 - \sum_{j=1}^{m-1} R_{ki_j}\right) \pi(k; i_1, \dots, i_{m-1}).$$

Then, for $m = 2$ and $i < j$,

$$\pi(k+1; i, j) = \begin{cases} \pi(k; i, j) (1 - W_k (R_{ki} + R_{kj})) & , \quad j < k+1 \\ \pi(k; i) (1 - R_{ki}) W_k & , \quad j = k+1 \end{cases}$$

2.4.2 *Tillé sampling*

Given a population $U = \{1, 2, \dots, N\}$ of size N , we start with a sample of N units, and we sequentially remove one unit until reaching the desired sample size n . The procedure is composed of $N - n$ steps; at each step k , $k = N - 1, \dots, n$, first-order inclusion probabilities are computed, then one unit among the remaining in the sample is removed at random.

The algorithm is as follows:

1. At each step k , with $k = N - 1, \dots, n$, compute the first-order inclusion probabilities proportional to a size measure X for all units in the population U :

$$\pi(i|k) = \frac{kx_i}{\sum_{j \in U} x_j}, \quad U = \{1, \dots, N\}.$$

If there is any $\pi(i|k) \geq 1$, it is set equal to 1 and the computation is repeated on the remaining units, as described in section 1.2.2.

2. Select one unit from the current sample with probability

$$r_{ki} = 1 - \frac{\pi(i|k)}{\pi(i|k+1)}, \quad i \in U$$

and remove it (see Remarks 2 and 3 below).

The inclusion probabilities can be expressed as

$$\pi(i|k) = (1 - r_{ki}) \pi(i|k+1),$$

that is, the probability of inclusion of unit i at step k is given by its probability of inclusion at step $k+1$ (the previous one, as $k = N - 1, \dots, n$) times the probability of *not* being removed from sample at step k .

Joint-inclusion probabilities

The inclusion probabilities of the m -th order are given by

$$\pi(i_1, \dots, i_m|n) = \prod_{k=n}^{N-1} \left(1 - \sum_{j=1}^m r_{ki_j} \right),$$

that is, the joint probability of *not* removing any of the units i_1, \dots, i_m in any of the k steps. As a particular case, second-order inclusion probabilities are obtained as

$$\pi(i, j|n) = \prod_{k=n}^{N-1} (1 - r_{ki} - r_{kj})$$

and they are such that $\pi(i, j|n) \leq \pi(i|n) \pi(j|n)$, for any $i, j \in U$, $i \neq j$.

Unfortunately, the algorithm can sometimes produce null joint-inclusion probabilities, even if $\pi(i|n) < 1$ for all $i \in \mathcal{U}$. A necessary and sufficient condition for strictly positive probabilities is $r_{ki} + r_{kj} < 1$ for all $k = n, \dots, N-1$

Remarks

1. We have that

$$\begin{aligned} \pi(i|N) &= 1, \quad i \in \mathcal{U}; \\ \pi(i|k) &\leq \pi(i|k+1), \quad i \in \mathcal{U}, k = n, \dots, N-1; \\ \sum_{i \in \mathcal{U}} \pi(i|k) &= k, \quad k = 1, \dots, N. \end{aligned}$$

2. We have that $0 \leq r_{ki} < 1$ and that

$$r_{ki} = \begin{cases} 0 & \text{if } i \in A_k, \\ 1 - \pi(i|k) & \text{if } i \in B_k, \\ \frac{1 - \sum_{i \in B_k} [1 - \pi(i|k)]}{k+1 - \#A_k - \#B_k} & \text{if } i \in C_k. \end{cases}$$

with

$$\begin{aligned} A_k &= \{i : \pi(i|k) = 1\}, \\ B_k &= \{i : \pi(i|k) = 1 \text{ and } \pi(i|k+1) = 1\}, \\ C_k &= \{i : \pi(i|k+1) < 1\}. \end{aligned}$$

3. Once k is such that $\pi(i|k+1) < 1$ for all $i \in \mathcal{U}$, the units are eliminated with equal probabilities. In fact, in that case all units belong to C_k and, for all $i \in \mathcal{U}$

$$r_{ki} = \frac{1 - \sum_{i \in B_k} [1 - \pi(i|k)]}{k+1 - \#A_k - \#B_k} = \frac{1 - 0}{k+1 - 0 - 0} = \frac{1}{k+1}$$

2.5 BOOTSTRAP

Bootstrap is a technique originally introduced by Efron (1979) for infinite population inference and i.i.d. random variables. Given a random sample $s = \{y_1, \dots, y_n\}$ with i.i.d. units with unknown distribution F , for which we want to estimate the variance of parameter $\vartheta(F)$, the core idea of Efron's bootstrap is to estimate the F distribution by the empirical distribution function \hat{F} :

$$\hat{F}_n(z) = \frac{1}{n} \sum_{i=1}^n \delta_{y_i \leq z},$$

where $\delta_{y_i \leq z}$ is a Dirac measure, defined as

$$\delta_{y_i \leq z} = \begin{cases} 1, & y_i \leq z \\ 0, & \text{otherwise} \end{cases}$$

and to draw a large number of (bootstrap) samples s^* from \hat{F} , which is equivalent to sampling with equal probabilities and with replacement from the original sample s . For each bootstrap replicate the parameter of interest $\hat{\vartheta}(\hat{F}_b^*)$ is estimated, and an approximation of the variance is obtained as Monte Carlo variance of the bootstrap estimators.

2.5.1 Bootstrap for finite populations

The classic bootstrap cannot be transposed to the finite population setting in a straightforward way, as in this scenario inference is not carried out with respect to the distribution function of a Y variable (which is assumed deterministic), but rather on the random indicator variable δ_i (see section 1.1). Moreover, in this context classic bootstrap does not work, as resampling would destroy heteroskedasticity and correlation among sample units, which are not necessarily i.i.d. (Bertail and Combris, 1997).

Thus, successful application of the bootstrap in a finite population setting requires appropriate modifications. In UPS designs, bootstrap procedures are mostly employed to estimate the variance of the HT total estimator. To do so, a large number B of bootstrap samples is generated and the HT estimator is computed for each of them, then the bootstrap variance is estimated as

$$v^* = \frac{1}{B-1} \sum_{b=1}^B \left(\hat{\vartheta}_b^* - \hat{\vartheta}^* \right)^2,$$

where $\hat{\vartheta}^* = \frac{1}{B} \sum_{b=1}^B \hat{\vartheta}_b^*$.

Finite-population bootstrap methods are generally classified into three groups: *plug-in* or *pseudo-population* methods, *direct* or *ad-hoc* bootstrap methods, and *weights* methods. An extensive survey of available bootstrap methods for finite population sampling is presented in Mashreghi et al. (2016).

Pseudo-population methods

While in infinite population inference the data generating process is a distribution function F , in finite population inference this role is assumed by the population U and the sampling design. Hence, *pseudo-population* (or *plug-in*) bootstrap methods pursue the spirit of Efron's bootstrap by estimating the original population by means of sample units, from which samples are drawn according to the design that generated the original sample. By employing the original sampling scheme, the finite population correction factors are naturally captured by the bootstrap variance estimator (Mashreghi et al., 2016).

The pseudo-population U^* is generated by repeating sample units a certain number of times. Gross (1980) first applied this technique to simple random sampling by repeating each sample unit $f^{-1} = N/n$

times, with N/n integer. Bickel and Freedman (1984), Booth et al. (1994), Chao and Lo (1985, 1994), and Sitter (1992) extended Gross' bootstrap to cases with non-integer N/n . These methods are slightly biased except for Sitter's method (Mashreghi et al., 2016).

A similar procedure for unequal probability sampling was later developed by Holmberg (1998), who proposed the following algorithm:

1. Repeat $\lfloor \pi_i^{-1} \rfloor$ times the pair (y_i, π_i) to create U^f , where $\lfloor z \rfloor$ indicates the greatest integer less than or equal to z ;
2. Perform a Poisson sampling on the original sample, drawing units with probability $\pi^{-1} - \lfloor \pi^{-1} \rfloor$ and call this sample U^c . Add these units to U^f to complete the pseudo-population, obtaining $U^* = U^f \cup U^c$;
3. Denote with $(\check{y}_i, \check{\pi}_i)$, $i = 1, \dots, \text{card}(U^*)$, the pairs of observations and inclusion probabilities in population U^* , compute new inclusion probabilities

$$\pi'_i = n \frac{\check{\pi}_i}{\sum_{j \in U^*} \check{\pi}_j};$$

4. Draw a large number B of bootstrap samples s^* from U^* according to the original design. Per each iteration compute the HT estimator $\hat{\vartheta}_b^*$, then compute the bootstrap variance as

$$v^* = \frac{1}{B-1} \sum_{b=1}^B \left(\hat{\vartheta}_b^* - \hat{\vartheta}^* \right)^2,$$

$$\text{with } \hat{\vartheta}^* = \frac{1}{B} \sum_{b=1}^B \hat{\vartheta}_b^*;$$

5. Repeat steps 2-4 a large number of times D to get v_1^*, \dots, v_D^* ; estimate the variance by

$$v^* = \frac{1}{D} \sum_{d=1}^D v_d^*.$$

Note that U^c , the random part of the pseudo-population U^* , introduces a double form of randomness; indeed, it is random both in the units included and in its size, due to the use of Poisson sampling (see section 2.3.3). Note also that, in the original paper by Holmberg, the last step is not present; however, it is needed to account for the source of variability introduced by U^c (Chauvet, 2007). Chauvet (2007) proposed a modification to Holmberg's procedure valid only for Poisson sampling, which produces correct variance estimates. Chauvet's algorithm is as above, but at step 3 inclusion probabilities are not recomputed, that is $\pi'_i = \check{\pi}_i$.

Direct bootstrap

Direct methods, also called *ad-hoc* methods, draw bootstrap samples directly from the original sample, with no need of generating a pseudo-population, which leads to clear computational advantages in terms of memory usage and computational time. However, the sampling procedure needs some modifications in order to capture the variability of the original sample.

Examples of direct bootstrap methods for SRS and UPS designs are the Rescaling Bootstrap (RSB) procedures proposed by Rao and Wu (1988). Their methods manage to match the variance of the original design by rescaling sample observations and drawing bootstrap samples according to a simple random sampling with replacement.

Antal and Tillé (2011, 2014) provide *ad-hoc* bootstrap procedures that do not need to rescale the observations. Their implementation is complex, but the authors show by means of simulation studies that these methods have good performances even with some non-linear estimators.

Weights approach

These methods consist in generating bootstrap weights to attach to the observations in the original sample. Weights are expressed in the general form

$$w_i^* = \alpha_i^* w_i, \quad i \in s$$

where w_i are the original weights and α_i^* are some bootstrap adjustments. The bootstrap estimator is then obtained as

$$\hat{\vartheta}^* = \sum_{i \in s} w_i^* y_i. \quad (2.10)$$

This approach is efficient because it avoids the generation of a pseudo-population, and it is convenient for the release of public data both for users, who can use the provided bootstrap weights in their studies without effort, and for public agencies, that do not need to unveil details on their surveys, which could compromise the confidentiality of information.

Potentially, any bootstrap method could be treated as a weighted bootstrap, as long as a proper set of weights is found.

The concept of bootstrap weights was introduced by Rao, Wu, and Yue (1992) where, with regard to the Rescaling Bootstrap by Rao and Wu (1988), the authors point out that the rescaling could be performed on the weights rather than on the observed values. Chipperfield and Preston (2007) proposed a modification of this method, showing gains in efficiency with small samples.

An important method belonging to this class is the *generalised bootstrap*, proposed by Bertail and Combris (1997) and further improved

by Beaumont and Patak (2012), who proposed an alternative procedure, described below, which is independent from the variable of interest. The generalised bootstrap can be applied to any sampling design without the need of selecting bootstrap samples, and most bootstrap methods can be seen as special cases.

Let us consider the Horvitz–Thompson estimator of the true parameter ϑ , which can be expressed as $\hat{\vartheta} = \sum_{i \in s} w_i y_i$. For a positive integer p , let

$$m_p = E(\hat{\vartheta} - \vartheta)^p$$

indicate the p -th moment of the sampling error and let \hat{m}_p be any design-unbiased estimator of m_p , we shall consider the estimators

$$\begin{aligned} \hat{m}_1 &= 0 \\ \hat{m}_2 &= \sum_{i \in s} \sum_{j \in s} \sigma_{ij} \check{y}_i \check{y}_j = \check{\mathbf{y}}^T \boldsymbol{\Sigma} \check{\mathbf{y}}, \end{aligned}$$

where \hat{m}_1 is clear from the unbiasedness of the HT estimator, \hat{m}_2 is the Horvitz–Thompson variance estimator expressed in quadratic form, $\check{\mathbf{y}}$ is the vector whose elements are $\check{y}_i = w_i y_i$ and $\boldsymbol{\Sigma}$ is an $n \times n$ matrix of elements σ_{ij} , given by

$$\sigma_{ij} = \begin{cases} (\pi_{ij} - \pi_i \pi_j) / \pi_{ij}, & \text{for } i \neq j \\ (1 - \pi_i), & \text{for } i = j \end{cases}$$

Alternatively, if n is fixed, the Sen–Yates–Grundy estimator may be considered by letting

$$\sigma_{ij} = \begin{cases} (\pi_{ij} - \pi_i \pi_j) / \pi_{ij}, & \text{for } i \neq j \\ (1 - \pi_i) - \sum_{k \in s} (\pi_{ik} - \pi_i \pi_k) / \pi_{ik}, & \text{for } i = j \end{cases}$$

Moreover, consider a bootstrap estimator expressed as in equation (2.10); for the bootstrap method to be effective, the expectation of the bootstrap error should match the estimated sampling error \hat{m}_p , that is, the identity

$$E(\hat{\vartheta}^* - \hat{\vartheta})^p = \hat{m}_p$$

should be satisfied for at least $p = 2$.

The bootstrap error can be written as $\hat{\vartheta}^* - \hat{\vartheta} = (\mathbf{a}^* - \mathbf{1}_n)^T \check{\mathbf{y}}$, where \mathbf{a}^* is a vector whose elements are the adjustments a_i^* and $\mathbf{1}_n$ is a vector of length n whose elements are all 1. From the previous identity, we have that the generalised bootstrap must satisfy the conditions

$$E^*[\mathbf{a}^*] = \mathbf{1}_n, \quad (2.11)$$

$$E^*[(\mathbf{a}^* - \mathbf{1}_n)(\mathbf{a}^* - \mathbf{1}_n)^T] = \boldsymbol{\Sigma}. \quad (2.12)$$

Hence, the vector of adjustments \mathbf{a}^* can be generated from any distribution that satisfies the conditions above.

Table 2.1: Examples of generation of the adjustments α_i^* (Beaumont and Patak, 2012).

Distribution	Description
Normal	$\alpha_i^* \sim N(\mu = 1, \sigma^2 = 1 - \pi_i)$
Uniform	$\alpha_i^* \sim U(1 - \sqrt{3(1 - \pi_i)}, 1 + \sqrt{3(1 - \pi_i)})$
Exponential	$\alpha_i^* = 1 + (b_i - 1)\sqrt{1 - \pi_i}$, with $b_i \sim \text{Exp}(1)$
Lognormal	$\alpha_i^* = \exp(b_i)$, with $b_i \sim N(\mu = -0.5 \log(2 - \pi_i), \sigma^2 = \log(2 - \pi_i))$

Beaumont and Patak (2012) propose a list of candidate distributions, some of which are reported in table 2.1. Through a simulation study, the authors show that the distribution used for generating the adjustments α_i^* does not influence the outcome much, as long as the distribution of bootstrap errors is not too far from the Normal distribution. Thus, they suggest avoiding highly skewed distributions. Moreover, the authors provide a tool for the determination of a proper number of bootstrap replicates, and suggest to always use at least 750 replications in order to obtain reliable results.

2.6 APPROXIMATE VARIANCE ESTIMATORS

2.6.1 Entropy

We shall briefly discuss the concept of entropy (Shannon, 1948), as it will be useful in the next sections and throughout the thesis.

Using the same notation introduced in chapter 1, the entropy of a sampling design is defined as

$$H = - \sum_{s \in S} p(s) \log(p(s)) = - E_p[\log(p(s))] \quad (2.13)$$

and it represents a measure of how much the selection probability is *spread* among the samples of the sample space S . In other terms, entropy may be considered a measure of *randomness* of the sampling design (Grafström, 2010).

As explained in Grafström (2010), high-entropy does not guarantee high efficiency, however, the high randomisation provided by the high entropy protects against the lack of additional auxiliary information, or low proportionality between the auxiliary and the target variable.

Simple random sampling is the design with the highest entropy in the class of fixed-size, equal probability sampling methods, as it spreads the probability of selection over all samples in the sample space evenly. In the class of unequal probability sampling designs, Poisson sampling has the largest entropy, while the Conditional Poisson sampling has highest entropy among UPS designs with fixed size

n (Hájek, 1981), for this reason it is also called Maximum Entropy sampling. CPS design is described in the next subsection. Other designs are known to have high entropy, such as Randomised Systematic sampling and the Rao–Sampford design, however this information is not known for all designs. Nonetheless, it is important to note that as population and sample size increase, and so does the number of possible samples, the entropy gets closer to the maximum for any design.

For more information about entropy of unequal probability sampling designs see Grafström (2010) and Grafström and Lundström (2013).

2.6.2 Conditional Poisson Sampling

Hájek (1981) defined the Conditional Poisson Sampling (CPS) as a Poisson sampling conditioned on a sample size n . He proposed to draw samples by means of a rejective procedure, by which Poisson samples are drawn until one with the desired sample size is selected. For this reason, Hájek’s procedure is also referred to as the rejective sampling. However, there are other ways to implement this design. This design is also known as Maximum Entropy sampling, because it can be obtained by maximizing the entropy measure (2.13).

An algorithm for the computation of inclusion probabilities have been given by Chen et al. (1994), who found the link between conditional Poisson sampling and the exponential distribution family, for which there exists a one to one relationship between the parameter and the mean, that is, the inclusion probabilities. Chen’s algorithm was later improved by Deville (2000). A different algorithm was proposed by Aires (1999)

The Conditional Poisson sampling can be expressed as an exponential design (Tillé, 2006):

$$p(\mathbf{s}, \boldsymbol{\lambda}, \mathbf{S}_n) = \frac{\exp\{\boldsymbol{\lambda}^T \mathbf{s}\}}{\sum_{\mathbf{s} \in \mathbf{S}_n} \exp\{\boldsymbol{\lambda}^T \mathbf{s}\}},$$

where \mathbf{S}_n is the set of samples with size n , $\boldsymbol{\lambda}$ is the parameter and \mathbf{s} is a vector of N elements such that $s_i = 1$ if unit i is in the sample, and 0 otherwise.

Given a set of first-order inclusion probabilities, Chen et al. (1994) provided an algorithm to derive the parameter $\boldsymbol{\lambda}$, which can then be used to estimate joint inclusion probabilities. Deville (2000) developed a fast recursive equation for π_{ij} :

$$\pi_{ij}(\boldsymbol{\lambda}, \mathbf{S}_n) = \frac{n(n-1) \exp\{\lambda_i\} \exp\{\lambda_j\} [1 - \pi_i(\boldsymbol{\lambda}, \mathbf{S}_{n-2}) - \pi_j(\boldsymbol{\lambda}, \mathbf{S}_{n-2}) + \pi_{ij}(\boldsymbol{\lambda}, \mathbf{S}_{n-2})]}{\sum_{i \in U} \sum_{j \neq i} \exp\{\lambda_i\} \exp\{\lambda_j\} [1 - \pi_i(\boldsymbol{\lambda}, \mathbf{S}_{n-2}) - \pi_j(\boldsymbol{\lambda}, \mathbf{S}_{n-2}) + \pi_{ij}(\boldsymbol{\lambda}, \mathbf{S}_{n-2})]},$$

with $\pi_{ij}(\boldsymbol{\lambda}, 0) = 0$ and $\pi_{ij}(\boldsymbol{\lambda}, 1) = 0$, $\forall i, j \in U, i \neq j$. For a detailed description of Conditional Poisson sampling see Tillé (2006).

2.6.3 *Approximate estimators*

A number of approximations to the joint-inclusion probabilities have been proposed for high-entropy designs. They are easy to compute and can be plugged in the Horvitz–Thompson or the Sen–Yates–Grundy estimator (see equations (1.9) and (1.10), respectively) to obtain variance estimates.

Hartley and Rao (1962) provided an approximation of the π_{ij} of order $O(N^{-4})$ for the Randomised Systematic sampling, assuming $N \rightarrow \infty$ and n fixed:

$$\begin{aligned} \tilde{\pi}_{ij} = & \frac{n-1}{n} \pi_i \pi_j + \frac{n-1}{n^2} (\pi_i^2 \pi_j + \pi_i \pi_j^2) - \frac{n-1}{n^3} \pi_i \pi_j \sum_{i \in U} \pi_j^2 + \\ & + \frac{2(n-1)}{n^3} (\pi_i^3 \pi_j + \pi_i \pi_j^3 + \pi_i^2 \pi_j^2) - \frac{3(n-1)}{n^4} (\pi_i^2 \pi_j + \pi_i \pi_j^2) \sum_{i \in U} \pi_i^2 + \\ & + \frac{3(n-1)}{n^5} \pi_i \pi_j \left(\sum_{i \in U} \pi_i^2 \right)^2 - \frac{2(n-1)}{n^4} \pi_i \pi_j \sum_{i \in U} \pi_j^3. \end{aligned}$$

Asok and Sukhatme (1976) found a similar approximation for Rao–Sampford sampling of order $O(n^3 N^{-3})$, given by

$$\begin{aligned} \tilde{\pi}_{ij} = & \frac{1}{2} \pi_i \pi_j (c_i + c_j), \\ \text{where } c_i = & \frac{n-1}{n} \left[1 - \frac{1}{n^2} \sum_{j \in U} \pi_j^2 + \frac{2}{n} \pi_i \right], \end{aligned}$$

which is equivalent to the first three terms of Hartley and Rao’s approximation.

Hájek (1964) used less restrictive conditions to derive an approximation under CPS sampling. This approximation holds for

$$d = \sum_{i \in U} \pi_i (1 - \pi_i) \rightarrow \infty,$$

which implies that $N \rightarrow \infty$ and $N - n \rightarrow \infty$. Hence, contrary to Hartley and Rao (1962) approximation, this one does not require the sample size to be small compared to the population. Hájek’s approximation is given by

$$\tilde{\pi}_{ij} = \pi_i \pi_j \left[1 - (1 - \pi_i)(1 - \pi_j) d^{-1} \right]. \quad (2.14)$$

Berger (1998a) gave more general conditions for Hájek’s approximation and showed that it can be used with a larger class of high-entropy sampling designs, such as the Rao–Sampford procedure.

Brewer and Donadio (2003) develop a series of approximations based on properties (1.4b) to (1.4d) and the approximation $\pi_{ij} \approx \pi_i \pi_j$, which holds for several high-entropy designs, provided that N and n

are large enough. Such approximations are represented in the general form

$$\tilde{\pi}_{ij} = \pi_i \pi_j (c_i + c_j)/2,$$

where possible choices for c_i are

$$\begin{aligned} c_i &= \frac{n-1}{n-\pi_i}, \\ c_i &= \frac{n-1}{n - \frac{1}{n} \sum_{j \in U} \pi_j^2}, \\ c_i &= \frac{n-1}{n - 2\pi_i + \frac{1}{n} \sum_{j \in U} \pi_j^2}, \\ c_i &= \frac{n-1}{n - \frac{2n-1}{n-1} \pi_i + \frac{1}{n-1} \sum_{j \in U} \pi_j^2}. \end{aligned}$$

Approximating the joint-inclusion probabilities and using them with the HT or SYG estimators gives easy-to-compute variance estimates. However, one still has to cope with the double summation of such estimators, which may be computationally cumbersome. For this reason, several approximate variance estimators have been proposed that only require a single sum and are function of the first-order inclusion probabilities only.

Based on the quantities that variance estimators require to be computed, Matei and Tillé (2005) divided them into three classes:

1. Estimators that require both first and second-order inclusion probabilities;
2. Estimators that use only first-order inclusion probabilities and only for sample units;
3. Estimators that need only first-order inclusion probabilities, for the entire population.

The first class is composed of the Horvitz-Thompson and the Sen-Yates-Grundy estimators, while the second and third classes include approximate variance estimators, which are listed below. The estimator nomenclature (e.g. Brewer estimator 1, Deville estimator 1, ...) have been chosen to reflect the one used in Matei and Tillé (2005).

Haziza et al. (2008) provide a common form to express most of the estimators in classes 2 and 3:

$$v(\hat{\vartheta}_{HT}) = \sum_{i \in s} c_i e_i^2,$$

where $e_i = \frac{y_i}{\pi_i} - \hat{B}$, with

$$\hat{B} = \frac{\sum_{i \in s} a_i (y_i / \pi_i)}{\sum_{i \in s} a_i}$$

and a_i and c_i are parameters that define different estimators:

- Estimators of class 2:

- Hájek estimator (v_H) [Hájek (1964)]

$$c_i = \frac{n}{n-1}(1 - \pi_i); \quad \alpha_i = c_i$$

- Deville estimator 2 (v_{D2}) [see Deville (1993, 1999)]

$$c_i = (1 - \pi_i) \left\{ 1 - \sum_{j \in s} \left[\frac{1 - \pi_j}{\sum_{k \in s} (1 - \pi_k)} \right]^2 \right\}^{-1}; \quad \alpha_i = c_i$$

- Deville estimator 3 (v_{D3}) [Deville (1993)]

$$c_i = (1 - \pi_i) \left\{ 1 - \sum_{j \in s} \left[\frac{1 - \pi_j}{\sum_{k \in s} (1 - \pi_k)} \right]^2 \right\}^{-1}; \quad \alpha_i = 1$$

- Rosén estimator (v_R) [Rosén (1991)]

$$c_i = \frac{n}{n-1}(1 - \pi_i); \quad \alpha_i = (1 - \pi_i) \log(1 - \pi_i) / \pi_i$$

- Brewer estimator 1 (v_{B1}) [see Brewer (2002) and Brewer and Donadio (2003)]

$$c_i = \frac{n}{n-1}(1 - \pi_i); \quad \alpha_i = 1$$

- Estimators of class 3:

- Brewer estimator 2 (v_{B2}) [see Brewer (2002) and Brewer and Donadio (2003)]

$$c_i = \frac{n}{n-1} \left(1 - \pi_i + \frac{\pi_i}{n} - n^{-2} \sum_{j \in U} \pi_j^2 \right); \quad \alpha_i = 1$$

- Brewer estimator 3 (v_{B3}) [see Brewer (2002) and Brewer and Donadio (2003)]

$$c_i = \frac{n}{n-1} \left(1 - \pi_i - \frac{\pi_i}{n} - n^{-2} \sum_{j \in U} \pi_j^2 \right); \quad \alpha_i = 1$$

- Brewer estimator 4 (v_{B4}) [see Brewer (2002) and Brewer and Donadio (2003)]

$$c_i = \frac{n}{n-1} \left(1 - \pi_i - \frac{\pi_i}{n-1} + n^{-1}(n-1)^{-1} \sum_{j \in U} \pi_j^2 \right); \quad \alpha_i = 1$$

- Berger estimator (v_{Be}) [Berger (2004)]

$$c_i = (1 - \pi_i) \frac{n}{n-1} \left[\frac{\sum_{j \in s} (1 - \pi_j)}{\sum_{j \in U} (1 - \pi_j)} \right]; \quad \alpha_i = c_i$$

– Hartley–Rao estimator (v_{HR}) [Hartley and Rao (1962)]

$$c_i = \frac{n}{n-1} \left(1 - \pi_i - n^{-1} \sum_{j \in s} \pi_i + n^{-1} \sum_{j \in U} \pi_j^2 \right); \quad a_i = 1$$

Additional estimators are defined in Matei and Tillé (2005):

- Deville estimator $\bar{1}$ [Deville (1993)]

$$v_{D1}(\hat{\vartheta}_{HT}) = \sum_{i \in s} \frac{c_i}{\pi_i^2} (y_i - y_i^*)^2$$

where

$$y_i^* = \pi_i \frac{\sum_{j \in s} c_j y_j / \pi_j}{\sum_{j \in s} c_j}$$

and $c_i = (1 - \pi_i) \frac{n}{n-1}$

- Tillé estimator [Tillé (1996)]

$$v_T(\hat{\vartheta}_{HT}) = \left(\sum_{i \in s} \omega_i \right) \sum_{i \in s} \omega_i (\tilde{y}_i - \bar{\tilde{y}}_\omega)^2 - n \sum_{i \in s} \left(\tilde{y}_i - \frac{\hat{t}_{HT}}{n} \right)^2$$

where $\tilde{y}_i = y_i / \pi_i$, $\omega_i = \pi_i / \beta_i$ and $\bar{\tilde{y}}_\omega = \left(\sum_{i \in s} \omega_i \right)^{-1} \sum_{i \in s} \omega_i \tilde{y}_i$

The coefficients β_i are computed iteratively according to the following algorithm:

1. $\beta_i^{(0)} = \pi_i, \forall i \in U$
 2. $\beta_i^{(2k-1)} = \frac{(n-1)\pi_i}{\beta_i^{(2k-2)} - \beta_i^{(2k-2)}}$
 3. $\beta_i^{2k} = \beta_i^{(2k-1)} \left(\frac{n(n-1)}{(\beta_i^{(2k-1)})^2 - \sum_{i \in U} (\beta_i^{(2k-1)})^2} \right)^{(1/2)}$
- with $\beta^{(k)} = \sum_{i \in U} \beta_i^k, k = 1, 2, 3, \dots$

- Fixed–Point estimator [see Deville and Tillé (2005)]

$$v_{FP}(\hat{\vartheta}_{HT}) = \sum_{i \in s} \frac{y_i^2}{\pi_i^2} \left(c_i - \frac{c_i^2}{\sum_{j \in s} c_j} \right) - \frac{1}{\sum_{j \in s} c_j} \sum_{i \in s} \sum_{\substack{j \in s \\ j \neq i}} \frac{y_i y_j c_i c_j}{\pi_i \pi_j}$$

where the coefficients c_i are obtained using the following recurrence equation, until convergence is reached:

$$c_i^{(k)} = \frac{\left(c_i^{(k-1)} \right)^2}{\sum_{j \in s} c_j^{(k-1)}} + (1 - \pi_i)$$

with the initialisation

$$c_i^{(0)} = (1 - \pi_i) \frac{n}{n-1}$$

a necessary condition for the existence of a solution is

$$\frac{1 - \pi_i}{\sum_{j \in s} (1 - \pi_j)} < \frac{1}{2}, \quad \forall i \in s$$

and, if not satisfied, one may consider a variant that uses only one iteration:

$$c_i^{(1)} = (1 - \pi_i) \left(\frac{n(1 - \pi_i)}{(n-1) \sum_{j \in s} (1 - \pi_j)} + 1 \right)$$

- Matei–Tillé estimator 1 [Matei and Tillé (2005)]

$$v_{MT1}(\hat{\vartheta}_{HT}) = \frac{n(N-1)}{N(n-1)} \sum_{i \in s} \frac{b_i}{\pi_i^3} (y_i - \hat{y}_i^*)^2$$

where

$$\hat{y}_i^* = \pi_i \frac{\sum_{i \in s} b_i y_i / \pi_i^2}{\sum_{i \in s} b_i / \pi_i}$$

and the coefficients b_i are computed iteratively by the algorithm:

1. $b_i^{(0)} = \pi_i(1 - \pi_i) \frac{N}{N-1}, \forall i \in U$
2. $b_i^{(k)} = \frac{(b_i^{(k-1)})^2}{\sum_{j \in U} b_j^{(k-1)}} + \pi_i(1 - \pi_i)$

a necessary condition for convergence is

$$\frac{\pi_i(1 - \pi_i)}{\sum_{j \in U} \pi_j(1 - \pi_j)} < \frac{1}{2}, \quad \forall i \in U$$

and, if not satisfied, an alternative solution that uses only one iteration is provided:

$$b_i = \pi_i(1 - \pi_i) \left(\frac{N\pi_i(1 - \pi_i)}{(N-1) \sum_{j \in U} \pi_j(1 - \pi_j)} + 1 \right)$$

- Matei–Tillé estimator 2 [Matei and Tillé (2005)]

$$v_{MT2}(\hat{\vartheta}_{HT}) = \frac{1}{1 - \sum_{i \in U} \frac{d_i^2}{\pi_i}} \sum_{i \in s} (1 - \pi_i) \left(\frac{y_i}{\pi_i} - \frac{\hat{\vartheta}_{HT}}{n} \right)^2$$

where

$$d_i = \frac{\pi_i(1 - \pi_i)}{\sum_{j \in U} \pi_j(1 - \pi_j)}$$

- Matei–Tillé estimator 3 [Matei and Tillé (2005)]

$$v_{MT3}(\hat{\vartheta}_{HT}) = \frac{1}{1 - \sum_{i \in U} \frac{d_i^2}{\pi_i}} \sum_{i \in s} (1 - \pi_i) \left(\frac{y_i}{\pi_i} - \frac{\sum_{j \in s} (1 - \pi_j) \frac{y_j}{\pi_j}}{\sum_{j \in s} (1 - \pi_j)} \right)^2$$

where d_i is defined as in the previous estimator.

- Matei–Tillé estimator 4 (v_{MT4}) [Matei and Tillé (2005)]

$$v_{MT4}(\hat{\vartheta}_{HT}) = \frac{1}{1 - \frac{1}{n^2} \sum_{i \in U} b_i} \sum_{i \in s} \frac{b_i}{\pi_i^3} (y_i - y_i^*)^2$$

where

$$y_i^* = \pi_i \frac{\sum_{j \in s} b_j y_j / \pi_j^2}{\sum_{j \in s} b_j / \pi_j}$$

and

$$b_i = \frac{\pi_i(1 - \pi_i)N}{N - 1}$$

- Matei–Tillé estimator 5 (v_{MT5}) [Matei and Tillé (2005)]

This estimator is defined as estimator v_{MT4} , and the b_i values are obtained as in v_{MT1} .

2.6.4 Performances

The performances of the previous estimators were analysed in Brewer and Donadio (2003), Berger (2004), Matei and Tillé (2005) and Haziza et al. (2008).

Brewer and Donadio (2003) studied the performances of estimators of the Brewer family (v_{B1} to v_{B4}) plus estimators v_H and v_{D2} . They used Brewer and Tillé sampling designs for $n = 2$ with nine small populations and Randomised Systematic and Tillé designs for $n > 2$ with two real population with $N = 220$ and $N = 281$. Results were compared in terms of Relative Bias and Stability, measured through Coefficient of Variation (CV). In all scenarios considered, the estimators had low bias and, in particular, when $n = 2$ Brewer estimator 4 had a considerably lower bias than the other estimators. No meaningful differences were found among estimators for what concerns stability. In addition, the authors found that relative bias was always positive for Tillé sampling, and higher than with Randomised Systematic design. Moreover, variances were lower for the former than the latter. Considering these results, the authors conjectured that Tillé sampling has lower entropy than Random Systematic sampling.

Berger (2004) analysed his estimator v_{Be} under Chao (1982) sampling for an artificial stratified population with $N = 7000$, comparing

it to the Sen–Yates–Grundy estimator and Brewer estimator v_{BE} with respect to relative bias and stability, measured by root mean square error (RMSE). The results showed that v_{BE} had a relatively low bias, comparable with that of Brewer estimator v_B , and that the RMSE of both approximate estimators was only slightly higher than that of the SYG estimator. The author also performed simulations considering only v_H under Systematic sampling, which has low–entropy, in which situation the estimator performed poorly.

Matei and Tillé (2005) evaluated the behaviour of all estimators listed above in terms of Relative Bias (RB), Mean Squared Error (MSE) and Relative Coverage (RC), under Conditional Poisson Sampling and for three population: a real one, a simulated one with high correlation between the variable of interest Y and the auxiliary variable X , and a simulated one with $\rho_{XY} = -0.4$. In their simulations, estimators v_{FP} , v_R , v_{MT1} and v_{MT3} generally returned the lowest Relative Bias, but not the best MSE. The authors also analysed the performances of the Horvitz–Thompson estimator and concluded that v_{HT} has worse performances than other estimators in terms of MSE and RC when X and Y are strongly correlated, while it has similar performances when the correlation is low. No big differences were found among the SYG and other estimators. The authors concluded that approximate variance estimators have similar performances among each other, and that estimators that require only first–order inclusion probabilities for sample units are as accurate as those that require the π_i for all population units, or even those that demand the knowledge of joint–inclusion probabilities. Thus, they recommended the use of a simple estimator, such as v_{D1} .

As Henderson (2006) pointed out, there was a large discrepancy between the results obtained in Brewer and Donadio (2003) and Matei and Tillé (2005) regarding the estimators belonging to Brewer’s family (v_{B1} to v_{B4}), for which the former found negligible values for the bias, while the latter reported very high RB. Nonetheless, results from simulations carried out by Henderson (2006) and Haziza et al. (2008) under the same settings agree with those of Brewer and Donadio (2003). For a detailed discussion about this topic see Henderson (2006).

Haziza et al. (2008) analysed estimators v_H , v_{Be} , v_R , v_{D2} , v_{D3} , v_R , v_{B1} , v_{B2} , v_{B3} and v_{B4} under Rao–Sampford sampling design for several populations: 10 real populations with different size, coefficient of variation of X and Y , and correlation between X and Y , and 24 simulated populations with different generating model, N and $CV(X)$. Performances were assessed in terms of Relative Bias, MSE and Relative Stability, computed as the ratio of the MSE of a given estimator over that of the SYG estimator. The authors concluded that when $CV(X) < 0.5$, all such estimators perform well in terms of RB. In all situations considered, estimators v_{D2} , v_{B3} and v_{B4} returned a Relative Bias not larger than 5%. When n , N are not large, however, these esti-

mators did not perform well, and this was especially true when $CV(X)$ was high. The authors explain this behaviour by the fact that the Rao–Sampford design has high entropy only when population and sample size are large. On the other hand, for large values of n and N all estimators performed well. Moreover, approximate estimators showed better or same stability with respect to the Sen–Yates–Grundy estimator, v_{SYG} , over all scenarios. Overall, approximate estimators performed as well as or better than v_{SYG} . However, the authors warn about the use of approximate variance estimators when stratification is applied. Indeed, being these estimators biased, the sum of biased estimates over strata may lead to a considerable bias if the biases are in the same direction, even if each of them is small.

Henderson (2006) proved the following theorem

Theorem 1. *For any given sampling algorithm and any given set of first order inclusion probabilities:*

$$\begin{aligned} v_{B_1} &\geq v_{D_1}, \\ v_{D_3} &\geq v_{B_1}, \\ v_{D_3} &\geq v_{D_2}, \end{aligned}$$

where equality holds if and only if (i) inclusion probabilities π_i are all equal, or (ii) $\pi_i = \left(n x_i / \sum_{i \in U} x_i \right)$ and x_i is exactly proportional to y_i , $\forall i \in U$.

From this theorem the following corollary is derived:

Corollary 1. *For any given sampling algorithm the following is true for all simulations where the variances are estimated for the same samples:*

$$\begin{aligned} RB(v_{B_1}) &\geq RB(v_{D_1}), \\ RB(v_{D_3}) &\geq RB(v_{B_1}), \\ RB(v_{D_3}) &\geq RB(v_{D_1}), \\ RB(v_{D_3}) &\geq RB(v_{D_2}), \end{aligned}$$

where RB indicates the Relative Bias.

2.6.5 Conclusions

The empirical studies available in literature suggest that the knowledge of joint-inclusion probabilities is not necessary for estimating the variance of the HT total estimator, as the approximate estimators appear to perform well both in terms of bias and MSE, as long as a high-entropy design is employed. Although the behaviour of such estimators has not been broadly explored with regard to small populations and low-entropy designs, the information available suggests that under these scenarios approximate estimators have larger bias.

There is no consistently better estimator and, as pointed out by Henderson (2006), usually exists a trade-off between Relative Bias and

Mean Squared Error, so that the choice of the more appropriate estimator should be done on the basis of the need of the study.

It is also important to note that, while in the HT and SYG estimators self-selecting units do not contribute to final estimate, this is not true for all approximate estimators, thus computation of these estimators should be performed without taking into account self-selecting units.

2.7 MONTE CARLO APPROACH

Another effective approach to the approximation of inclusion probabilities is represented by Monte Carlo simulation (Fattorini, 2006). Estimation is carried out by performing K independent random draws of a sample from the target population according to the sampling design that generated the original sample. First-order inclusion probabilities may then be approximated by the proportion of occurrences of unit i in the Monte Carlo sample replicates:

$$\tilde{\pi}_i = \frac{W_i + 1}{K + 1}$$

and joint-inclusion probabilities are given by the proportion of occurrences of the couple (i, j) over the Monte Carlo replicates:

$$\tilde{\pi}_{ij} = \frac{W_{ij} + 1}{K + 1},$$

where W_i and W_{ij} are, respectively, the number of occurrences of unit i and of couple (i, j) over the K replicates. Both numerator and denominator are incremented by one unit to assure strictly positive estimates.

Considering the estimation of both π_i and π_{ij} through Monte Carlo approximations $\tilde{\pi}_i$ and $\tilde{\pi}_{ij}$, by replacing these quantities in the Horvitz-Thompson total and variance estimators, one obtains:

$$\hat{\vartheta}_{MC} = \sum_{i \in s} \frac{y_i}{\tilde{\pi}_i} \quad (2.18)$$

and

$$v_{MC}[\hat{\vartheta}_{HT}] = \sum_{i \in s} \frac{1 - \tilde{\pi}_i}{\tilde{\pi}_i^2} y_i^2 + 2 \sum_{i \in s} \sum_{j > i} \frac{(\tilde{\pi}_{ij} - \tilde{\pi}_i \tilde{\pi}_j)}{\tilde{\pi}_i \tilde{\pi}_j \tilde{\pi}_{ij}} y_i y_j. \quad (2.19)$$

Taking into account that W_i is a binomial random variable with probability π_i , $W_i \sim \text{Bin}(K, \pi_i)$, Fattorini (2006) found that the expectation and variance of $\hat{\vartheta}_{MC}$ are, respectively:

$$\begin{aligned}
E[\hat{\vartheta}_{MC}] &= \vartheta - \sum_{i \in U} y_i (1 - \pi_i)^{K+1}, \\
\text{Var}[\hat{\vartheta}_{MC}] &= \sum_{i \in U} y_i^2 \left\{ \pi_i E \left[\left(\frac{K+1}{W_i+1} \right)^2 \right] - \pi_i^2 \left(E \left[\frac{K+1}{W_i+1} \right] \right)^2 \right\} \\
&\quad + 2 \sum_{i \in U} \sum_{j > i} y_i y_j \left\{ \pi_{ij} E \left[\frac{(K+1)^2}{(W_i+1)(W_j+1)} \right] - \pi_i \pi_j E \left[\frac{K+1}{W_i+1} \right] E \left[\frac{K+1}{W_j+1} \right] \right\}
\end{aligned}$$

and proved that the Monte Carlo total (2.18) and variance (2.19) are asymptotically unbiased, as the number of Monte Carlo replicates increases. Indeed, he showed that

$$\frac{|E[\hat{\vartheta}_{MC}] - \vartheta|}{\vartheta} \leq \frac{1}{(K+2)\pi_0},$$

where $\pi_0 = \min\{\pi_i\}$, and that

$$\frac{|E[v_{MC}] - \text{Var}[\hat{\vartheta}_{MC}]|}{\text{Var}[\hat{\vartheta}_{MC}]} \leq \frac{|E[v_{MC}] - \text{Var}[\hat{\vartheta}_{MC}]|}{\text{Var}[\hat{\vartheta}_{HT}]} \leq \frac{1}{(K+2) \text{CV}[\hat{\vartheta}_{HT}]^2 \pi_{00}}, \quad (2.20)$$

where $\pi_{00} = \min\{\pi_{ij}\}$ and $\text{CV}[\hat{\vartheta}_{HT}] = \{\text{Var}_{HT}[\hat{\vartheta}_{HT}]\}^{1/2} / \vartheta$.

Moreover, he proved that the MSE of the Monte Carlo total estimator $\hat{\vartheta}_{MC}$ converges to the variance of the HT estimator as K increases:

$$\frac{|\text{MSE}[\hat{\vartheta}_{MC}] - \text{Var}[\hat{\vartheta}_{HT}]|}{\text{Var}[\hat{\vartheta}_{HT}]} \leq \frac{9}{(K+2)\pi_0} \left[1 + \frac{1}{\text{CV}[\hat{\vartheta}_{HT}]^2} \right]. \quad (2.21)$$

Despite being more computationally demanding, this method is extremely flexible and can be employed even with designs for which first-order inclusion probabilities cannot be computed (Barabesi et al., 1997; Fattorini and Ridolfi, 1997) or when units are substituted due to nonresponse (Thompson and Wu, 2008).

VARIANCE APPROXIMATIONS: A COMPARISON THROUGH A SIMULATION STUDY

In section 2.6 the approximate variance estimators were introduced, which allow for simple estimation of the variance of the Horvitz–Thompson total by only requiring first–order inclusion probabilities. Various authors, such as Matei and Tillé (2005) and Haziza et al. (2008) have showed that, overall, these estimators have a reasonably low bias and MSE when used with high–entropy designs and large populations. On the other hand, although theory and empirical studies suggest that these estimators are affected by a larger bias when employed with small populations and low–entropy designs, their behaviour has not been broadly explored under these scenarios.

The Monte Carlo method, described in section 2.7, is another approach to variance approximation, which consists in replacing the π_{ij} in the HT and SYG variance estimators with approximations obtained through Monte Carlo simulation. It is computer–intensive, but more flexible than approximate estimators. Indeed, this approach can be employed under any sampling design and with any population size, as long as an adequate number of replicates is generated. Moreover, the Monte Carlo method can be used even with designs for which first–order inclusion probabilities are not available (Barabesi et al., 1997; Fattorini and Ridolfi, 1997) or when units are substituted due to nonresponse (Thompson and Wu, 2008). Although the effectiveness of this method has not been explored widely, results by Fattorini (2006) guarantee convergence as the number K of Monte Carlo replications increases. Considering that larger populations tend to spread more the inclusion probabilities among units, thus generating lower values of π_o , the upper bound in equation (2.21) suggests that, for a given K , this approach may converge more quickly for relatively small populations.

Considering the characteristics of these two approaches, it seems reasonable to consider them as complementary, rather than alternative, to each other. In particular, the Monte Carlo method might be preferable with smaller populations and designs with lower entropy, while approximate variance estimators may be a better choice with high–entropy designs. In this chapter, the behaviour of the two approaches is analysed and the hypothesis above is studied by means of an extensive simulation study. The simulation set–up will be described in section 3.1, and results will be discussed in section 3.2.

3.1 METHODOLOGY

In this section, the characteristics of the simulation study that we carried out to evaluate and compare the behaviour of the approximate variance estimators and Monte Carlo estimators will be described.

Estimates have been computed under a large number of scenarios, in an attempt to cover as broadly as possible the situations that may occur in a real study. Samples were drawn from four different sampling designs, namely the Maximum Entropy, Randomised Systematic, Tillé (1996) and Brewer (1975) sampling designs.

Two real populations were considered, the first is the small population by Sukhatme (1954, p. 183) with data on cultivated areas in $N = 34$ villages in Lucknow subdivision (India), where the target variable is the amount of area under wheat in 1937 (W_{37} , in acres), and the size measure is the area under wheat in 1936 (W_{36} , in acres). The second population is obtained by the 284 belgian municipalities dataset (MU_{284}) by Särndal et al. (1992, appendix B), taking the 1985 population (P_{85} , in thousands) as target variable and variable P_{75} , representing the 1975 population (in thousands), as size measure. The 11 units with highest value of the size measure were removed to avoid self-selecting units, so the actual population size was $N = 273$ and in the sequel we will refer to this population as MU_{273} . These information have been summarised in table 3.1.

Table 3.1: Summary of population characteristics.

Population	N	Y	Y
Sukhatme	34	W_{37} : amount of area under wheat in 1937 (in acres)	W_{36} : amount of area under wheat in 1936 (in acres)
MU_{273}	273	P_{85} : 1985 population (in thousands)	P_{75} : 1975 population (in thousands)

By using the *population space* approach by Stehman and Overton (1994), both populations were transformed to obtain new populations with different coefficients of variation, and thus explore a wider range of scenarios. Both the target variable Y and the size measure X were transformed by using the relation

$$Z' = \left| Z + (c \text{sd}(Z) - \bar{Z}) \right|,$$

where Z' is the transformed variable, Z is the variable to be transformed, c is a constant indicating the approximate CV of the new variable Z' , $\text{sd}(Z)$ is the standard deviation of Z and \bar{Z} is its mean. For both populations, a population space was obtained by generating X and Y variables with coefficients of variations that are equal to the

Table 3.2: Correlation between X and Y for the Sukhatme population space, by coefficients of variation of Y and X.

		CV(Y)				
		1.05	0.95	0.76	0.67	0.5
CV(X)	1.09	0.90	0.86	0.84	0.84	0.84
	0.96	0.86	0.93	0.92	0.92	0.92
	0.77	0.84	0.92	0.93	0.93	0.93
	0.67	0.84	0.92	0.93	0.93	0.93
	0.5	0.84	0.92	0.93	0.93	0.93

Table 3.3: Correlation between X and Y for the MU273 population space, by coefficients of variation of Y and X.

		CV(Y)				
		1.05	0.95	0.76	0.67	0.5
CV(X)	1.27	0.99	0.96	0.96	0.96	0.96
	1	0.95	0.99	0.99	0.99	0.99
	0.81	0.95	0.99	0.99	0.99	0.99
	0.67	0.95	0.99	0.99	0.99	0.99
	0.5	0.95	0.99	0.99	0.99	0.99

values reported in table 3.4; each of the generated population spaces included 25 populations. Tables 3.2 and 3.3 show the correlation between the target variable Y and the size measure X for all combinations of CV(X) and CV(Y) for the Sukhatme and MU273 population spaces, respectively.

Samples were drawn with fixed size and different sampling fractions f . The size of the samples generated was 1%, 5%, 10%, 15% and 20% of the population size for the MU273 population space, and 5%, 10%, 15%, 20% for Sukhatme population.

The simulation was carried out by generating $K = 5 \times 10^4$ samples from each population in the two population spaces. For each sample, variance estimates were obtained by both the Monte Carlo method and the approximate variance estimators. In total, 900 scenarios were generated: 400 for the Sukhatme population space, and 500 for the MU273 population space.

Fifteen approximate variance estimators listed in section 2.6 were considered: v_{D_1} , v_{D_2} , v_{D_3} , v_{FP} , v_{Be} , v_{Ti} , v_{MT_1} , v_{MT_2} , v_{MT_3} , v_{MT_4} , v_{MT_5} , v_{B_1} , v_{B_2} , v_{B_3} and v_{B_4} . As for the Monte Carlo method, joint-inclusion probabilities were estimated using 10^7 replications, and were then plugged into the Horvitz–Thompson and Sen–Yates–Grundy variance estimators. The HT and SYG estimators (1.9) and (1.10) were also computed for comparison. The characteristics of the simulation study have been summarised in table 3.4.

Table 3.4: Characteristics of the simulation study.

	Sukhatme population space	MU273 population space
Sampling design	Maximum Entropy, Randomised Systematic, Tillé, Brewer	
Number of replicates for Monte Carlo π_{ij} estimates	10^7	
Number of replicates for the simulation study	5×10^4	
N	34	273
CV(Y)	0.5, 0.67, 0.76, 0.95, 1.05	0.5, 0.67, 0.82, 1, 1.29
CV(X)	0.5, 0.67, 0.77, 0.96, 1.1	0.5, 0.67, 0.81, 1, 1.27
Sampling fraction	5%, 10%, 15%, 20%	1%, 5%, 10%, 15%, 20%

Performances were measured through Monte Carlo Relative Bias, Mean Square Error, and Relative Stability, expressed as in Haziza et al. (2008):

$$\begin{aligned} \text{RB}_{\text{MC}}(v) &= 100 \times \frac{E_{\text{MC}}[v] - \text{MSE}_{\text{MC}}[\hat{\vartheta}]}{\text{MSE}_{\text{MC}}[\hat{\vartheta}]}, \\ \text{MSE}_{\text{MC}}(v) &= E_{\text{MC}} \left[v - \text{MSE}_{\text{MC}}[\hat{\vartheta}] \right]^2, \\ \text{RS}_{\text{MC}}(v) &= \frac{\text{MSE}[v]}{\text{MSE}[v_{\text{SYG}}]}, \end{aligned}$$

where

$$\begin{aligned} E_{\text{MC}}[\hat{\vartheta}] &= \frac{1}{K} \sum_{k \in K} \hat{\vartheta}^{(k)}, \\ \text{MSE}_{\text{MC}}[\hat{\vartheta}] &= E_{\text{MC}}[(\hat{\vartheta} - \vartheta)^2] \end{aligned}$$

and v_{SYG} is the Sen–Yates–Grundy variance estimator. For the computation of the SYG variance under the Randomised Systematic and Brewer designs, for which exact joint-inclusion probability are not available, the π_{ij} approximations by Hartley and Rao (1962) and by Brewer and Donadio (2003, eq. (18)) were employed, respectively.

All computations were made through R software. Computation of first-order inclusion probabilities and sample selection were performed by means of package `sampling` by Tillé and Matei (2016), while the rest was implemented through self-written code. We collected the main routines for the approximation of inclusion probabilities and for variance estimation in two R packages: `jipApprox` (Sichera, 2018a) and `UPSvarApprox` (Sichera, 2018b), which are described in appendix C.

3.2 RESULTS

The results of the simulation study are reported by tables B.13 to B.62 in appendix B, which show the Monte Carlo Relative Bias (RB), Relative Stability (RS) and Mean Square Error (MSE) of the estimators con-

sidered, by N , $CV(X)$, $CV(Y)$, sampling design and sampling fraction. Tables B.1 to B.6 give an overview of the results by presenting some summary statistics of the distribution of RB, RS and MSE over all scenarios, while tables B.7 to B.12 show these summaries divided by sampling design. Tables B.1 to B.12 are also reported in this chapter for an easier consultation. In the tables, RB values greater than 5 and RS values greater than 1 are in bold. As the Horvitz–Thompson and Sen–Yates–Grundy variance estimators are design–unbiased, their RB was omitted. Furthermore, in some scenarios, under Tillé’s elimination procedure some null joint–inclusion probabilities were produced, leading to inadmissible values for the relative stability, which were omitted.

The nomenclature used for the approximate variance estimators will be consistent with the notation defined in the previous chapter, while v_{MC1} and v_{MC2} will indicate the estimators obtained by using Monte Carlo approximations to the joint–inclusion probabilities with the Horvitz–Thompson and Sen–Yates–Grundy estimators, respectively.

In tables 3.5 and 3.6 some summary statistics for the distribution of the Relative Bias of the variance estimators are reported. The violin plots in figure 3.1 illustrate the distribution of the Relative bias for each estimator, in the two population spaces. Overall, with the larger populations the RB have similar distribution among the estimators, while for the small populations approximate variance estimators are generally more biased than Monte Carlo estimators, and especially the estimators belonging to the Matei–Tillé and Brewer families. In particular, the v_{B4} estimator always underestimates the variance, thus its use does not appear to be appropriate with small populations; this can be explained by the fact that it was derived under large population and high–entropy assumptions. Moreover, the bias of the v_{MC2} estimator is always comparable or lower than that of approximate estimators for each percentile considered, the only exception being its maximum value for the MU273 population space. Furthermore, the maximum bias reported for the Monte Carlo estimators is lower than that of approximate estimators when $N = 34$. Tables 3.7 and 3.8 report the same results conditioned on the sampling design, showing that for $N = 273$ the Relative Bias is generally low, with the exception of Tillé sampling, where a large proportion ($\approx 25\%$) of results are highly biased for all estimators considered, which is likely due to the presence of some null π_{ij} . However, when $N = 34$, approximate estimators have larger bias than Monte Carlo estimators, and many of them have RB larger than 5% in about 25% of the cases under all sampling designs.

As for the MSE, summary quantities are reported in tables 3.5 and 3.6. In figure 3.2 the MSE distribution for each estimator is summarised by boxplots; values are on a logarithmic scale to improve

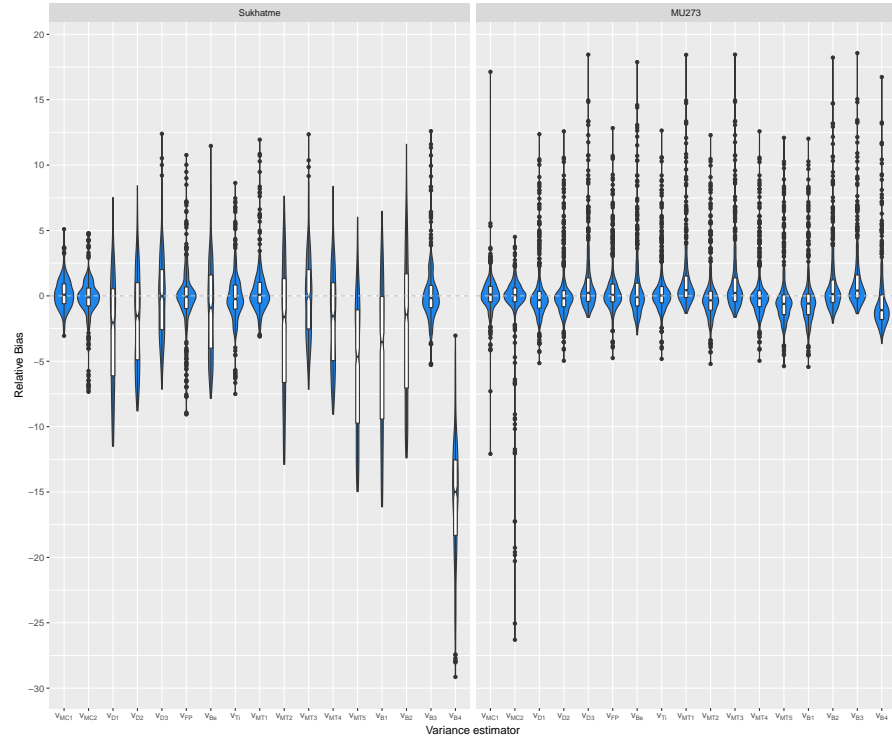


Figure 3.1: Distribution of the Relative Bias of the variance estimators over all scenarios, by population space.

readability. There does not seem to be any relevant difference among the estimators, however, for $N = 273$ Monte Carlo estimators have maximum value higher than approximate estimators under Brewer sampling (see table 3.15).

Overall, relative stability is good (tables 3.9 and 3.10), especially under high-entropy designs (see tables 3.11 and 3.12), where for over about 90% of the scenarios it was lower than or equal to 1. An exception is Tillé sampling, for which RS was over 1 in about 50% of the scenarios. Moreover, the v_{HT} estimator showed to be highly unstable in some cases, assuming extremely high values under all designs and at a larger extent for $N = 273$. The same can be said for the Monte Carlo estimator v_{MC1} , which is based on v_{HT} .

We will now consider the behaviour of the variance estimators more in details through the analysis of the results reported in tables B.13 to B.62, in which the RB, RS, and MSE are reported for each of the 900 scenarios under study.

When $N = 34$, the bias of the Monte Carlo estimators is low and it is generally comparable or lower than that of approximate estimators, with a few exceptions for v_{MC2} when $CV(X) = 1.1$, in particular when $f = 5\%$ with Brewer and Randomised Systematic sampling, and when $f = 20\%$ for Tillé's elimination procedure. Approximate variance estimators perform well when $CV(X)$ is small and in particular when $CV(X) = \{0.5, 0.67\}$. When $0.77 \leq CV(X) \leq 1.1$ most approximate

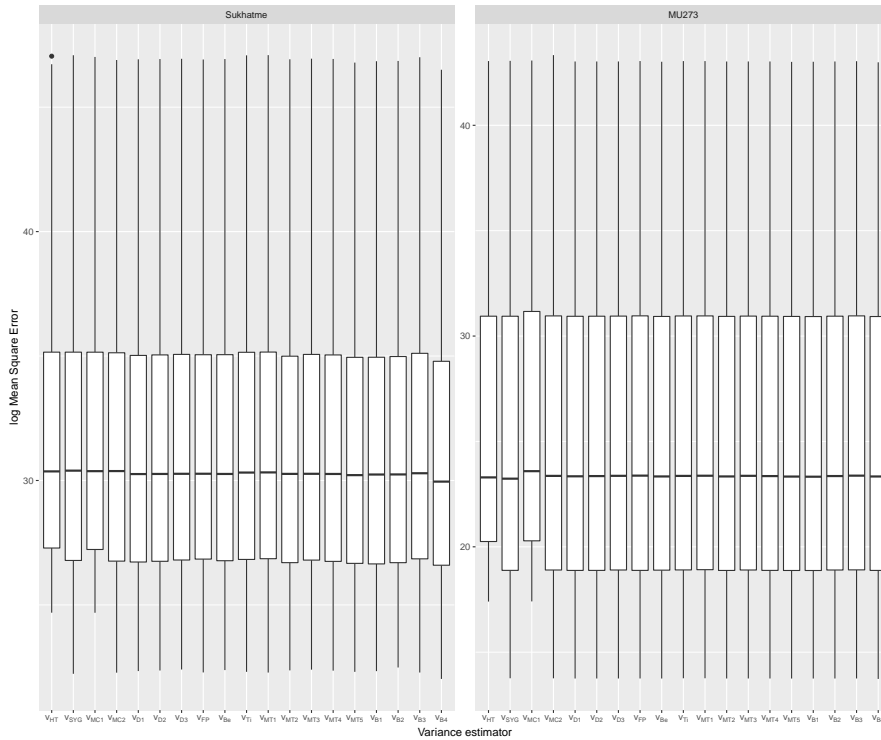


Figure 3.2: Distribution of the MSE of the variance estimators over all scenarios, on logarithmic scale.

estimators have relative bias close or greater than 5, especially those in the Matei–Tillé and Brewer families. Among the approximate estimators, v_{Ti} and v_{Be} have low bias even in those cases where other approximate estimators perform poorly, with some exceptions. The v_{Ti} estimator had large bias in some scenarios under Tillé sampling and with $f = \{15\%, 20\%\}$, while v_{Be} had large bias in many scenarios where $f = 20\%$, which is reasonable, as this method was derived under the assumption that $N - n \rightarrow \infty$. Finally, v_{B4} is always strongly negatively biased and should be avoided when population size is small. Furthermore, the value of $CV(Y)$ does not seem to influence the Relative Bias.

As for the Relative Stability, approximate variance estimators are usually more stable than the SYG estimator. Some exceptions occurred when $CV(Y) \approx CV(X) < 1$, when these estimators, and especially the Matei–Tillé, Brewer and Deville families, assumed large MSE values, up to 38% larger than the MSE of the SYG estimator. Large RS values are encountered more frequently with Tillé sampling. The Monte Carlo estimator v_{MC2} , on the other hand, tend to have Relative Stability close to 1 most of the times, with a few exceptions when $CV(Y) < 1$ and $CV(X) \approx 1$, and when $CV(Y) \approx CV(X)$. This is not surprising, as v_{MC2} is directly obtained from the SYG estimator. Estimator v_{MC1} , on the other hand, appears to be highly unstable, especially when $CV(Y) \approx CV(X) < 1$ and f is large. The MSE of Ap-

proximate estimators is usually comparable or lower than that of the Monte Carlo estimators.

With regards to the MU273 population space, both the approaches have low bias when $CV(X) < 1$. Nonetheless, when $CV(X) \geq 1$, Monte Carlo estimates are more biased. However, v_{MC2} performs better than approximate variance estimators for Tillé sampling and large sampling fractions. Monte Carlo estimates behave reasonably well even with larger populations, but they can be biased with small sampling fractions ($f = 1\%$) and under Tillé, CPS, and Randomised Systematic designs, while under Brewer sampling they performed generally well, with only two cases in which $RB > 5$ (one for v_{MC1} and one for v_{MC2}). Suggesting a preferable approximate estimator is more complicated here, because it depends on the characteristics of the scenario, however, the estimators v_{FP} (class 2) and v_{Be} (class 3) seem to behave well in most scenarios.

The relative stability is very close to 1 in most scenarios for both Approximate and Monte Carlo estimators. Monte Carlo estimator v_{MC2} is more or as efficient as the SYG estimator ($RS \leq 1$) in the 50% of the scenarios, and less stable generally when $CV(Y) < 1$ and $CV(X) \geq 1$, especially for small sampling fractions ($f = 1\%, 5\%$). On the other hand, v_{MC1} is again very unstable, with $RB \leq 1$ in less than 25% of the cases, with some extremely large RS values in several scenarios, in particular when $CV(X) < 1$. Estimators are mostly unstable under Tillé sampling, with very large values of relative stability for $CV(Y) \geq 1$ and $CV(X) \geq 1$. Finally, all estimators are unstable ($RS > 1$) for $CV(Y) = 1.29$, $CV(X) = 0.5$.

Table 3.5: Summary statistics for the distribution of Monte Carlo Relative Bias of different variance estimators for the MU273 population.

Estimator	min	5%	25%	50%	75%	95%	max
v_{HT}	—	—	—	—	—	—	—
v_{SYG}	—	—	—	—	—	—	—
v_{MC1}	0.01	0.08	0.27	0.57	1.16	2.66	17.14
v_{MC2}	0	0.05	0.23	0.5	0.93	3.8	26.3
v_{D1}	0	0.07	0.33	0.73	1.56	4.89	12.37
v_{D2}	0	0.06	0.27	0.67	1.43	4.95	12.58
v_{D3}	0.01	0.07	0.32	0.72	1.5	6.48	18.45
v_{FP}	0	0.05	0.3	0.62	1.24	5.03	12.83
v_{Be}	0	0.07	0.29	0.59	1.25	4.97	12.64
v_{Ti}	0	0.06	0.29	0.7	1.52	6.53	18.43
v_{MT1}	0	0.1	0.36	0.78	1.67	4.95	12.29
v_{MT2}	0.01	0.07	0.32	0.71	1.5	6.47	18.45
v_{MT3}	0	0.06	0.27	0.67	1.43	4.95	12.58
v_{MT4}	0.01	0.09	0.48	0.99	1.82	4.76	12.09
v_{MT5}	0.01	0.08	0.45	1	1.88	4.75	12.02
v_{B1}	0	0.06	0.3	0.72	1.47	6.33	18.22
v_{B2}	0	0.07	0.37	0.84	1.71	6.1	17.88
v_{B3}	0	0.03	0.28	0.69	1.6	6.66	18.56
v_{B4}	0.01	0.27	0.9	1.45	2.13	5.06	16.73

Table 3.6: Summary statistics for the distribution of Monte Carlo Relative Bias of different variance estimators for the Sukhatme population.

Estimator	min	5%	25%	50%	75%	95%	max
v_{HT}	—	—	—	—	—	—	—
v_{SYG}	—	—	—	—	—	—	—
v_{MC1}	0	0.08	0.36	0.73	1.19	2.48	5.1
v_{MC2}	0	0.07	0.28	0.69	1.3	4.04	7.34
v_{D1}	0.02	0.22	1.2	2.92	6.15	10.33	11.5
v_{D2}	0.01	0.24	1.24	2.96	5.13	7.48	8.8
v_{D3}	0.01	0.17	1.1	2.29	3.56	6.06	12.4
v_{FP}	0.01	0.06	0.3	0.82	2.18	6.94	10.77
v_{Be}	0	0.1	0.47	0.98	1.9	5.67	8.64
v_{Ti}	0	0.07	0.32	0.79	1.43	4.3	11.95
v_{MT1}	0	0.25	1.48	3.79	6.82	11.38	12.89
v_{MT2}	0.03	0.17	1.05	2.34	3.6	6.12	12.35
v_{MT3}	0	0.24	1.22	2.97	5.22	7.64	9.05
v_{MT4}	0.02	0.41	1.61	4.76	9.74	13.99	14.97
v_{MT5}	0.01	0.36	1.75	4.22	9.42	14.62	16.13
v_{B1}	0	0.16	1.16	2.85	4.54	6.55	11.47
v_{B2}	0.01	0.21	1.6	4.57	7.24	11.09	12.39
v_{B3}	0	0.1	0.43	0.88	1.95	5.17	12.6
v_{B4}	3.04	8.75	12.55	14.99	18.32	24.55	29.14

Table 3.7: Summary statistics for the distribution of Monte Carlo Relative Bias of different variance estimators for the MU273 population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max
Brewer	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC1	0.02	0.09	0.21	0.4	1.18	2.78	17.14
	VMC2	0	0.06	0.21	0.4	0.9	1.78	9.35
	VD1	0	0.06	0.3	0.57	0.95	1.57	2.13
	VD2	0	0.03	0.19	0.49	0.9	1.38	1.81
	VD3	0.01	0.06	0.32	0.61	1.11	3.01	3.58
	VFP	0	0.02	0.22	0.45	0.83	1.14	1.6
	VBe	0.01	0.07	0.22	0.41	0.75	1.19	1.7
	VTi	0.02	0.06	0.25	0.64	1.12	2.89	3.52
	VMT1	0.01	0.07	0.32	0.56	1.15	1.82	2.39
	VMT2	0.01	0.06	0.32	0.61	1.11	3.01	3.58
	VMT3	0	0.03	0.19	0.49	0.9	1.38	1.81
	VMT4	0.01	0.07	0.41	0.78	1.37	2.22	2.49
	VMT5	0.01	0.08	0.39	0.76	1.55	2.33	2.75
	VB1	0.01	0.03	0.23	0.62	1.08	2.73	3.46
VB2	0.01	0.09	0.3	0.65	1.45	2.72	3.36	
VB3	0	0.02	0.25	0.63	1.04	3.08	3.56	
VB4	0.04	0.23	0.75	1.14	1.81	2.28	3.66	
Tillé	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC1	0.02	0.09	0.35	0.72	1.02	2.68	7.29
	VMC2	0.01	0.04	0.25	0.57	0.89	6.64	25.05
	VD1	0.01	0.08	0.6	1.68	4.21	8.48	12.37
	VD2	0.01	0.09	0.58	1.76	4.27	8.6	12.58
	VD3	0.01	0.07	0.65	2.05	5.89	12.18	18.45
	VFP	0	0.06	0.6	2.15	4.63	8.8	12.83
	VBe	0	0.11	0.56	1.96	4.44	8.7	12.64
	VTi	0	0.11	0.7	2.27	5.79	12	18.43
	VMT1	0	0.13	0.58	1.53	4.3	8.45	12.29
	VMT2	0.02	0.07	0.65	2.05	5.89	12.18	18.45
	VMT3	0.01	0.08	0.58	1.76	4.27	8.6	12.58
	VMT4	0.01	0.12	0.59	1.62	4.1	8.23	12.09
	VMT5	0.03	0.11	0.54	1.44	4.16	8.2	12.02
	VB1	0.01	0.09	0.61	2.03	5.65	12.05	18.22
VB2	0.03	0.1	0.63	1.83	5.13	12.05	17.88	
VB3	0.01	0.07	0.64	2.29	5.82	12.06	18.56	
VB4	0.04	0.3	1.07	2.25	4.22	10.37	16.73	
CPS	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC1	0.01	0.08	0.28	0.57	1.11	2.5	3.71
	VMC2	0.01	0.06	0.2	0.5	1.05	11.87	26.3
	VD1	0	0.03	0.2	0.47	0.74	1.69	2.08
	VD2	0	0.04	0.2	0.41	0.69	1.43	1.7
	VD3	0.01	0.06	0.27	0.63	1.2	3.52	4.96
	VFP	0	0.05	0.21	0.43	0.72	1.31	1.59
	VBe	0	0.03	0.2	0.44	0.69	1.33	1.61
	VTi	0.01	0.04	0.34	0.63	1.23	3.28	5
	VMT1	0.02	0.09	0.26	0.52	0.88	1.83	2.37
	VMT2	0.01	0.06	0.27	0.63	1.2	3.52	4.96
	VMT3	0	0.04	0.2	0.41	0.69	1.44	1.7
	VMT4	0.02	0.1	0.38	0.65	1.23	2.13	2.67
	VMT5	0.01	0.07	0.32	0.7	1.33	2.29	2.95
	VB1	0.01	0.06	0.26	0.66	1.2	3.27	4.76
VB2	0.02	0.07	0.39	0.8	1.34	3.05	4.45	
VB3	0.01	0.05	0.29	0.64	1.23	3.5	5.06	
VB4	0.01	0.1	0.88	1.35	1.88	2.57	3.44	
Rand. Sys.	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC1	0.03	0.07	0.28	0.73	1.38	2.64	12.08
	VMC2	0	0.08	0.27	0.54	0.92	3.11	10.18
	VD1	0.01	0.12	0.5	1.08	1.88	3.61	5.14
	VD2	0.02	0.14	0.42	1.04	1.79	3.14	4.96
	VD3	0.02	0.08	0.25	0.57	0.92	1.58	2.3
	VFP	0.04	0.11	0.35	0.71	1.29	2.69	4.75
	VBe	0.02	0.1	0.37	0.84	1.39	2.74	4.82
	VTi	0	0.05	0.23	0.45	0.81	1.24	2.37
	VMT1	0.05	0.12	0.57	1.19	2.09	3.64	5.2
	VMT2	0.02	0.08	0.25	0.57	0.92	1.58	2.3
	VMT3	0.02	0.14	0.42	1.04	1.79	3.13	4.96
	VMT4	0.01	0.08	0.68	1.43	2.31	3.82	5.37
	VMT5	0.02	0.08	0.65	1.46	2.44	3.96	5.43
	VB1	0	0.07	0.25	0.58	1	1.88	2.3
VB2	0	0.06	0.31	0.72	1.57	2.47	2.99	
VB3	0	0.03	0.21	0.44	0.82	1.2	2.35	
VB4	0.12	0.44	0.99	1.45	1.99	2.85	3.16	

Table 3.8: Summary statistics for the distribution of Monte Carlo Relative Bias of different variance estimators for the Sukhatme population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max
Brewer	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0.01	0.09	0.46	0.73	1.16	2.54	3.72
	VMC ₂	0.01	0.15	0.42	0.76	1.18	4.76	7.24
	VD ₁	0.02	0.28	1.08	3.24	6.34	9.84	11.33
	VD ₂	0.04	0.29	1.05	2.46	4.92	7.18	7.53
	VD ₃	0.04	0.15	1.06	2.01	3.34	5.01	5.9
	VFP	0.02	0.04	0.28	0.62	1.07	5.41	7.71
	VBe	0.01	0.11	0.52	0.78	1.39	2.62	6.24
	VTi	0	0.06	0.38	0.68	1.15	2	2.36
	VMT ₁	0.1	0.3	1.41	3.72	6.76	11.31	12.84
	VMT ₂	0.03	0.16	1.04	2.08	3.4	5.06	5.86
	VMT ₃	0.02	0.32	1.06	2.49	4.99	7.35	7.84
	VMT ₄	0.07	0.23	1.58	5.28	9.79	13.37	14.72
	VMT ₅	0.07	0.35	1.33	4.67	9.83	14.62	16.06
	VB ₁	0.01	0.16	1	2.38	4.04	6.37	6.57
VB ₂	0.05	0.22	1.53	4.72	6.78	11.09	11.31	
VB ₃	0	0.11	0.45	0.85	1.69	3.05	3.69	
VB ₄	9.75	11.15	12.56	15.14	17.89	24.91	27.41	
Tillé	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0.03	0.07	0.37	0.76	1.31	2.36	2.55
	VMC ₂	0.03	0.07	0.39	1.05	1.84	4.03	6.44
	VD ₁	0.02	0.22	1.57	2.65	5.63	9.1	11.34
	VD ₂	0.32	0.61	1.68	2.96	5.08	6.57	8.42
	VD ₃	0.1	0.37	1.75	2.94	4.36	8.47	12.4
	VFP	0.05	0.23	1.4	2.25	4.31	7.29	10.77
	VBe	0.04	0.23	1.03	2.06	3.71	6.83	8.64
	VTi	0.06	0.12	0.75	1.59	3.21	8.42	11.95
	VMT ₁	0.08	0.45	1.58	3.54	5.94	10.35	12.8
	VMT ₂	0.13	0.42	1.84	2.94	4.4	8.5	12.35
	VMT ₃	0.29	0.56	1.7	2.98	5.12	6.72	8.38
	VMT ₄	0.04	0.45	1.35	3.27	7.96	12.57	14.72
	VMT ₅	0.05	1.15	2.32	3.46	7.94	13.69	16.02
	VB ₁	0.02	0.61	2.12	3.33	4.93	7.58	11.47
VB ₂	0.14	0.73	2.65	3.96	7.15	10.17	11.6	
VB ₃	0.04	0.1	0.59	1.78	3.23	9.19	12.6	
VB ₄	3.04	5.61	9.73	13.6	17.75	24.48	27.71	
CPS	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0	0.08	0.3	0.72	1.17	1.86	5.1
	VMC ₂	0	0.08	0.25	0.44	1.17	3.11	3.86
	VD ₁	0.03	0.3	1.11	2.9	6.28	9.48	10.63
	VD ₂	0.02	0.15	1.02	2.8	4.85	7.83	8.06
	VD ₃	0.04	0.31	0.86	1.84	3.25	6.04	7.14
	VFP	0.01	0.03	0.23	0.5	1.16	5.15	9.06
	VBe	0.01	0.05	0.45	0.77	1.34	2.62	3
	VTi	0	0.1	0.27	0.66	1.1	2.2	2.56
	VMT ₁	0	0.15	1.37	3.77	7.35	10.97	12.21
	VMT ₂	0.03	0.31	0.86	1.85	3.3	6.11	7.15
	VMT ₃	0	0.12	0.99	2.79	4.9	7.83	8.12
	VMT ₄	0.15	0.6	1.92	5.36	9.76	14.04	14.36
	VMT ₅	0.03	0.4	1.65	5.18	9.72	14.29	15.45
	VB ₁	0	0.14	0.89	2.27	3.97	7.4	7.85
VB ₂	0.1	0.26	1.05	4.5	6.86	11.06	11.9	
VB ₃	0.02	0.13	0.43	0.74	1.35	2.91	5.28	
VB ₄	10.13	11.35	12.97	15.07	18.37	25.21	29.14	
Rand. Sys.	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0.02	0.07	0.27	0.71	1.28	2.1	3.65
	VMC ₂	0	0.03	0.23	0.55	1.26	4.28	7.34
	VD ₁	0.03	0.14	1.26	3.69	6.58	10.69	11.5
	VD ₂	0.01	0.21	0.97	3.2	5.73	8.24	8.8
	VD ₃	0.01	0.07	0.66	2.54	3.73	5.53	7.1
	VFP	0.01	0.08	0.26	0.6	1.81	5.3	6.98
	VBe	0	0.09	0.31	0.79	1.34	4.51	5.97
	VTi	0.01	0.04	0.21	0.52	1.07	2.02	3.09
	VMT ₁	0.02	0.15	1.57	4.19	7.35	12.28	12.89
	VMT ₂	0.03	0.09	0.68	2.42	3.77	5.34	7.1
	VMT ₃	0.06	0.22	1.01	3.24	5.75	8.43	9.05
	VMT ₄	0.02	0.5	1.58	5.26	10.45	14.09	14.97
	VMT ₅	0.01	0.26	1.63	4.79	9.91	15.52	16.13
	VB ₁	0.01	0.16	0.57	3.17	4.98	6.1	7.37
VB ₂	0.01	0.14	1.02	5.06	8.27	11.41	12.39	
VB ₃	0.02	0.06	0.34	0.94	1.61	2.74	3.06	
VB ₄	10.35	11.2	13.22	15.54	18.08	24.49	27.48	

Table 3.9: Summary statistics for the distribution of Monte Carlo Relative Stability of different variance estimators for the MU273 population.

Estimator	min	5%	25%	50%	75%	95%	max
v _{HT}	0.87	0.96	1	1.01	2.06	73.94	654.26
v _{SYG}	1	1	1	1	1	1	1
v _{MC1}	0.86	0.97	1.01	1.11	3.61	97.2	661.75
v _{MC2}	0.54	0.94	1	1	1.02	1.13	2.04
v _{D1}	0.85	0.93	0.97	0.98	1	1.15	1.78
v _{D2}	0.86	0.93	0.98	0.98	1	1.15	1.79
v _{D3}	0.92	0.96	0.98	0.99	1	1.21	2.37
v _{FP}	0.86	0.94	0.99	1	1	1.15	1.76
v _{Be}	0.9	0.97	0.99	1	1	1.17	1.68
v _{Ti}	0.96	0.99	1	1	1.01	1.2	2.16
v _{MT1}	0.85	0.93	0.97	0.98	1	1.15	1.79
v _{MT2}	0.92	0.96	0.98	0.99	1	1.21	2.37
v _{MT3}	0.86	0.93	0.98	0.98	1	1.15	1.79
v _{MT4}	0.85	0.93	0.96	0.97	0.99	1.14	1.76
v _{MT5}	0.85	0.92	0.96	0.97	0.99	1.14	1.78
v _{B1}	0.92	0.96	0.98	0.99	1	1.2	2.36
v _{B2}	0.9	0.95	0.97	0.98	1	1.2	2.4
v _{B3}	0.93	0.96	0.99	1	1.01	1.21	2.32
v _{B4}	0.88	0.94	0.96	0.97	0.98	1.14	2.01

Table 3.10: Summary statistics for the distribution of Monte Carlo Relative Stability of different variance estimators for the Sukhatme population.

Estimator	min	5%	25%	50%	75%	95%	max
v _{HT}	0.48	0.93	0.98	1.03	1.38	22.97	58.43
v _{SYG}	1	1	1	1	1	1	1
v _{MC1}	0.43	0.92	0.98	1.04	1.31	22.88	53.99
v _{MC2}	0.54	0.96	0.99	1	1.01	1.08	1.23
v _{D1}	0.37	0.77	0.85	0.89	0.97	1.12	1.25
v _{D2}	0.37	0.82	0.88	0.91	0.97	1.13	1.27
v _{D3}	0.38	0.86	0.91	0.94	1	1.15	1.34
v _{FP}	0.36	0.85	0.96	0.99	1.01	1.11	1.26
v _{Be}	0.43	0.93	0.97	0.99	1	1.08	1.3
v _{Ti}	0.43	0.94	0.99	1	1.01	1.08	1.36
v _{MT1}	0.37	0.75	0.83	0.88	0.97	1.16	1.28
v _{MT2}	0.38	0.86	0.91	0.94	1	1.15	1.34
v _{MT3}	0.37	0.82	0.88	0.9	0.97	1.12	1.27
v _{MT4}	0.32	0.71	0.79	0.85	0.94	1.08	1.23
v _{MT5}	0.34	0.69	0.78	0.84	0.95	1.14	1.24
v _{B1}	0.37	0.83	0.89	0.92	0.98	1.14	1.33
v _{B2}	0.35	0.75	0.82	0.86	0.97	1.21	1.38
v _{B3}	0.4	0.91	0.96	0.98	1.01	1.1	1.33
v _{B4}	0.24	0.56	0.7	0.75	0.8	0.85	1.03

Table 3.11: Summary statistics for the distribution of Monte Carlo Relative Stability of different variance estimators for the MU273 population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max
Brewer	v _{HT}	0.92	0.96	1	1.01	2.17	72.32	644.61
	v _{SYG}	1	1	1	1	1	1	1
	v _{MC1}	0.86	0.97	1	1.08	3.2	90.13	642.14
	v _{MC2}	0.82	0.93	0.99	1	1.03	1.28	1.56
	v _{D1}	0.85	0.89	0.97	0.98	0.99	1	1.06
	v _{D2}	0.86	0.89	0.97	0.98	0.99	1	1.06
	v _{D3}	0.94	0.96	0.97	0.99	1	1.01	1.15
	v _{FP}	0.86	0.9	0.98	0.99	1	1.01	1.06
	v _{Be}	0.9	0.94	0.99	0.99	1	1	1.04
	v _{Ti}	0.99	0.99	1	1	1	1.01	1.01
	v _{MT1}	0.85	0.89	0.96	0.97	0.99	1.01	1.07
	v _{MT2}	0.94	0.96	0.97	0.99	1	1.01	1.15
	v _{MT3}	0.86	0.89	0.97	0.98	0.99	1	1.06
	v _{MT4}	0.85	0.89	0.96	0.97	0.98	1	1.06
	v _{MT5}	0.85	0.88	0.95	0.97	0.98	1	1.06
	v _{B1}	0.94	0.96	0.97	0.99	0.99	1.01	1.14
v _{B2}	0.93	0.95	0.96	0.97	0.99	1.01	1.15	
v _{B3}	0.94	0.97	0.98	0.99	1	1.01	1.13	
v _{B4}	0.88	0.93	0.95	0.96	0.97	0.99	1.08	
Tillé	v _{HT}	0.87	0.96	1	1.05	2.21	72.69	603.09
	v _{SYG}	1	1	1	1	1	1	1
	v _{MC1}	0.87	0.97	1.02	1.15	5	102.02	609.86
	v _{MC2}	0.54	0.96	1	1	1.02	1.1	1.17
	v _{D1}	0.94	0.96	0.99	1.01	1.11	1.38	1.78
	v _{D2}	0.94	0.97	0.99	1.01	1.11	1.38	1.79
	v _{D3}	0.94	0.97	0.99	1.01	1.13	1.67	2.37
	v _{FP}	0.95	0.98	1	1.02	1.12	1.38	1.76
	v _{Be}	0.97	0.99	1	1.01	1.13	1.39	1.68
	v _{Ti}	0.97	0.99	1	1.02	1.15	1.68	2.16
	v _{MT1}	0.94	0.96	0.99	1	1.1	1.38	1.79
	v _{MT2}	0.94	0.97	0.99	1.01	1.13	1.67	2.37
	v _{MT3}	0.94	0.97	0.99	1.01	1.11	1.38	1.79
	v _{MT4}	0.93	0.96	0.98	1	1.1	1.37	1.76
	v _{MT5}	0.93	0.95	0.98	1	1.1	1.37	1.78
	v _{B1}	0.94	0.96	0.99	1.01	1.13	1.66	2.36
v _{B2}	0.93	0.95	0.98	1	1.12	1.65	2.4	
v _{B3}	0.95	0.97	1	1.02	1.14	1.66	2.32	
v _{B4}	0.92	0.94	0.97	0.99	1.1	1.48	2.01	
CPS	v _{HT}	0.92	0.96	1	1.01	1.88	70.62	635.98
	v _{SYG}	1	1	1	1	1	1	1
	v _{MC1}	0.92	0.96	1	1.09	3.4	88.02	642.73
	v _{MC2}	0.78	0.82	1	1	1.01	1.07	1.17
	v _{D1}	0.92	0.95	0.97	0.98	0.99	1	1.11
	v _{D2}	0.92	0.95	0.98	0.98	0.99	1.01	1.12
	v _{D3}	0.92	0.97	0.99	0.99	1	1.04	1.34
	v _{FP}	0.92	0.95	0.98	1	1	1.01	1.11
	v _{Be}	0.98	0.99	1	1	1	1	1
	v _{Ti}	0.96	1	1	1	1.01	1.08	1.18
	v _{MT1}	0.92	0.94	0.96	0.98	0.99	1.01	1.12
	v _{MT2}	0.92	0.97	0.99	0.99	1	1.04	1.34
	v _{MT3}	0.92	0.95	0.98	0.98	0.99	1.01	1.12
	v _{MT4}	0.91	0.94	0.96	0.97	0.99	1	1.11
	v _{MT5}	0.91	0.94	0.95	0.97	0.99	1	1.11
	v _{B1}	0.92	0.97	0.98	0.99	1	1.04	1.33
v _{B2}	0.9	0.96	0.97	0.98	1	1.04	1.35	
v _{B3}	0.93	0.98	0.99	1	1.01	1.04	1.32	
v _{B4}	0.9	0.94	0.96	0.97	0.97	1.01	1.21	
Rand. Sys.	v _{HT}	0.91	0.96	1	1	1.86	69.98	654.26
	v _{SYG}	1	1	1	1	1	1	1
	v _{MC1}	0.9	0.97	1.01	1.1	3.62	97.14	661.75
	v _{MC2}	0.74	0.99	1	1.01	1.03	1.22	2.04
	v _{D1}	0.9	0.93	0.96	0.98	0.99	1	1.09
	v _{D2}	0.9	0.93	0.97	0.98	0.99	1	1.09
	v _{D3}	0.93	0.95	0.98	0.99	1	1.01	1.13
	v _{FP}	0.9	0.93	0.98	0.99	1	1	1.09
	v _{Be}	0.95	0.97	0.99	1	1	1	1.06
	v _{Ti}	1	1	1	1	1	1	1
	v _{MT1}	0.9	0.92	0.96	0.97	0.99	1	1.09
	v _{MT2}	0.93	0.95	0.98	0.99	1	1.01	1.13
	v _{MT3}	0.9	0.93	0.97	0.98	0.99	1	1.09
	v _{MT4}	0.9	0.92	0.96	0.97	0.98	1	1.09
	v _{MT5}	0.89	0.92	0.95	0.96	0.98	1	1.09
	v _{B1}	0.92	0.95	0.98	0.98	0.99	1.01	1.13
v _{B2}	0.92	0.94	0.96	0.97	0.99	1.01	1.13	
v _{B3}	0.93	0.96	0.99	1	1	1.01	1.13	
v _{B4}	0.92	0.93	0.95	0.96	0.97	0.99	1.12	

Table 3.12: Summary statistics for the distribution of Monte Carlo Relative Stability of different variance estimators for the Sukhatme population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max
Brewer	VHT	0.9	0.94	0.98	1.03	1.52	23.22	58.43
	VSYG	1	1	1	1	1	1	1
	VMC ₁	0.89	0.92	0.98	1.02	1.23	21.17	53.31
	VMC ₂	0.77	0.94	0.99	1.01	1.03	1.09	1.23
	VD ₁	0.81	0.82	0.85	0.88	0.95	1.17	1.21
	VD ₂	0.82	0.83	0.88	0.9	0.96	1.18	1.23
	VD ₃	0.83	0.84	0.91	0.94	1	1.19	1.26
	VFP	0.8	0.82	0.97	0.99	1.03	1.13	1.2
	VB _e	0.93	0.93	0.97	0.98	1	1.1	1.12
	VT _i	0.93	0.94	0.99	1	1.02	1.06	1.1
	VMT ₁	0.8	0.81	0.83	0.85	0.95	1.24	1.28
	VMT ₂	0.83	0.84	0.91	0.94	1	1.19	1.25
	VMT ₃	0.82	0.83	0.88	0.89	0.96	1.18	1.22
	VMT ₄	0.71	0.74	0.79	0.84	0.93	1.14	1.19
	VMT ₅	0.75	0.75	0.78	0.82	0.93	1.22	1.24
	VB ₁	0.82	0.83	0.89	0.91	0.98	1.19	1.24
VB ₂	0.77	0.78	0.81	0.84	0.96	1.27	1.38	
VB ₃	0.87	0.88	0.96	0.98	1.01	1.11	1.16	
VB ₄	0.52	0.56	0.71	0.76	0.8	0.84	0.86	
Tillé	VHT	0.48	0.85	1.03	1.07	1.4	23.08	53.8
	VSYG	1	1	1	1	1	1	1
	VMC ₁	0.43	0.85	1.01	1.08	1.44	23.07	53.99
	VMC ₂	0.54	0.96	0.99	1	1.01	1.03	1.06
	VD ₁	0.37	0.74	0.86	0.92	1.03	1.14	1.25
	VD ₂	0.37	0.74	0.89	0.94	1.04	1.15	1.27
	VD ₃	0.38	0.77	0.92	0.99	1.06	1.18	1.34
	VFP	0.36	0.82	0.95	1.01	1.08	1.17	1.26
	VB _e	0.43	0.87	0.96	1	1.04	1.16	1.3
	VT _i	0.43	0.88	0.97	1.01	1.07	1.18	1.36
	VMT ₁	0.37	0.73	0.83	0.9	1.04	1.15	1.26
	VMT ₂	0.38	0.76	0.92	0.99	1.06	1.17	1.34
	VMT ₃	0.37	0.75	0.89	0.94	1.04	1.15	1.27
	VMT ₄	0.32	0.65	0.81	0.86	0.99	1.12	1.23
	VMT ₅	0.34	0.68	0.8	0.85	1.01	1.12	1.24
	VB ₁	0.37	0.74	0.9	0.96	1.06	1.17	1.33
VB ₂	0.35	0.69	0.84	0.89	1.04	1.19	1.32	
VB ₃	0.4	0.81	0.96	1.01	1.06	1.18	1.33	
VB ₄	0.24	0.57	0.69	0.76	0.81	0.97	1.03	
CPS	VHT	0.88	0.93	0.98	1	1.22	20.86	50.62
	VSYG	1	1	1	1	1	1	1
	VMC ₁	0.89	0.94	0.98	1.02	1.24	20.9	50.74
	VMC ₂	0.96	0.98	1	1	1.01	1.11	1.17
	VD ₁	0.76	0.77	0.85	0.89	0.96	1.09	1.14
	VD ₂	0.82	0.82	0.87	0.9	0.97	1.1	1.15
	VD ₃	0.88	0.89	0.91	0.93	0.99	1.12	1.17
	VFP	0.86	0.87	0.96	0.99	1	1.03	1.09
	VB _e	0.93	0.95	0.98	0.99	1	1.02	1.04
	VT _i	0.96	0.97	1	1	1	1.02	1.04
	VMT ₁	0.73	0.75	0.82	0.88	0.97	1.15	1.19
	VMT ₂	0.88	0.89	0.91	0.93	0.99	1.12	1.17
	VMT ₃	0.82	0.82	0.87	0.9	0.97	1.1	1.15
	VMT ₄	0.7	0.71	0.78	0.84	0.94	1.06	1.12
	VMT ₅	0.68	0.69	0.78	0.84	0.95	1.12	1.16
	VB ₁	0.83	0.84	0.89	0.91	0.98	1.11	1.16
VB ₂	0.74	0.75	0.82	0.85	0.97	1.19	1.29	
VB ₃	0.93	0.94	0.95	0.98	1	1.04	1.06	
VB ₄	0.55	0.6	0.68	0.75	0.79	0.85	0.89	
Rand. Sys.	VHT	0.88	0.93	0.97	1	1.22	18.26	45.85
	VSYG	1	1	1	1	1	1	1
	VMC ₁	0.88	0.93	0.97	1	1.26	19.52	48.08
	VMC ₂	0.82	0.96	0.99	1	1	1.03	1.09
	VD ₁	0.77	0.81	0.85	0.89	0.94	1.08	1.13
	VD ₂	0.83	0.83	0.87	0.9	0.95	1.09	1.14
	VD ₃	0.86	0.87	0.91	0.93	0.98	1.1	1.17
	VFP	0.84	0.88	0.95	0.98	0.99	1.07	1.09
	VB _e	0.94	0.95	0.98	0.99	1	1.01	1.03
	VT _i	0.97	0.98	1	1	1	1.01	1.01
	VMT ₁	0.74	0.78	0.83	0.87	0.94	1.14	1.17
	VMT ₂	0.87	0.87	0.91	0.93	0.98	1.1	1.16
	VMT ₃	0.83	0.84	0.87	0.9	0.95	1.09	1.14
	VMT ₄	0.71	0.74	0.78	0.85	0.92	1.04	1.11
	VMT ₅	0.69	0.72	0.78	0.84	0.93	1.12	1.15
	VB ₁	0.84	0.85	0.89	0.91	0.96	1.1	1.15
VB ₂	0.75	0.76	0.81	0.85	0.96	1.19	1.28	
VB ₃	0.92	0.93	0.96	0.97	0.99	1.03	1.05	
VB ₄	0.56	0.59	0.7	0.74	0.78	0.83	0.85	

Table 3.13: Summary statistics for the distribution of Monte Carlo MSE of different variance estimators for the MU273 population.

Estimator	min	5%	25%	50%	75%	95%	max
v _H T	3.61e+07	1.15e+08	6.24e+08	1.31e+10	2.74e+13	5.20e+16	4.96e+18
v _S Y _G	9.49e+05	5.51e+06	1.59e+08	1.23e+10	2.74e+13	5.21e+16	5.02e+18
v _M C ₁	3.64e+07	1.11e+08	6.45e+08	1.76e+10	3.43e+13	5.29e+16	5.09e+18
v _M C ₂	9.46e+05	5.77e+06	1.61e+08	1.40e+10	2.78e+13	4.95e+16	6.57e+18
v _D 1	9.38e+05	6.23e+06	1.58e+08	1.38e+10	2.73e+13	5.09e+16	4.86e+18
v _D 2	9.38e+05	6.25e+06	1.58e+08	1.39e+10	2.74e+13	5.09e+16	4.86e+18
v _D 3	9.39e+05	6.67e+06	1.61e+08	1.41e+10	2.76e+13	5.09e+16	4.87e+18
v _F P	9.39e+05	6.26e+06	1.59e+08	1.42e+10	2.79e+13	5.18e+16	4.96e+18
v _B e	9.48e+05	6.65e+06	1.62e+08	1.40e+10	2.77e+13	5.19e+16	4.95e+18
v _T i	9.49e+05	7.04e+06	1.64e+08	1.41e+10	2.78e+13	5.20e+16	4.95e+18
v _M T ₁	9.38e+05	6.23e+06	1.58e+08	1.37e+10	2.72e+13	5.08e+16	4.86e+18
v _M T ₂	9.39e+05	6.67e+06	1.61e+08	1.41e+10	2.75e+13	5.09e+16	4.87e+18
v _M T ₃	9.38e+05	6.25e+06	1.58e+08	1.39e+10	2.74e+13	5.09e+16	4.87e+18
v _M T ₄	9.36e+05	6.22e+06	1.58e+08	1.36e+10	2.71e+13	5.00e+16	4.78e+18
v _M T ₅	9.37e+05	6.21e+06	1.58e+08	1.35e+10	2.70e+13	5.02e+16	4.80e+18
v _B 1	9.38e+05	6.64e+06	1.61e+08	1.39e+10	2.74e+13	5.09e+16	4.87e+18
v _B 2	9.37e+05	6.60e+06	1.60e+08	1.37e+10	2.71e+13	5.02e+16	4.80e+18
v _B 3	9.39e+05	6.68e+06	1.62e+08	1.42e+10	2.78e+13	5.16e+16	4.93e+18
v _B 4	9.20e+05	6.43e+06	1.58e+08	1.37e+10	2.70e+13	4.89e+16	4.68e+18

Table 3.14: Summary statistics for the distribution of Monte Carlo MSE of different variance estimators for the Sukhatme population.

Estimator	min	5%	25%	50%	75%	95%	max
v _H T	5.27e+10	1.23e+11	7.08e+11	1.54e+13	1.85e+15	5.40e+18	2.71e+20
v _S Y _G	4.53e+09	1.73e+10	4.31e+11	1.59e+13	1.85e+15	5.25e+18	2.81e+20
v _M C ₁	5.26e+10	1.21e+11	6.68e+11	1.56e+13	1.85e+15	5.32e+18	2.63e+20
v _M C ₂	4.73e+09	1.74e+10	4.19e+11	1.56e+13	1.81e+15	5.39e+18	2.32e+20
v _D 1	5.05e+09	1.74e+10	4.05e+11	1.38e+13	1.63e+15	4.40e+18	2.39e+20
v _D 2	5.16e+09	1.77e+10	4.17e+11	1.39e+13	1.66e+15	4.65e+18	2.41e+20
v _D 3	5.38e+09	1.82e+10	4.40e+11	1.40e+13	1.69e+15	4.87e+18	2.44e+20
v _F P	4.78e+09	1.87e+10	4.57e+11	1.40e+13	1.67e+15	5.40e+18	2.37e+20
v _B e	4.90e+09	1.89e+10	4.51e+11	1.47e+13	1.84e+15	5.20e+18	2.80e+20
v _T i	4.76e+09	1.93e+10	4.65e+11	1.48e+13	1.86e+15	5.40e+18	2.83e+20
v _M T ₁	5.19e+09	1.71e+10	3.95e+11	1.39e+13	1.58e+15	4.24e+18	2.39e+20
v _M T ₂	5.35e+09	1.82e+10	4.39e+11	1.40e+13	1.69e+15	4.86e+18	2.45e+20
v _M T ₃	5.14e+09	1.77e+10	4.16e+11	1.39e+13	1.66e+15	4.64e+18	2.43e+20
v _M T ₄	4.90e+09	1.69e+10	3.83e+11	1.33e+13	1.51e+15	4.06e+18	2.09e+20
v _M T ₅	5.04e+09	1.67e+10	3.75e+11	1.36e+13	1.51e+15	3.93e+18	2.21e+20
v _B 1	5.25e+09	1.79e+10	4.28e+11	1.39e+13	1.68e+15	4.72e+18	2.41e+20
v _B 2	5.83e+09	1.72e+10	3.94e+11	1.36e+13	1.56e+15	4.24e+18	2.24e+20
v _B 3	4.76e+09	1.86e+10	4.63e+11	1.43e+13	1.77e+15	5.21e+18	2.60e+20
v _B 4	3.68e+09	1.58e+10	3.57e+11	1.02e+13	1.28e+15	3.71e+18	1.57e+20

Table 3.15: Summary statistics for the distribution of Monte Carlo MSE of different variance estimators for the MU273 population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max
Brewer	VHT	5.16e+07	1.33e+08	6.45e+08	1.39e+10	2.78e+13	4.55e+16	4.96e+18
	VSYG	9.50e+05	7.16e+06	1.65e+08	1.38e+10	2.78e+13	4.60e+16	5.02e+18
	VMC1	4.36e+07	1.22e+08	6.36e+08	1.64e+10	3.44e+13	4.60e+16	5.09e+18
	VMC2	9.46e+05	6.84e+06	1.63e+08	1.37e+10	2.70e+13	6.08e+16	6.57e+18
	VD1	9.39e+05	6.37e+06	1.58e+08	1.35e+10	2.73e+13	4.46e+16	4.86e+18
	VD2	9.40e+05	6.38e+06	1.58e+08	1.36e+10	2.74e+13	4.47e+16	4.86e+18
	VD3	9.41e+05	6.88e+06	1.62e+08	1.38e+10	2.76e+13	4.48e+16	4.87e+18
	VFP	9.41e+05	6.40e+06	1.59e+08	1.39e+10	2.79e+13	4.55e+16	4.96e+18
	VBe	9.50e+05	6.73e+06	1.64e+08	1.38e+10	2.77e+13	4.57e+16	4.95e+18
	VTi	9.51e+05	7.19e+06	1.65e+08	1.39e+10	2.78e+13	4.57e+16	4.95e+18
	VMT1	9.40e+05	6.37e+06	1.58e+08	1.34e+10	2.71e+13	4.44e+16	4.86e+18
	VMT2	9.41e+05	6.88e+06	1.62e+08	1.38e+10	2.75e+13	4.48e+16	4.87e+18
	VMT3	9.40e+05	6.38e+06	1.58e+08	1.36e+10	2.74e+13	4.47e+16	4.87e+18
	VMT4	9.38e+05	6.35e+06	1.57e+08	1.33e+10	2.71e+13	4.41e+16	4.78e+18
	VMT5	9.38e+05	6.35e+06	1.57e+08	1.32e+10	2.69e+13	4.39e+16	4.80e+18
	VB1	9.40e+05	6.85e+06	1.62e+08	1.36e+10	2.74e+13	4.47e+16	4.87e+18
VB2	9.39e+05	6.82e+06	1.61e+08	1.34e+10	2.70e+13	4.40e+16	4.80e+18	
VB3	9.41e+05	6.89e+06	1.63e+08	1.39e+10	2.78e+13	4.55e+16	4.93e+18	
VB4	9.23e+05	6.61e+06	1.59e+08	1.34e+10	2.70e+13	4.37e+16	4.68e+18	
Tillé	VHT	3.61e+07	7.94e+07	5.56e+08	8.26e+09	2.63e+13	5.01e+16	3.59e+18
	VSYG	9.51e+05	5.28e+06	1.16e+08	6.96e+09	2.59e+13	4.98e+16	3.60e+18
	VMC1	3.64e+07	8.91e+07	6.14e+08	1.90e+10	3.43e+13	5.01e+16	3.67e+18
	VMC2	9.49e+05	5.39e+06	1.34e+08	1.22e+10	3.09e+13	4.20e+16	4.00e+18
	VD1	9.39e+05	7.17e+06	1.59e+08	1.23e+10	3.03e+13	4.79e+16	3.53e+18
	VD2	9.40e+05	7.21e+06	1.60e+08	1.23e+10	3.04e+13	4.79e+16	3.53e+18
	VD3	9.40e+05	8.45e+06	1.76e+08	1.24e+10	3.05e+13	4.80e+16	3.53e+18
	VFP	9.40e+05	7.22e+06	1.60e+08	1.25e+10	3.09e+13	4.88e+16	3.60e+18
	VBe	9.49e+05	7.50e+06	1.62e+08	1.25e+10	3.07e+13	4.85e+16	3.60e+18
	VTi	9.50e+05	8.70e+06	1.65e+08	1.25e+10	3.08e+13	4.85e+16	3.61e+18
	VMT1	9.39e+05	7.17e+06	1.61e+08	1.22e+10	3.00e+13	4.78e+16	3.53e+18
	VMT2	9.40e+05	8.45e+06	1.76e+08	1.24e+10	3.05e+13	4.80e+16	3.53e+18
	VMT3	9.40e+05	7.21e+06	1.60e+08	1.23e+10	3.04e+13	4.79e+16	3.53e+18
	VMT4	9.37e+05	7.11e+06	1.58e+08	1.22e+10	3.00e+13	4.74e+16	3.47e+18
	VMT5	9.38e+05	7.12e+06	1.60e+08	1.21e+10	2.98e+13	4.74e+16	3.48e+18
	VB1	9.39e+05	8.40e+06	1.76e+08	1.23e+10	3.04e+13	4.80e+16	3.53e+18
VB2	9.38e+05	8.35e+06	1.77e+08	1.21e+10	3.00e+13	4.75e+16	3.49e+18	
VB3	9.40e+05	8.45e+06	1.74e+08	1.25e+10	3.09e+13	4.85e+16	3.58e+18	
VB4	9.20e+05	7.78e+06	1.59e+08	1.21e+10	3.00e+13	4.66e+16	3.40e+18	
CPS	VHT	4.23e+07	1.17e+08	6.37e+08	1.46e+10	2.70e+13	4.93e+16	3.89e+18
	VSYG	9.50e+05	6.67e+06	1.62e+08	1.44e+10	2.70e+13	4.93e+16	3.89e+18
	VMC1	4.25e+07	1.18e+08	6.55e+08	1.64e+10	2.77e+13	5.02e+16	3.89e+18
	VMC2	9.50e+05	6.69e+06	1.62e+08	1.44e+10	2.79e+13	4.18e+16	3.16e+18
	VD1	9.39e+05	6.31e+06	1.57e+08	1.41e+10	2.68e+13	4.83e+16	3.81e+18
	VD2	9.40e+05	6.32e+06	1.57e+08	1.42e+10	2.68e+13	4.83e+16	3.81e+18
	VD3	9.40e+05	6.79e+06	1.60e+08	1.44e+10	2.68e+13	4.84e+16	3.82e+18
	VFP	9.40e+05	6.33e+06	1.58e+08	1.45e+10	2.71e+13	4.93e+16	3.89e+18
	VBe	9.49e+05	6.65e+06	1.62e+08	1.42e+10	2.70e+13	4.93e+16	3.89e+18
	VTi	9.50e+05	7.10e+06	1.62e+08	1.44e+10	2.70e+13	4.93e+16	3.89e+18
	VMT1	9.40e+05	6.30e+06	1.57e+08	1.40e+10	2.68e+13	4.82e+16	3.81e+18
	VMT2	9.41e+05	6.79e+06	1.60e+08	1.44e+10	2.68e+13	4.84e+16	3.82e+18
	VMT3	9.40e+05	6.32e+06	1.57e+08	1.42e+10	2.68e+13	4.83e+16	3.81e+18
	VMT4	9.37e+05	6.29e+06	1.57e+08	1.39e+10	2.66e+13	4.75e+16	3.75e+18
	VMT5	9.38e+05	6.29e+06	1.56e+08	1.38e+10	2.67e+13	4.76e+16	3.76e+18
	VB1	9.40e+05	6.77e+06	1.60e+08	1.42e+10	2.68e+13	4.83e+16	3.81e+18
VB2	9.39e+05	6.74e+06	1.59e+08	1.40e+10	2.67e+13	4.77e+16	3.77e+18	
VB3	9.41e+05	6.80e+06	1.61e+08	1.45e+10	2.70e+13	4.90e+16	3.86e+18	
VB4	9.22e+05	6.54e+06	1.57e+08	1.39e+10	2.61e+13	4.66e+16	3.67e+18	
Rand. Sys.	VHT	4.88e+07	1.32e+08	6.63e+08	1.64e+10	2.68e+13	5.40e+16	4.16e+18
	VSYG	9.49e+05	6.98e+06	1.66e+08	1.50e+10	2.70e+13	5.44e+16	4.19e+18
	VMC1	4.99e+07	1.32e+08	6.66e+08	1.88e+10	2.71e+13	5.35e+16	4.01e+18
	VMC2	9.50e+05	7.00e+06	1.64e+08	1.50e+10	2.77e+13	5.56e+16	4.46e+18
	VD1	9.38e+05	6.41e+06	1.58e+08	1.50e+10	2.63e+13	5.30e+16	4.09e+18
	VD2	9.38e+05	6.42e+06	1.59e+08	1.50e+10	2.64e+13	5.31e+16	4.09e+18
	VD3	9.39e+05	6.64e+06	1.63e+08	1.50e+10	2.66e+13	5.31e+16	4.09e+18
	VFP	9.39e+05	6.43e+06	1.59e+08	1.50e+10	2.69e+13	5.40e+16	4.16e+18
	VBe	9.48e+05	6.78e+06	1.66e+08	1.50e+10	2.70e+13	5.43e+16	4.19e+18
	VTi	9.49e+05	6.98e+06	1.66e+08	1.50e+10	2.70e+13	5.44e+16	4.19e+18
	VMT1	9.38e+05	6.41e+06	1.58e+08	1.51e+10	2.61e+13	5.29e+16	4.08e+18
	VMT2	9.39e+05	6.64e+06	1.63e+08	1.51e+10	2.66e+13	5.31e+16	4.09e+18
	VMT3	9.38e+05	6.42e+06	1.59e+08	1.50e+10	2.64e+13	5.31e+16	4.09e+18
	VMT4	9.36e+05	6.41e+06	1.58e+08	1.50e+10	2.61e+13	5.22e+16	4.01e+18
	VMT5	9.37e+05	6.40e+06	1.58e+08	1.50e+10	2.59e+13	5.23e+16	4.03e+18
	VB1	9.38e+05	6.63e+06	1.63e+08	1.50e+10	2.64e+13	5.31e+16	4.09e+18
VB2	9.37e+05	6.60e+06	1.62e+08	1.50e+10	2.61e+13	5.23e+16	4.03e+18	
VB3	9.39e+05	6.65e+06	1.64e+08	1.50e+10	2.68e+13	5.39e+16	4.14e+18	
VB4	9.21e+05	6.52e+06	1.60e+08	1.46e+10	2.60e+13	5.12e+16	3.94e+18	

Table 3.16: Summary statistics for the distribution of Monte Carlo MSE of different variance estimators for the Sukhatme population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max
Brewer	V _{HT}	6.93e+10	1.53e+11	7.30e+11	1.56e+13	1.74e+15	4.49e+18	1.96e+20
	V _{SYG}	4.53e+09	1.92e+10	4.78e+11	1.56e+13	1.78e+15	4.36e+18	2.08e+20
	V _{MC1}	5.87e+10	1.34e+11	6.54e+11	1.58e+13	1.76e+15	4.96e+18	1.92e+20
	V _{MC2}	4.73e+09	1.92e+10	4.67e+11	1.54e+13	1.78e+15	5.29e+18	1.61e+20
	V _{D1}	5.05e+09	1.74e+10	4.30e+11	1.41e+13	1.53e+15	3.71e+18	1.68e+20
	V _{D2}	5.16e+09	1.77e+10	4.37e+11	1.42e+13	1.57e+15	3.82e+18	1.71e+20
	V _{D3}	5.38e+09	1.82e+10	4.60e+11	1.43e+13	1.62e+15	4.01e+18	1.73e+20
	V _{F_P}	4.78e+09	1.87e+10	4.64e+11	1.42e+13	1.67e+15	4.49e+18	1.67e+20
	V _{Be}	4.90e+09	1.89e+10	4.64e+11	1.51e+13	1.74e+15	4.26e+18	1.93e+20
	V _{Ti}	4.76e+09	1.93e+10	4.82e+11	1.52e+13	1.76e+15	4.36e+18	1.95e+20
	V _{MT1}	5.19e+09	1.71e+10	4.21e+11	1.42e+13	1.49e+15	3.63e+18	1.70e+20
	V _{MT2}	5.35e+09	1.82e+10	4.59e+11	1.43e+13	1.61e+15	4.01e+18	1.73e+20
	V _{MT3}	5.14e+09	1.77e+10	4.36e+11	1.42e+13	1.56e+15	3.82e+18	1.71e+20
	V _{MT4}	4.90e+09	1.69e+10	4.17e+11	1.32e+13	1.45e+15	3.40e+18	1.48e+20
	V _{MT5}	5.04e+09	1.67e+10	4.07e+11	1.37e+13	1.43e+15	3.36e+18	1.58e+20
	V _{B1}	5.25e+09	1.79e+10	4.53e+11	1.42e+13	1.58e+15	3.89e+18	1.71e+20
V _{B2}	5.85e+09	1.72e+10	4.27e+11	1.38e+13	1.47e+15	3.51e+18	1.60e+20	
V _{B3}	4.76e+09	1.86e+10	4.75e+11	1.47e+13	1.70e+15	4.29e+18	1.82e+20	
V _{B4}	3.84e+09	1.58e+10	3.83e+11	1.04e+13	1.30e+15	3.05e+18	1.07e+20	
Tillé	V _{HT}	5.27e+10	1.17e+11	7.60e+11	1.62e+13	1.68e+15	4.28e+18	1.82e+20
	V _{SYG}	4.99e+09	1.73e+10	4.31e+11	1.66e+13	1.61e+15	3.91e+18	2.18e+20
	V _{MC1}	5.26e+10	1.17e+11	7.64e+11	1.62e+13	1.68e+15	4.44e+18	1.85e+20
	V _{MC2}	5.00e+09	1.74e+10	4.19e+11	1.67e+13	1.60e+15	4.04e+18	1.84e+20
	V _{D1}	5.07e+09	1.78e+10	4.04e+11	1.37e+13	1.45e+15	3.55e+18	1.61e+20
	V _{D2}	5.18e+09	1.81e+10	4.16e+11	1.38e+13	1.48e+15	3.66e+18	1.62e+20
	V _{D3}	5.38e+09	1.86e+10	4.40e+11	1.39e+13	1.53e+15	3.83e+18	1.64e+20
	V _{F_P}	4.82e+09	1.91e+10	4.68e+11	1.40e+13	1.60e+15	4.28e+18	1.59e+20
	V _{Be}	4.94e+09	1.91e+10	4.50e+11	1.46e+13	1.62e+15	4.06e+18	1.89e+20
	V _{Ti}	4.81e+09	1.95e+10	4.64e+11	1.47e+13	1.64e+15	4.17e+18	1.91e+20
	V _{MT1}	5.20e+09	1.75e+10	3.94e+11	1.38e+13	1.41e+15	3.47e+18	1.60e+20
	V _{MT2}	5.35e+09	1.86e+10	4.38e+11	1.39e+13	1.52e+15	3.83e+18	1.65e+20
	V _{MT3}	5.15e+09	1.81e+10	4.15e+11	1.38e+13	1.48e+15	3.65e+18	1.63e+20
	V _{MT4}	4.92e+09	1.72e+10	3.82e+11	1.29e+13	1.37e+15	3.25e+18	1.41e+20
	V _{MT5}	5.05e+09	1.70e+10	3.74e+11	1.33e+13	1.35e+15	3.21e+18	1.48e+20
	V _{B1}	5.26e+09	1.83e+10	4.27e+11	1.38e+13	1.50e+15	3.71e+18	1.62e+20
V _{B2}	5.83e+09	1.76e+10	3.94e+11	1.34e+13	1.39e+15	3.35e+18	1.50e+20	
V _{B3}	4.80e+09	1.90e+10	4.62e+11	1.43e+13	1.61e+15	4.10e+18	1.75e+20	
V _{B4}	3.99e+09	1.58e+10	3.55e+11	1.01e+13	1.24e+15	2.92e+18	1.06e+20	
CPS	V _{HT}	5.81e+10	1.35e+11	6.61e+11	1.48e+13	1.85e+15	5.40e+18	1.57e+20
	V _{SYG}	4.79e+09	1.92e+10	4.62e+11	1.50e+13	1.87e+15	5.44e+18	1.51e+20
	V _{MC1}	5.82e+10	1.35e+11	6.66e+11	1.50e+13	1.85e+15	5.57e+18	1.56e+20
	V _{MC2}	4.74e+09	1.92e+10	4.71e+11	1.50e+13	1.85e+15	5.77e+18	1.67e+20
	V _{D1}	5.10e+09	1.74e+10	4.23e+11	1.40e+13	1.63e+15	4.45e+18	1.32e+20
	V _{D2}	5.22e+09	1.78e+10	4.30e+11	1.41e+13	1.66e+15	4.66e+18	1.34e+20
	V _{D3}	5.43e+09	1.82e+10	4.53e+11	1.42e+13	1.70e+15	4.87e+18	1.36e+20
	V _{F_P}	4.83e+09	1.87e+10	4.56e+11	1.39e+13	1.70e+15	5.40e+18	1.31e+20
	V _{Be}	4.96e+09	1.89e+10	4.57e+11	1.50e+13	1.84e+15	5.32e+18	1.49e+20
	V _{Ti}	4.81e+09	1.93e+10	4.75e+11	1.50e+13	1.86e+15	5.41e+18	1.51e+20
	V _{MT1}	5.24e+09	1.72e+10	4.14e+11	1.41e+13	1.58e+15	4.29e+18	1.35e+20
	V _{MT2}	5.40e+09	1.82e+10	4.52e+11	1.42e+13	1.70e+15	4.86e+18	1.36e+20
	V _{MT3}	5.19e+09	1.77e+10	4.29e+11	1.41e+13	1.66e+15	4.65e+18	1.34e+20
	V _{MT4}	4.94e+09	1.69e+10	4.10e+11	1.31e+13	1.51e+15	4.10e+18	1.16e+20
	V _{MT5}	5.08e+09	1.67e+10	4.03e+11	1.36e+13	1.51e+15	3.98e+18	1.25e+20
	V _{B1}	5.30e+09	1.79e+10	4.46e+11	1.41e+13	1.68e+15	4.72e+18	1.34e+20
V _{B2}	5.91e+09	1.73e+10	4.26e+11	1.36e+13	1.56e+15	4.26e+18	1.27e+20	
V _{B3}	4.81e+09	1.86e+10	4.67e+11	1.45e+13	1.78e+15	5.21e+18	1.42e+20	
V _{B4}	3.86e+09	1.58e+10	3.77e+11	1.02e+13	1.28e+15	3.71e+18	8.24e+19	
Rand. Sys.	V _{HT}	5.96e+10	1.31e+11	6.53e+11	1.58e+13	1.86e+15	4.31e+18	2.71e+20
	V _{SYG}	4.94e+09	1.95e+10	4.79e+11	1.57e+13	1.85e+15	4.45e+18	2.81e+20
	V _{MC1}	6.11e+10	1.38e+11	6.65e+11	1.60e+13	1.86e+15	4.21e+18	2.63e+20
	V _{MC2}	4.76e+09	1.95e+10	4.82e+11	1.55e+13	1.85e+15	4.57e+18	2.32e+20
	V _{D1}	5.10e+09	1.75e+10	4.08e+11	1.46e+13	1.62e+15	3.77e+18	2.39e+20
	V _{D2}	5.22e+09	1.78e+10	4.19e+11	1.48e+13	1.66e+15	3.83e+18	2.41e+20
	V _{D3}	5.45e+09	1.83e+10	4.41e+11	1.49e+13	1.71e+15	3.89e+18	2.44e+20
	V _{F_P}	4.81e+09	1.88e+10	4.68e+11	1.46e+13	1.79e+15	4.06e+18	2.37e+20
	V _{Be}	4.94e+09	1.91e+10	4.70e+11	1.56e+13	1.84e+15	4.41e+18	2.80e+20
	V _{Ti}	4.78e+09	1.95e+10	4.79e+11	1.56e+13	1.85e+15	4.47e+18	2.83e+20
	V _{MT1}	5.25e+09	1.72e+10	3.97e+11	1.48e+13	1.58e+15	3.77e+18	2.39e+20
	V _{MT2}	5.42e+09	1.82e+10	4.40e+11	1.49e+13	1.71e+15	3.90e+18	2.45e+20
	V _{MT3}	5.20e+09	1.78e+10	4.18e+11	1.47e+13	1.66e+15	3.84e+18	2.43e+20
	V _{MT4}	4.93e+09	1.70e+10	3.86e+11	1.37e+13	1.54e+15	3.32e+18	2.09e+20
	V _{MT5}	5.08e+09	1.67e+10	3.77e+11	1.42e+13	1.51e+15	3.49e+18	2.21e+20
	V _{B1}	5.32e+09	1.79e+10	4.28e+11	1.48e+13	1.68e+15	3.83e+18	2.41e+20
V _{B2}	5.95e+09	1.73e+10	3.95e+11	1.43e+13	1.56e+15	3.55e+18	2.24e+20	
V _{B3}	4.79e+09	1.87e+10	4.64e+11	1.52e+13	1.80e+15	4.13e+18	2.60e+20	
V _{B4}	3.68e+09	1.58e+10	3.58e+11	1.02e+13	1.33e+15	2.95e+18	1.57e+20	

3.3 CONCLUSIONS

The simulation study presented in this chapter compared the performances of the main approximate variance estimators with the behaviour of variance estimates obtained by means of the Monte Carlo approach.

The study confirmed the results available in literature concerning the approximate variance estimators, in particular their low bias when $CV(X)$ is low (Haziza et al., 2008). This behaviour may be explained by considering that when the variability of X is low, so is the variability of the inclusion probabilities π_i . The high homogeneity among the π_i leads to a high-entropy condition, which is when these estimators are more effective.

Simulation results suggested that Monte Carlo estimators may be preferable when the population is small, due to their lower bias, especially in those cases when approximate estimators show poor performances, namely when $CV(X)$ is high. In particular, if n is fixed, the choice should go for the v_{MC2} estimator, as it is more stable than v_{MC1} . In the class of Approximate variance estimators, v_{Ti} seems to be the best choice with small populations. With larger populations, the approximate estimators performed well overall, especially with designs with higher entropy, and generally had lower bias than Monte Carlo estimates. However, the latter behaved reasonably well when $CV(X) < 1$, in particular with Brewer sampling, under which showed lower bias than with the designs with higher entropy. The coefficient of variation of the target variable Y does not seem to influence the bias, however it appears to influence the relative stability, in combination with the coefficient of variation of X .

Thus, Monte Carlo estimators appear to be more effective with small populations, while approximate variance estimators seem preferable with larger ones. Among approximate estimators, v_{Ti} seems to be the best choice with small N , while estimators v_{FP} and v_{Be} perform generally well with larger populations.

As noted by Haziza et al. (2008), in stratified designs the bias of the approximate variance estimators may sum up over the strata and generate a considerable bias. In such cases, the Monte Carlo estimates may be a good alternative, provided that strata are not too large.

Finally, it should be considered that the bias of the Monte Carlo estimators can be reduced by increasing the number of replications performed to approximate the inclusion probabilities. This would naturally increase the computational time required, however, with modern computers and the use of dedicated routines, such as function `jip_MonteCarlo()` in the `jipApprox` R package, one can perform many millions Monte Carlo replicates in a relatively short time. This is especially true for real studies, where joint-inclusion probabilities must be computed only for sample units.

The Horvitz–Thompson estimator is sensitive to outliers because it is linear in the observed values. Thus, large observations have large influence on the total estimates, in particular when associated with small inclusion probabilities. Discarding an outlier always leads to variance reduction, however, if such outlier is a correct value, ignoring it would introduce bias in the estimation (Hulliger, 1995). Hence, particular care must be taken when identifying outliers.

In this chapter, the Robust Horvitz–Thompson estimator by Beaumont, Haziza, et al. (2013) will be introduced after giving a brief review of some of the finite–population influence measures available in literature. In particular, the conditional bias approach will be described, on which the robust total estimator is based.

Finally, variance estimation of the robust total estimator will be discussed. Some new MSE estimators will be defined, and a robust bootstrap procedure will be proposed.

4.1 INFLUENCE MEASURES

Hulliger (1995) and Barranco–Chamorro et al. (2007) proposed influence measures based on models. The latter developed a diagnostic procedure for local influence in survey sampling based on Cook (1986) distance and the second–order method by Wu and Luo (1993), while the former introduced an influence measure based on a sensitivity curve, obtained by expressing the Horvitz–Thompson estimator as a least squares functional. Hulliger’s sensitivity curve is given by

$$\widehat{SC}_i = \frac{y_i - \beta_{LS} x_i}{x_i \sum_{j \in s} \pi_j^{-1}}, \quad (4.1)$$

where $\beta_{LS} = \frac{\sum_{i \in s} y_i \pi_i^{-1}}{\sum_{i \in s} x_i \pi_i^{-1}}$.

On the other hand, Moreno–Rebollo et al. (1999) proposed a design–based approach using conditional bias to measure the influence of the i –th sample unit on the Horvitz–Thompson total. For the i –th sample unit, the conditional bias of the Horvitz–Thompson estimator, due to the inclusion of the i –th unit in the sample, is defined as

$$B_i^{HT} = E[\hat{\vartheta}_{HT} | \delta_i = 1] - \vartheta = \sum_{j \in U} \left(\frac{\pi_{ij}}{\pi_i \pi_j} - 1 \right) y_j, \quad (4.2)$$

which can be estimated by

$$\hat{B}_i^{\text{HT}} = \sum_{j \in s} \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_j \pi_{ij}} \right) y_j \quad (4.3)$$

and has the appealing property of being null when $\pi_i = 1$ (Beaumont, Haziza, et al., 2013).

By considering the quantities $A_i = \hat{B}_i^{\text{HT}} / \hat{\vartheta}$, Moreno–Rebollo et al. (1999) showed through an example that their influence measure is highly correlated with Hulliger’s sensitivity curve in equation (4.1).

Poisson sampling is an interesting special case because, reminding that in this case $\pi_{ij} = \pi_i \pi_j$, equation (4.2) reduces to

$$B_i^{\text{HT}} = \left(\frac{1}{\pi_i} - 1 \right) y_i. \quad (4.4)$$

Hence, conditional bias can be computed exactly under Poisson sampling.

4.2 ROBUST HORVITZ–THOMPSON ESTIMATOR

Beaumont, Haziza, et al. (2013) have used the idea of conditional bias as an influence measure to develop a Robust Horvitz–Thompson estimator:

$$\hat{\vartheta}_{\text{RHT}} = \hat{\vartheta}_{\text{HT}} - \sum_{i \in s} B_i^{\text{HT}} + \sum_{i \in s} \psi(B_i^{\text{HT}}) = \hat{\vartheta}_{\text{HT}} + \sum_{i \in s} \left[\psi(B_i^{\text{HT}}) - B_i^{\text{HT}} \right],$$

which is obtained by reducing the contribution of influential units by the quantity $(\psi(B_i^{\text{HT}}) - B_i^{\text{HT}})$, where $\psi(\cdot)$ is any bounded function such that $\psi(z) \approx z$ when z is close to 0. In general B_i^{HT} is unknown and must be estimated as in equation (4.3), which leads to

$$\hat{\vartheta}_{\text{RHT}} = \hat{\vartheta}_{\text{HT}} + \sum_{i \in s} \left[\psi(\hat{B}_i^{\text{HT}}, c) - \hat{B}_i^{\text{HT}} \right].$$

A typical choice for $\psi(\cdot)$ is the Huber function (Huber, 1964):

$$\psi(\hat{B}_i^{\text{HT}}, c) = \text{sign}(\hat{B}_i^{\text{HT}}) \min(|\hat{B}_i^{\text{HT}}|, c),$$

where c is a constant that determines the magnitude of the bias reduction operated by the robust estimator, where $c \in [B_{\min}^{\text{HT}}, B_{\max}^{\text{HT}}]$, with $B_{\min}^{\text{HT}} = \min(B_i^{\text{HT}})$ and $B_{\max}^{\text{HT}} = \max(B_i^{\text{HT}})$, for $i \in s$. In other words, c is a tuning constant that acts as a threshold for the bias–variance trade–off: if $B_i^{\text{HT}} > c$, unit i will be considered influential, and the Horvitz–Thompson estimate will be curbed by a quantity equal to $\psi(\hat{B}_i^{\text{HT}}, c) - \hat{B}_i^{\text{HT}}$. Otherwise, the conditional bias related to the i -th will produce no penalisation as, for $B_i^{\text{HT}} \leq c$, we have that $\psi(\hat{B}_i^{\text{HT}}, c) - \hat{B}_i^{\text{HT}} = \hat{B}_i^{\text{HT}} - \hat{B}_i^{\text{HT}} = 0$.

However, c is an arbitrary constant and must be estimated in some way. Beaumont, Haziza, et al. (2013) suggest to do so by minimising the maximum absolute conditional bias of the robust estimator. That is, to solve

$$\arg \min_c \left\{ \max \left(\left| \hat{B}_i^{\text{RHT}}(c) \right| ; i \in s \right) \right\},$$

where $B_i^{\text{RHT}}(c)$ is the conditional bias of the robust estimator, given unit i and a threshold c , and is given by

$$\begin{aligned} B_i^{\text{RHT}}(c) &= E[\hat{\vartheta}^{\text{RHT}}(c) \mid \delta_i = 1] - \vartheta = \\ &= B_i^{\text{HT}} + E \left[\sum_{j \in s} \{ \psi(\hat{B}_i^{\text{HT}}, c) - \hat{B}_i^{\text{HT}} \} \mid \delta_i = 1 \right]. \end{aligned}$$

The second term of the right-hand side of the previous equation can be estimated without bias by $\Delta(c) = \sum_{i \in s} [\psi(\hat{B}_i^{\text{HT}}, c) - \hat{B}_i^{\text{HT}}]$. Hence an estimator for $B_i^{\text{RHT}}(c)$ is

$$\hat{B}_i^{\text{RHT}}(c) = \hat{B}_i^{\text{HT}}(c) + \Delta(c). \quad (4.5)$$

The authors show that the optimum threshold c_{opt} must be such that

$$\Delta(c_{\text{opt}}) = -\frac{\hat{B}_{\min}^{\text{HT}} + \hat{B}_{\max}^{\text{HT}}}{2}.$$

Hence, the Robust Horvitz–Thompson estimator with optimum c is obtained as

$$\hat{\vartheta}_{\text{RHT}}(c_{\text{opt}}) = \hat{\vartheta}_{\text{HT}} + \Delta(c_{\text{opt}}) = \hat{\vartheta}_{\text{HT}} - \frac{1}{2} (\hat{B}_{\min}^{\text{HT}} + \hat{B}_{\max}^{\text{HT}}). \quad (4.6)$$

The $\hat{\vartheta}_{\text{RHT}}$ estimator is biased, however, the authors show through a simulation study that such bias is negligible with relatively large populations.

As for the estimation of the variability of $\hat{\vartheta}_{\text{RHT}}$, Beaumont, Haziza, et al. (2013) expressed the MSE as

$$\text{MSE}[\hat{\vartheta}_{\text{RHT}}] = \text{Var}[\hat{\vartheta}_{\text{RHT}}] + \left\{ E[\hat{\vartheta}_{\text{RHT}} - \hat{\vartheta}_{\text{HT}}]^2 + \text{Var}[\hat{\vartheta}_{\text{RHT}} - \hat{\vartheta}_{\text{HT}}] \right\},$$

for which they proposed the following estimator:

$$\text{mse}_o[\hat{\vartheta}_{\text{RHT}}] = v[\hat{\vartheta}_{\text{RHT}}] + \max \left\{ 0, (\hat{\vartheta}_{\text{RHT}} - \hat{\vartheta}_{\text{HT}})^2 + v[\hat{\vartheta}_{\text{RHT}} - \hat{\vartheta}_{\text{HT}}] \right\}.$$

4.3 ESTIMATING THE MSE OF THE ROBUST HORVITZ–THOMPSON ESTIMATOR

4.3.1 Analytic estimators

In this subsection, we will define a set of estimators for the MSE of the Robust Horvitz–Thompson estimator. Then, they will be analysed and compared to Beaumont, Haziza, et al. (2013) MSE estimator, mse_o , by means of a simulation study.

The MSE may be written as

$$\begin{aligned}
\text{MSE}[\hat{\vartheta}_{\text{RHT}}] &= E[\hat{\vartheta}_{\text{RHT}} - \vartheta]^2 = \\
&= E[\hat{\vartheta}_{\text{HT}} + \Delta(c) - \vartheta]^2 = E[(\hat{\vartheta}_{\text{HT}} - \vartheta) + \Delta(c)]^2 = \\
&= E[(\hat{\vartheta}_{\text{HT}} - \vartheta)^2 + \Delta(c)^2 + 2(\hat{\vartheta}_{\text{HT}} - \vartheta)\Delta(c)] = \\
&= E[\hat{\vartheta}_{\text{HT}} - \vartheta]^2 + E[\Delta(c)^2] + 2 E[(\hat{\vartheta}_{\text{HT}} - \vartheta)\Delta(c)] = \\
&= \text{Var}[\hat{\vartheta}_{\text{HT}}] + E[\Delta(c)^2] + 2 E[(\hat{\vartheta}_{\text{HT}} - \vartheta)\Delta(c)] = \\
&= \text{Var}[\hat{\vartheta}_{\text{HT}}] + \text{Var}[\Delta(c)] + \left\{ E[\Delta(c)] \right\}^2 + 2 E[(\hat{\vartheta}_{\text{HT}} - \vartheta)\Delta(c)],
\end{aligned} \tag{4.7}$$

where the last equality is obtained by considering that $\text{Var}[\Delta(c)] = E[\Delta(c)^2] - \{E[\Delta(c)]\}^2$. Alternatively, by considering the decomposition in variance plus square bias, the MSE may be written as

$$\begin{aligned}
\text{MSE}[\hat{\vartheta}_{\text{RHT}}] &= \text{Var}[\hat{\vartheta}_{\text{RHT}}] + \left\{ E[\hat{\vartheta}_{\text{RHT}} - \vartheta] \right\}^2 = \\
&= \left\{ \text{Var}[\hat{\vartheta}_{\text{HT}} + \Delta(c)] \right\} + \left\{ E[\hat{\vartheta}_{\text{HT}} + \Delta(c) - \vartheta] \right\}^2 = \\
&= \left\{ \text{Var}[\hat{\vartheta}_{\text{HT}}] + \text{Var}[\Delta(c)] + 2 \text{Cov}[\hat{\vartheta}_{\text{HT}}, \Delta(c)] \right\} + \left\{ E[\hat{\vartheta}_{\text{HT}}] + E[\Delta(c)] - \vartheta \right\}^2 = \\
&= \text{Var}[\hat{\vartheta}_{\text{HT}}] + \text{Var}[\Delta(c)] + 2 \text{Cov}[\hat{\vartheta}_{\text{HT}}, \Delta(c)] + \left\{ E[\Delta(c)] \right\}^2.
\end{aligned} \tag{4.8}$$

Moreover, it should be noted that the variability of the robust total estimator $\hat{\vartheta}_{\text{RHT}}$ is affected by two different components: the variability of the estimator itself, and that of the tuning constant c , which is estimated on the sample as explained in the previous section. Taking this into account, and considering the decomposition (Cochran, 1977, p. 276):

$$\text{Var}(\hat{\vartheta}) = E_1 \left\{ E_2[\hat{\vartheta} - \vartheta]^2 \right\} = \text{Var}_1 \left\{ E_2[\hat{\vartheta}] \right\} + E_1 \left\{ \text{Var}_2[\hat{\vartheta}] \right\} \tag{4.9}$$

one may write

$$\begin{aligned}
\text{MSE}[\hat{\vartheta}_{\text{RHT}}] &= \text{Var}[\hat{\vartheta}_{\text{RHT}}] + \left\{ E[\hat{\vartheta}_{\text{RHT}} - \vartheta] \right\}^2 = \\
&= E_p \left\{ E_c[\hat{\vartheta}_{\text{RHT}} - \vartheta]^2 \right\} + \left\{ E[\Delta(c)] \right\}^2 = \\
&= E_p \left\{ \text{Var}_c[\hat{\vartheta}_{\text{RHT}}] \right\} + \text{Var}_p \left\{ E_c[\hat{\vartheta}_{\text{RHT}}] \right\} + E_p \left\{ E_c[\Delta(c)] \right\}^2,
\end{aligned} \tag{4.10}$$

where Var_p and E_p are the design variance and design expectation, while Var_c and E_c are the variance and expectation taken over the c values.

From what precedes, the following estimators may be derived:

$$\begin{aligned} \text{mse}_0[\hat{\vartheta}_{\text{RHT}}] &= v[\hat{\vartheta}_{\text{RHT}}] + \max\left\{0, (\hat{\vartheta}_{\text{RHT}} - \hat{\vartheta}_{\text{HT}})^2 + v[\Delta(c)]\right\}, \\ \text{mse}_1[\hat{\vartheta}_{\text{RHT}}] &= v[\hat{\vartheta}_{\text{HT}}] + v[\Delta(c)] + \Delta(c)^2 + 2(\hat{\vartheta}_{\text{RHT}} - \hat{\vartheta}_{\text{HT}})\Delta(c), \\ \text{mse}_2[\hat{\vartheta}_{\text{RHT}}] &= v[\hat{\vartheta}_{\text{HT}}] + v[\Delta(c)] + \Delta(c)^2 + 2\text{cov}[\hat{\vartheta}_{\text{HT}}, \Delta(c)], \\ \text{mse}_{\text{pc0}}[\hat{\vartheta}_{\text{RHT}}] &= v[\hat{\vartheta}_{\text{RHT}}] + v_p\{e_c[\hat{\vartheta}_{\text{RHT}}]\} + \max\left\{0, (\hat{\vartheta}_{\text{RHT}} - \hat{\vartheta}_{\text{HT}})^2 + v[\Delta(c)]\right\}, \\ \text{mse}_{\text{pc1}}[\hat{\vartheta}_{\text{RHT}}] &= v[\hat{\vartheta}_{\text{RHT}}] + v_p\{e_c[\hat{\vartheta}_{\text{RHT}}]\} + \Delta(c)^2, \end{aligned}$$

where mse_0 is the estimator proposed by Beaumont, Haziza, et al. (2013), mse_1 and mse_2 are estimators of the MSE in equations (4.7) and (4.8), respectively, and mse_{pc0} , mse_{pc1} are estimators based on the decomposition (4.9) and equation (4.10). Additionally, the following estimator will be considered:

$$\text{mse}_3[\hat{\vartheta}_{\text{RHT}}] = v[\hat{\vartheta}_{\text{HT}}] + v[\Delta(c)] + 2\text{cov}[\hat{\vartheta}_{\text{HT}}, \Delta(c)] + 2\Delta(c)^2$$

In the estimators defined above, $v[\hat{\vartheta}_{\text{HT}}]$ is the variance of the Horvitz–Thompson estimator, $v[\hat{\vartheta}_{\text{RHT}}]$ is the Horvitz–Thompson variance of the robust estimator, expressed as

$$\hat{\vartheta}_{\text{RHT}} = \sum_{i \in s} \frac{\tilde{y}_i}{\pi_i^{-1}}$$

where $\tilde{y}_i = y_i + \pi_i[\psi(\hat{B}_i^{\text{HT}}, c) - \hat{B}_i^{\text{HT}}]$, and $v[\Delta(c)]$ is the variance of $\Delta(c) = \sum_{i \in s} \frac{\xi_i}{\pi_i}$, where $\xi_i = \pi_i\{\psi(\hat{B}_i^{\text{HT}}, c_{\text{opt}}) - \hat{B}_i^{\text{HT}}\}$. Moreover, $\text{cov}[\hat{\vartheta}_{\text{HT}}, \Delta(c)]$ is the covariance between the Horvitz–Thompson estimator and $\Delta(c)$, and is estimated by

$$\text{cov}[\hat{\vartheta}_{\text{HT}}, \Delta(c)] = \sum_{i \in s} \sum_{j \in s} \frac{(\pi_{ij} - \pi_i\pi_j)}{\pi_i\pi_j\pi_{ij}} y_i \xi_j$$

Finally, $v_p\{e_c[\hat{\vartheta}_{\text{RHT}}]\}$ can be estimated through the following procedure:

1. Given a sample s drawn from a population \mathcal{U} , perform B bootstrap resamples from s ;
2. For each bootstrap replicate $b = 1, \dots, B$:
 - a) Select a set of c values. For instance, take values equal to the set of conditional bias: $C = \{c_i \mid c_i = B_i^{\text{HT}}, i = 1, \dots, n\}$;
 - b) Compute the Robust Horvitz–Thompson estimator for every c value of the set C generated at the previous step, and denote it with $\hat{\vartheta}_{\text{RHT}}(c)$;
 - c) Compute $A_b = \frac{1}{n} \sum_{c \in C} \hat{\vartheta}_{\text{RHT}}(c)$.
3. Compute the bootstrap variance $v^* = \frac{1}{B-1} \sum_{b=1}^B (A_b - \bar{A})^2$, where $\bar{A} = \frac{1}{B} \sum_{b=1}^B A_b$.

Simulation study

For assessing the behaviour of the estimators defined above, we set up a Monte Carlo simulation with $K = 5000$ replicates.

Artificial asymmetric populations with different size and different percentages of outliers were generated similarly to Beaumont, Haziza, et al. (2013) by considering the model

$$y_i = (1 - u_i) | 2x_i + 3.7\sqrt{x_i} \varepsilon_i | + u_i w_i,$$

where $\varepsilon_i \sim N(0, 1)$, the w_i were generated from a Normal distribution with mean 1200 and standard deviation 200, and u_i were independently generated from a Bernoulli distribution with probability p equal to the desired proportion of outliers in the population. The values x_i of the size variable were generated from a Gamma distribution with mean 50 and variance 500.

Small and large populations were generated, with $N = 250, 500, 2500, 5000$ and with different amounts of outliers (0%, 2% and 5%). Samples were drawn with sampling fractions $f = 2\%, 5\%, 10\%$ of the population size, according to Poisson sampling. The choice of the Poisson sampling is convenient because, under this design, conditional bias can be computed exactly, thus there is no variability due to the estimation of B_i , giving more control on the study and better interpretation of the results. For estimators mse_{pc0} and mse_{pc1} , Chauvet bootstrap with $B = 100$ and $D = 5$ replications was employed for the estimation of the $v_p \{e_c[\hat{\vartheta}_{RHT}]\}$ component.

The estimators under study have been compared in terms of Relative Bias, expressed as

$$RB = 100 \times \frac{E_{MC}[mse_*(\hat{\vartheta}_{RHT})] - MSE_{MC}[\hat{\vartheta}_{RHT}]}{MSE_{MC}[\hat{\vartheta}_{RHT}]},$$

where $mse_*(\hat{\vartheta}_{RHT})$ is one of the estimators under study, and E_{MC} and MSE_{MC} are the Monte Carlo Expectation and Monte Carlo MSE.

Results are showed in table 4.1. Estimator mse_0 appears to be heavily biased, and to consistently underestimate the MSE. Estimator mse_2 also underestimates the MSE and is comparable to mse_0 when outliers are absent, but is less biased than mse_0 when the population has outliers. Estimator mse_3 shows reasonably low bias for large populations with $N = 2500, 5000$, and for populations without outliers, while it strongly overestimates under small populations with outliers. Taking into account the variability of the tuning constant c seems to improve the accuracy. Indeed, estimators mse_{pc0} and mse_{pc1} are overall less biased than mse_0 and mse_1 , respectively. Estimator mse_{pc1} performs reasonably well with large populations, however both estimators mse_{pc0} and mse_{pc1} show high bias when there are no outliers. Finally, mse_1 performs very poorly and should be avoided.

To sum up, estimator mse_2 should always be preferred to mse_0 , and estimator mse_3 seems to be a reasonable choice with large pop-

Table 4.1: Relative Bias of the MSE estimators of $\hat{\vartheta}_{RHT}$.

outliers	N	f	RB(mse ₀)	RB(mse ₁)	RB(mse ₂)	RB(mse ₃)	RB(mse _{pco})	RB(mse _{pcl})
0%	250	0.02	-0.77	82.91	5.29	21.94	93.11	99.17
		0.05	-15.45	15.56	-13.83	-7.43	59.96	61.59
		0.10	-10.15	5.37	-9.49	-6.40	65.42	66.08
	500	0.02	-18.34	24.98	-15.61	-6.78	55.12	57.84
		0.05	-9.12	9.72	-8.13	-4.41	65.72	66.71
		0.10	-2.38	7.95	-1.86	0.12	75.47	75.99
	2500	0.02	-0.75	9.60	-0.24	1.73	78.48	78.98
		0.05	-1.42	2.86	-1.22	-0.44	75.75	75.95
		0.10	-3.75	-1.56	-3.65	-3.26	71.23	71.33
	5000	0.02	-3.03	2.03	-2.81	-1.87	72.72	72.94
		0.05	-2.07	0.07	-1.98	-1.60	73.82	73.91
		0.10	-1.67	-0.54	-1.63	-1.43	74.63	74.67
2%	250	0.02	-19.48	289.47	22.90	70.26	60.80	103.18
		0.05	-28.95	167.92	-2.74	27.05	13.01	39.22
		0.10	-27.44	117.98	-9.14	12.81	8.30	26.60
	500	0.02	-33.28	190.85	-2.64	31.13	8.36	38.99
		0.05	-29.51	158.41	-4.23	23.64	0.21	25.49
		0.10	-28.07	117.65	-9.40	12.27	-2.47	16.20
	2500	0.02	-29.87	100.27	-13.05	6.38	-4.35	12.48
		0.05	-27.48	55.29	-17.76	-5.24	-5.96	3.76
		0.10	-22.21	34.67	-16.05	-7.35	-0.86	5.30
	5000	0.02	-29.05	86.59	-14.42	2.80	-7.57	7.06
		0.05	-26.94	44.21	-18.75	-7.99	-8.69	-0.50
		0.10	-18.61	31.75	-13.15	-5.47	0.33	5.78
5%	250	0.02	-25.56	291.16	18.17	66.14	45.17	88.90
		0.05	-34.12	154.43	-9.47	19.02	-2.58	22.07
		0.10	-31.15	102.74	-15.05	5.33	-6.94	9.16
	500	0.02	-36.47	205.53	-3.23	32.89	-2.33	30.90
		0.05	-34.07	131.31	-12.49	12.18	-13.63	7.95
		0.10	-29.34	89.85	-14.79	3.14	-11.37	3.18
	2500	0.02	-35.37	68.71	-22.77	-7.04	-22.34	-9.74
		0.05	-25.99	37.45	-19.09	-9.36	-14.03	-7.14
		0.10	-17.50	27.00	-12.95	-6.08	-4.55	0.00
	5000	0.02	-27.72	54.38	-18.33	-5.85	-17.74	-8.36
		0.05	-19.52	30.03	-14.36	-6.73	-9.91	-4.75
		0.10	-14.81	18.61	-11.49	-6.31	-4.48	-1.16

ulations. Nonetheless, further studies are needed to derive an estimator with improved performance and more consistent behaviour, especially with small populations.

In the next subsection an alternative approach for estimating the variability of $\hat{\vartheta}_{\text{RHT}}$ is discussed.

4.3.2 Bootstrap

Here, we propose a bootstrap approach for estimating the variance of the Robust Horvitz–Thompson estimator, to which we will refer as *robust bootstrap*. The method described below is inspired by the *robust bootstrap* proposed by Amado and Pires (2004) for infinite population bootstrap. The main idea is to perform a bootstrap procedure where the sample observations y_i are penalised by a function of the conditional bias.

The algorithm is described below. To illustrate our *robust bootstrap* procedure, we will consider Chauvet bootstrap (see section 2.5.1):

1. Given the original sample s with observations y_i , compute the Horvitz–Thompson estimator $\hat{\vartheta}_{\text{HT}}$ and the conditional bias B_i ;
2. Compute the relative conditional bias with regards to the Horvitz–Thompson estimator: $rb_i = B_i^{\text{HT}}/\hat{\vartheta}_{\text{HT}}$;
3. Estimate the optimum value \tilde{c} that minimizes the maximum of $|rb_i + \sum_{i \in s} [\psi(rb_i, c) - rb_i]|$, where $\psi(rb_i, c) = \text{sign}(rb_i) \min(|rb_i|, c)$ is the Huber function of rb_i . In practice, the optimum tuning constant \tilde{c} can be approximated by generating a grid of values in the interval $[\min(rb_i), \max(rb_i)]$ and by identifying the value that satisfies the optimal condition;
4. Penalise sample observations by a function of rb_i :

$$\tilde{y}_i = \max \left\{ 0, y_i \left(1 - [rb_i - \psi(rb_i, \tilde{c})] \right) \right\};$$

5. Repeat $\lfloor \pi_i^{-1} \rfloor$ times the pair (\tilde{y}_i, π_i) to create U^f , where $\lfloor z \rfloor$ indicates the greatest integer less than or equal to z ;
6. Perform a Poisson sampling on the original sample, drawing units with probability $\pi_i^{-1} - \lfloor \pi_i^{-1} \rfloor$, and call this sample U^c . Add these units to U^f to complete the pseudo-population, obtaining $U^* = U^f \cup U^c$;
7. Draw a large number B of bootstrap samples s^* from U^* through Poisson sampling. At each iteration compute the HT estimator $\hat{\vartheta}_b^* = \sum_{i \in s^*} \tilde{y}_i \pi_i^{-1}$, then compute the bootstrap variance as

$$v^* = \frac{1}{B-1} \sum_{b=1}^B \left(\hat{\vartheta}_b^* - \hat{\vartheta}^* \right)^2,$$

with $\hat{\vartheta}^* = \frac{1}{B} \sum_{b=1}^B \hat{\vartheta}_b^*$;

8. Repeat steps 6–7 a large number of times D to get v_1^*, \dots, v_D^* . Finally, estimate the variance of $\hat{\vartheta}_{\text{RHT}}$ by

$$v^* = \frac{1}{D} \sum_{d=1}^D v_d^*.$$

Note that, as this algorithm considers Chauvet bootstrap, which is employed under Poisson sampling, the conditional bias does not need to be estimated. Under different sampling designs (and different bootstrap procedures), the B_i^{HT} in the bootstrap algorithm must be replaced by \hat{B}_i^{HT} .

An alternative estimation method could be a Chauvet bootstrap performed using the robust total estimator, rather than $\hat{\vartheta}_{\text{HT}}$. The performances of both these approaches were assessed through an empirical study.

Simulation Study

A simulation study was carried out to study the behaviour of the robust bootstrap introduced above. Population data have been generated as in the previous simulation, with $N = 250, 500, 1000$ and with different percentages of outliers (0%, 2% and 5%). $K = 1000$ Monte Carlo samples were drawn from each population by means of Poisson sampling, with sampling fractions $f = 0.02, 0.10$. Bootstrap was performed with $B = 1000, D = 100$. Besides the robust bootstrap and the bootstrap with $\hat{\vartheta}_{\text{RHT}}$, a classic Chauvet bootstrap on the Horvitz–Thompson estimator was also performed for comparison.

Performances were evaluated by means of Monte Carlo Relative Bias. As the bootstrap procedures analysed aim to estimate the variability of different estimators, the Relative bias was computed in different ways. For the classic Chauvet bootstrap, Relative Bias is obtained with respect to the Horvitz–Thompson estimator:

$$\text{RB} = 100 \times \frac{E_{\text{MC}}[v^*(\hat{\vartheta}_{\text{HT}})] - \text{MSE}_{\text{MC}}[\hat{\vartheta}_{\text{HT}}]}{\text{MSE}_{\text{MC}}[\hat{\vartheta}_{\text{HT}}]}.$$

While for the robust bootstrap and the bootstrap with the robust estimator the Relative Bias was computed with regard to the robust total estimator:

$$\text{RB} = 100 \times \frac{E_{\text{MC}}[v^*(\hat{\vartheta}_{\text{RHT}})] - \text{MSE}_{\text{MC}}[\hat{\vartheta}_{\text{RHT}}]}{\text{MSE}_{\text{MC}}[\hat{\vartheta}_{\text{RHT}}]},$$

where $v^*(\hat{\vartheta}_{\text{HT}})$ and $v^*(\hat{\vartheta}_{\text{RHT}})$ are, respectively, the bootstrap variance of the Horvitz–Thompson and Robust Horvitz–Thompson estimators.

Table 4.2: Relative Bias of the bootstrap variance estimators under different scenario.

outliers	N	f	RB[$v_{bs}(\hat{\vartheta}_{HT})$]	RB[$v_{bs}(\hat{\vartheta}_{RHT})$]	RB[$v_{rbs}(\hat{\vartheta}_{RHT})$]
0	250	0.02	11.62	-9.13	-9.60
		0.10	3.23	-0.85	-1.17
	500	0.02	3.73	-6.06	-7.77
		0.10	0.89	0.16	0.25
	1000	0.02	1.78	-3.29	-3.78
		0.10	4.88	4.10	4.20
2	250	0.02	2.91	33.02	-5.91
		0.10	-8.01	11.20	7.65
	500	0.02	1.92	36.80	1.64
		0.10	9.48	23.46	23.57
	1000	0.02	5.30	38.78	16.96
		0.10	4.28	9.19	14.66
5	250	0.02	6.57	49.58	0.75
		0.10	-0.48	6.52	2.89
	500	0.02	3.00	43.11	6.33
		0.10	-3.51	-0.41	-0.69
	1000	0.02	0.74	15.74	2.14
		0.10	9.68	3.26	5.90

In table 4.2, results are presented by percentage of outliers, population size, and sampling fraction. The last three columns, from left to right, represent the relative bias of the bootstrap variance of the Horvitz–Thompson estimator ($v_{bs}[\hat{\vartheta}_{HT}]$), the bootstrap variance of the Robust Horvitz–Thompson estimator ($v_{bs}[\hat{\vartheta}_{RHT}]$), and the variance of the robust estimator computed under the robust bootstrap ($v_{rbs}[\hat{\vartheta}_{RHT}]$).

The use of the robust estimator in the classic bootstrap procedure is not suggested, as it returns very large bias under most scenarios. On the other hand, the proposed robust bootstrap shows reasonably low bias, overall comparable with that of Chauvet bootstrap, except for the scenarios with 2% of outliers in the population and $N = 1000$, and when $N = 500$ and $f = 0.1$, where the bias is large.

4.4 CONCLUSIONS

The Robust Horvitz–Thompson estimator $\hat{\vartheta}_{RHT}$ let us perform robust estimation of the Horvitz–Thomson total in a simple and intuitive way, by employing the appealing concept of conditional bias as a design–based influence measure.

The work of this chapter was dedicated to researching a proper estimator of the variability of $\hat{\vartheta}_{RHT}$, by pursuing two approaches. An analytic and a bootstrap approach were defined, with different results.

Further studies are certainly needed to refine the two approaches proposed, nonetheless the results of the study presented in these

pages suggest that the analytic estimator mse_3 may be reasonably reliable in estimating the MSE of $\hat{\vartheta}_{RHT}$ when the population is large, while the robust bootstrap should give acceptable results with small populations.

AN APPLICATION TO A REAL STUDY: ESTIMATING THE NUMBER OF ROAMING DOGS IN AN URBAN AREA

As was mentioned in the introduction to this thesis, the motivation of this PhD project comes from a real study that we carried out for estimating the number of roaming dogs in an area of the city of Palermo, Italy. This chapter will be dedicated to the application of the methods discussed in this thesis to such study. First, we will describe the characteristics and methodology of the survey. Then, estimates of the total number of roaming dogs in the first city district will be presented, and its variance will be estimated by employing some of the estimation methods that have been discussed throughout the thesis. Robust estimation will be performed as well.

5.1 BACKGROUND

The presence of roaming dogs in urban areas is an important issue that municipalities have to deal with, especially in big cities of Southern Italy, where such phenomenon is widespread and represents a source of risk for people and public security, as well as being a concern for public hygiene. In these areas, policy-makers need to undertake actions focused on keeping under control the growth of such population and aimed to reduce its size in the long run. In Italy, the main control technique applied is the capture–sterilization–release procedure which, to be effective, needs to be applied to most of the roaming dogs populating a given area (Jackman and Rowan, 2007). Knowing the roaming–dog–population size is thus necessary; however, despite its importance, the size estimation of roaming dogs is a little–covered subject.

Different sources presented figures about this phenomenon, but the methodology used is either omitted (partially or completely) or not statistically accurate. In Italy, among the attempts made to find a measure for this phenomenon, an example is a 2006 census arranged by the Italian Ministry of Health (Italian Ministry of Health, 2006). In accordance with this study, in that year around 590,000 roaming dogs were in Italy. However, only the dogs that spent some time in public kennels were considered and the estimation/counting procedures are unknown. More data have been released by Italian regions and animal associations, but without relevant results. In late 2010, the *World Society for the Protection of Animals* (WSPA, now *World Animal Protection*) released a report titled *Surveying roaming dog populations:*

guidelines on methodology, where a formal procedure to get a count of roaming dogs is proposed. WSPA suggests a way to elicit a sample of sub-areas from the area of interest and find an estimate of the total number of roaming dogs in that territory by means of a simple random sampling with systematic sample selection.

For these reasons, we carried out a survey in the city of Palermo, Italy, in collaboration with the local health authority, whose help was important to shape the survey procedure to the dog behaviour. The study consisted in two subsequent surveys and had two aims: first, to formalize a simple and reproducible survey sampling procedure, in order to obtain a statistically reliable estimate and be able to monitor the phenomenon through time in a consistent and comparable way; second, to provide local authorities with an estimate of the roaming-dog-population size and help them plan proper prevention measures.

Section 5.2 describes the characteristics of the two surveys and of the data collection procedure. Sampling procedure and estimation methodology are introduced in section 5.3, while estimation results are presented in section 5.4.

5.2 FEATURES OF THE SURVEY AND DATA COLLECTION

The study started in 2010 with a pilot survey. With the aim of setting up an effective methodology for the study, we first defined assumptions concerning the distribution of the dogs over territory and their presence along the day considering, together with veterinarians of the local public health authority, their nature and habits. These assumptions had naturally driven to the choices that defined the sampling design illustrated in this section.

However, the 2010 pilot survey came with some flaws, consisting mostly in lack of precision by data collectors. This led to some issues during the identification process of dogs observed at different times.

As our main focus was to assess the effectiveness of the method, in 2014 we set up a new survey with the same characteristics, except for data collection, which we updated with some modifications in order to remove, or at least reduce, non-sampling errors introduced by data collectors' imprecision.

The data collection procedure that we are going to describe is in some aspects similar to the method proposed in WSPA's report (World Society for the Protection of Animals, 2010).

5.2.1 *Data collection settings and procedure*

The surveys were carried out in the first district of Palermo, Sicily, a 249.7-hectares-large area, which was divided into 76 sub-areas, defined in such a way that each was large approximately 3 hectares and surrounded by a path that could be covered during data collec-

tion. The number of roaming dogs was observed in a sample of 12 sub-areas, selected according to a *FPDUST* sampling, which will be described in section 5.3.

Data collection occurred during the second week of June, due to its stable weather; per each area in the sample, data have been gathered at 13:00 and 20:00 on Monday, Thursday and Sunday. These choices are in line with the World Organization for Animal Health (2010) guidelines and they are aimed to maximize the chances of observing all roaming dogs which dwelt the sample areas.

At the times indicated, each person in charge for data collection visited the assigned location, where they followed a given path. Paths were designed to have approximately equal length among all areas and had been kept fixed over all observations. Data collection sessions lasted at least thirty minutes and the path were completed at least twice throughout each observation session. Data collectors were veterinarians and volunteers of the public local kennel, as well as components of local animal associations. In order to have consistent data entries over all the observations, each person was assigned to one specific area. Although being advised of the importance of collecting data with careful attention to any distinctive trait of the observed dogs and being asked to report all of them in the form, during the pilot survey of 2010, some of the field workers partially or completely neglected this task. The result was that a number of observations could not be identified as belonging to the same dog with a sufficient degree of confidence, leading to inaccurate data.

In 2014, the survey was repeated with exact conditions, the only addition made to this procedure was taking a photo of the observed roaming dogs, made possible by the larger diffusion of devices capable of taking pictures of sufficient quality.

Collected information has then been manipulated to get a consistent dataset of the overall presence of dogs per each sampled area, consisting in a dataset in which the total number of observed roaming dogs was recorded for each unit of the sample. In the next subsection we will talk about the form used to register observations and how we changed it to improve precision in 2014 survey.

5.2.2 *Data collection form*

In both surveys, each data collector received six forms to fill in with observed information, one per each observation time.

The 2010 form included information about gender, size and health conditions (good, bad) of the dogs observed, as well as a section for the observer to write down any other significant information he would have found. As said in the previous Section, the last point had been quite neglected, thus some changes were implemented in the 2014 survey. Dog coat colour, a field to indicate whether or not a pic-

ture of the observed dog had been taken and one to write down the picture file name were added to the observed variables. Furthermore, the section dedicated to additional information was reorganised to be easier to fill in and highlighted to be simply reminded to the field workers. In addition, a meeting was held with veterinarians and future field workers in order to explain the details of the data collection procedure.

Applying these simple changes, we achieved a much higher quality of data, with no more issues in identifying different observations of the same dog over different times, and thus improving precision. Figures 5.1 and 5.2 show the form for 2014 survey.

5.3 SAMPLING PROCEDURE

The sample was drawn according to the PPS FPDUST spatial sampling (Barabesi et al., 1997), which is a probability-proportional-to-size (PPS) version of the FPDUST sampling proposed by Fattorini and Ridolfi (1997). The use of this design was appealing for our study as it modifies the probability of selecting spatially contiguous units by a parameter β .

Denote with Y the variable of interest, with X the size variable, and with y_i and x_i their values for the i -th unit in population U . PPS FPDUST design draws a sample through the following procedure:

1. Select the first unit with probability

$$p_i = \frac{x_i}{\sum_{i \in U} x_i}$$

2. Select the k -th unit, $k = 2, \dots, n$, with probability

$$p_{i|i_1, \dots, i_{k-1}} = \begin{cases} \frac{(1-\beta)x_i}{\sum_{i \notin s_k} x_i - \beta \sum_{i \in c(s_k)} x_i}, & i \in c(s_k) \\ \frac{x_i}{\sum_{i \notin s_k} x_i - \beta \sum_{i \in c(s_k)} x_i}, & \text{otherwise} \end{cases} \quad (5.1)$$

where $s_k = \{i_1, \dots, i_{k-1}\}$ is the set of units selected up to step k , $|s_k|$ is its cardinality and $c(s_k)$ is the set of non-sampled units which are contiguous to at least one sampled unit. β is a parameter that modifies selection probabilities for units belonging to $c(s_k)$ and must take values in the interval $(-\infty, 1)$ to ensure the existence of the design; when $\beta = 0$, one obtains the simple random sampling.

In our study, the size variable X is represented by the walkable area of each unit, that is, the total area excluding buildings and closed spaces such as private parks and parkings, or any surface where dogs could not be found. We have set $\beta = 0.5$, in order to penalize selection of contiguous units, but not too much as sampling close units would have resulted in reduced costs.

Compilare in stampatello

Rilevatore: Nome _____ Cognome _____	Data: _____ Ora inizio: _____ Ora fine: _____
---	---

Sede Rilevazione: _____ IB _____

Presenza rifiuti 'liberi' SÌ NO

Condizioni meteorologiche _____

	Taglia	Sesso	Stato di Salute	Colore	Segni particolari (se presenti, specificare sul retro)	Foto
1	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
2	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
3	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
4	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
5	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
6	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
7	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
8	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
9	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
10	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
11	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
12	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
13	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
14	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
15	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
16	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
17	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
18	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____
19	<input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> G	<input type="checkbox"/> M <input type="checkbox"/> F <input type="checkbox"/> NS	<input type="checkbox"/> S <input type="checkbox"/> M <input type="checkbox"/> NS	<input type="checkbox"/> N <input type="checkbox"/> B <input type="checkbox"/> M <input type="checkbox"/> Altro _____	<input type="checkbox"/> Sì <input type="checkbox"/> No	<input type="checkbox"/> No <input type="checkbox"/> Sì _____

Figure 5.1: Data collection form from 2014 survey – front

Legenda

Taglia: P = piccolo, M = medio, G = grande;

Sesso: M = maschio, F = femmina, NS = non so;

Stato di salute: S = sano, M = malato, NS = non so

Colore: N = nero, B = bianco, M = marrone, Altro;

Età: A = adulto, C = cucciolo;

Segni particolari: qualsiasi informazione aggiuntiva (razza, microchip, ...);

Foto: Specificare se è stata scattata una foto o meno. In caso affermativo, inserire il nome del file nello spazio apposito accanto al campo "SI".

Segni particolari:

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____
11. _____
12. _____
13. _____
14. _____
15. _____
16. _____
17. _____
18. _____
19. _____

Figure 5.2: Data collection form from 2014 survey – back

5.4 ESTIMATION

Estimation of the total number of roaming dogs in the urban area under study can be carried out through the Horvitz–Thompson estimator. In order to reduce the effect of outliers on the estimates, and especially for the 2010 study, robust estimation of the total will also be performed by using the robust estimator defined in equation (4.6).

Unfortunately, for FPDUST sampling exact inclusion probabilities are not available, not even of the first order. However, they can be accurately approximated by using the Monte Carlo approach described in section 2.7.

In our study, first and second–order inclusion probabilities were estimated through Monte Carlo simulation with ten million replicates, which guarantees negligible bias of the Horvitz–Thompson estimator. Indeed, considering the upper bound equation (2.20) by Fattorini (2006), and a conservative value of 0.01 for $\pi_0 = \min(\pi_i)$, we can rely on the fact that the bias of the Horvitz–Thompson estimator will be lower than 10^{-5} .

Different variance estimation methods were applied, among those described throughout the thesis. For the Horvitz–Thompson total, variance estimates were obtained by means of the Monte Carlo method and Tillé’s approximate variance estimator v_{T_i} which, the use of these estimators is justified by the good performances showed with small populations in the simulation study presented in chapter 3. Additionally, these estimates will be compared to bootstrap variance estimates obtained by means of the Holmberg pseudo–population algorithm (see section 2.5.1), with $B = 1000$ and $D = 100$. For the robust estimator $\hat{\delta}_{RHT}$, estimates were obtained through the robust bootstrap proposed in section 4.3.2. The robust bootstrap used was based on the Holmberg algorithm, with $B = 1000$ and $D = 100$. The conditional bias was estimated by using Monte Carlo inclusion probabilities.

Computations were performed through the R software and the packages sampling by Tillé and Matei (2016), Frames2 by Arcos et al. (2015), and the self–written packages jipApprox, bootstrapFP, fpdust, and robustHT (see appendix C).

Table 5.1 presents sample data for the two studies of years 2010 and 2014. The table reports the unit ID, the size measure X , namely the walkable area of each unit, the Monte Carlo inclusion probabilities $\tilde{\pi}_i$, and the observed number of roaming dogs in the 2010 and 2014 study. The two last rows of table 5.1 show the Horvitz–Thompson and the Robust Horvitz–Thompson total estimates.

It can be seen that the number of observed dogs in 2010 presents some values that are noticeably higher than others, in particular unit 9. As we mentioned earlier, indeed, the first study was affected by errors during data collection, which nonetheless seem to have been

Table 5.1: Sample data for the roaming dogs studies carried out in 2010 and 2014.

ID	X	$\tilde{\pi}_i$	Y	
			2010	2014
1	3.17	0.31	5	6
2	2.74	0.29	15	7
3	1.20	0.13	5	1
4	1.13	0.14	3	0
5	1.14	0.13	1	3
6	2.94	0.24	14	9
7	1.46	0.16	1	2
8	2.07	0.22	6	4
9	1.07	0.12	36	13
10	2.66	0.25	8	2
11	1.28	0.15	6	5
12	1.14	0.12	12	1
		$\hat{\vartheta}_{HT}$	692	298
		$\hat{\vartheta}_{RHT}$	608	273

successfully reduced in the 2014 survey by the modifications applied to the data collection form. In fact, besides a general reduction in the presence of roaming dogs, which was likely caused by the interventions of local health authorities, the observed values seem more regular. Unit 9 still exhibits an unusually high value, however it is now much closer to the general distribution of the observations, by analysing data collection forms it seems to be a correct outlier, probably due to a particular concentration of dogs in that area.

The estimated total number of roaming dogs in the first district of Palermo, Italy, more than halved in the four-year-long period between the two studies. The reduction appears too large to be due only to outliers, and this is confirmed by the robust estimates, which show a very similar trend. Hence, it seems that the interventions made by the local health authorities to stop the spread of the dog population were effective.

The square root of the estimates of the variance for the Horvitz–Thompson total and the robust total estimator are presented in table 5.2. As for $\hat{\vartheta}_{HT}$, the Monte Carlo approach and Tillé’s approximate variance estimators returned very similar results, which could be expected as they both showed good performances in the simulation study described in chapter 3. On the other hand, the bootstrap returns higher estimates, which could be due to the presence of outliers. Finally, the standard errors for $\hat{\vartheta}_{RHT}$ produced by the robust

Table 5.2: Variance estimates for the Horvitz–Thompson and the Robust Horvitz–Thompson total estimators, using different approaches.

Estimation method	$\sqrt{v[\hat{\vartheta}]}$	
	2010	2014
HT estimation		
Monte Carlo	258.17	92.28
Approximate variance estimators	260.00	92.43
Bootstrap	302.29	104.84
RHT estimation		
Robust Bootstrap	263.30	94.26

bootstrap are similar to the Monte Carlo and approximate variance estimates for $\hat{\vartheta}_{HT}$.

5.5 CONCLUSIONS

The widespread of roaming–dog populations in urban areas is an important issue for local administrations. Despite being essential for a proper planning of health and security actions, no reliable and reproducible procedure for the estimation of such phenomenon have been adopted yet.

With this study we have proposed a first attempt to define a simple, reproducible survey sampling procedure to estimate the total of roaming dogs on an urban area. The procedure proposed was first implemented through a pilot survey on a reduced portion of the city of Palermo, Italy. Some imprecisions in data collection were identified, so a few years later the study was replicated with same characteristics and an enhanced data collection form.

A sample of sub–areas from the first district of Palermo has been drawn by means of a FPDUST spatial sampling with selection probabilities proportional to a size variable. For this sampling design exact inclusion probabilities are not available, however they can be easily estimated through Monte Carlo simulation.

The main estimation techniques discussed in this thesis were employed to obtain variance estimates for the total of roaming dogs. Moreover, robust estimation was possible by means of the Robust Horvitz–Thompson total estimator $\hat{\vartheta}_{RHT}$ and the robust bootstrap.

A reduction in the estimated size of the roaming–dog population was registered over the time period between the two studies. This was expected due to the interventions promoted by the health authority after the results of the first survey.

Although the work for this thesis let us obtain variance estimates for the this survey and the results were satisfactory for local health

authorities, further studies are required to improve the efficiency of the survey, for example by defining stratification variables and better size measures to improve precision. Also, the sampling design might be replaced with the *doubly balanced spatial sampling* proposed by Grafström and Tillé (2012). More work is also needed to extend the survey to larger areas; this could be done, depending on available human and financial resources, either by including all the city districts in the study or by means of a two-stage design with districts as primary units and their sub-areas as second-stage units.

CONCLUSIONS

This thesis focused on variance estimation for the Horvitz–Thompson total estimator in Unequal Probability Sampling, with particular focus on small population and the presence of outliers.

The motivation for this project came from a study carried out in Palermo, Italy, which proposed a survey procedure for estimating the roaming–dog total in the first city district. The study involved an Unequal Probability Sampling design, the PPS FPDUST spatial sampling, which however presented two issues. First, variance could not be directly estimated because exact inclusion probabilities are not available for this design. Second, observations presented some outliers.

By analysing the solutions available in literature, two of them appeared particularly appealing for their general applicability: approximate variance estimators and the Monte Carlo method. Approximate estimators are known to perform well with large populations and under high–entropy designs, while other scenarios have not been explored broadly. The Monte Carlo method is extremely flexible and asymptotically unbiased as the number of replicates increases, however its behaviour has not been largely studied under empirical scenarios.

We hypothesised that these two methods could be used complementarily, in the sense that approximate variance estimators should probably be preferred with large populations and with high–entropy designs, while the Monte Carlo method could behave better with smaller populations. This hypothesis was assessed through an extensive simulation study that considered a large number of scenarios, different for population and sampling size, coefficient of variation of the target variable and size measure, and for the sampling designs employed. Overall, the study confirmed our hypothesis, and suggested that Monte Carlo estimates should be preferred with small populations, while approximate estimators generally have better behaviour with larger ones.

Then, to reduce the sensitivity to outliers of our estimates, the Robust Horvitz–Thompson estimator was considered. An estimator for its MSE was proposed, that in a simulation study showed to be generally less biased than the one provided by the authors of the robust estimator. However, the bias was still too large for the estimator to be considered reliable. Another MSE estimator was proposed that had relatively low bias with large populations, but performed poorly with smaller ones. Finally, a robust bootstrap approach was proposed, that

in our simulations assumed reasonably low bias with small populations.

Finally, the results of the thesis were applied to the roaming-dog survey. Variance estimates for the Horvitz-Thompson total have been obtained through Monte Carlo approach and approximate variance estimators and robust estimation have been achieved by employing the Robust Horvitz-Thompson estimator and the robust bootstrap.

SAMPLING PROCEDURES

In this appendix are presented a few sampling procedures that were mentioned but not described in the body of the thesis.

A.1 SIMPLE RANDOM SAMPLING WITHOUT REPLACEMENT

Given a population $U = \{1, \dots, N\}$ with N units, random sampling without replacement draws a sample s of size n with probability

$$p(s) = \frac{1}{\binom{N}{n}}$$

Each population unit is included in s with probability $f = n/N$, according to the following procedure. At each step k , $k = 0, \dots, n - 1$, draw one unit from the population with probability $p_k = \frac{1}{N-k}$, then remove the selected unit from the population and repeat, until the sample is complete.

A.2 STRATIFIED RANDOM SAMPLING

Sometimes population units can be grouped, or stratified, by means of additional information provided by a so called *stratification variable*. If it is possible to stratify the population in such a way that units are homogeneous within the strata and heterogeneous among them, then a stratified sampling is usually more efficient than a simple random sampling.

A stratified simple random sampling consists in dividing population units in H strata and perform a simple random sample inside each stratum.

For more information about stratified sampling and how to allocate units to strata see Cochran (1977).

A.3 BREWER METHOD

Brewer, 1975 introduced a method for the selection of a sample of size n with probability proportional to size and without replacement. The method has simple working selection probabilities and first order inclusion probabilities, but joint inclusion probabilities are hard to compute for $n > 2$.

A.3.1 *Derivation of working probabilities*

Let us define the following quantities:

P_I normed measure of size;

$P_I(r)$ working selection probability defined on the $(n - r + 1)$ th draw (the r -last draw);

The selection procedure described below is such that $\pi_I = nP_I$ and has working selection probabilities $P_I(r)$ such that the conditional probability of inclusion of the i -th unit in the last r draws, given that it was not selected in the $n - r$, is

$$\Pr\{I \in s[r, n] \mid I \notin s[1, r - 1]\} = \frac{rP_I}{\left(1 - \sum_{i=1}^{n-r} p_i\right)} \quad (\text{A.1})$$

where the p_i are the sample values of the P_I .

The conditional probabilities of inclusion in the last r draws are therefore proportional to the measures of size of the remaining (not yet selected) units.

The formulae for the working probabilities of selection are obtained by induction. Let's assume that the method works for $(n - 1)$; selecting the first unit with working probability $P_I(n)$, the probability of including the I th unit in the next $(n - 1)$ draws, given that it hasn't been selected yet, is

$$\pi_I = nP_I = \underbrace{P_I(n)}_{\text{probability of selecting unit I at first draw}} + \underbrace{\sum_{J \neq I}^N P_J(n) \frac{(n-1)P_I}{1 - P_J}}_{\text{conditional inclusion probability for unit I in the next } (n-1) \text{ draws}} \quad (\text{A.2})$$

Putting

$$U_I(n) = \frac{P_I(n)}{1 - P_I}$$

$$U(n) = \sum_{J=1}^N U_J(n)$$

equation (A.2) becomes (for the proof, see Appendix A.3.4):

$$U_I(n) = \frac{P_I[n - (n - 1)U(n)]}{(1 - nP_I)} \quad (\text{A.3})$$

summing over (A.3),

$$U(n) = [n - (n - 1)U(n)] \sum_{I=1}^N \left[\frac{P_I}{1 - nP_I} \right]$$

putting $T(n) = \sum_{I=1}^N [P_I/(1 - nP_I)]$ and collecting terms in $U(n)$,

$$U(n) = \frac{nT(n)}{1 + (n-1)T(n)}. \quad (\text{A.4})$$

Substituting (A.4) in (A.3) we obtain the working selection probabilities

$$P_I(n) = \alpha(n) \frac{P_I(1 - P_I)}{1 - nP_I} \quad (\text{A.5})$$

where

$$\alpha(n) = n[1 + (n-1)T(n)]^{-1} = \sum_{J=1}^N \frac{P_J(1 - P_J)}{1 - nP_J}. \quad (\text{A.6})$$

The $P_I(n)$ sum to unity and each lies in the range $0 < P_I(n) < 1$, as $0 < nP_I < 1$ and $P_I(n) < nP_I = \pi_I$.

A.3.2 Sampling procedure

The procedure for selecting a sample of size n from population U with Brewer's method consists in selecting, at each selection step k , a unit from the remaining ones with probability

$$\frac{P_i(1 - P_i)}{D^{(k)}(1 - rP_i)}, \quad i \in U^{(k)}$$

where $r = n - k + 1$ and $U^{(k)}$ is the population at step k , consisting of population U minus the units already included in the sample. Moreover,

$$D^{(k)} = \sum_{j \in U^{(k)}} \frac{P_j(1 - P_j)}{1 - rP_j}$$

Thus, sample selection procedure is given by:

- Select the first unit with probability

$$P_i^{(1)} = \frac{P_i(1 - P_i)}{D_1(1 - 1P_i)}, \quad \text{with} \quad D_1 = \sum_{i=1}^N \frac{P_i(1 - P_i)}{1 - nP_i}$$

- Let I_1 be the unit selected at first step. At the second step ($k=2$), select unit i with probability

$$P_i^{(2)} = \frac{P_i(1 - P_i)}{D_2(1 - 1P_i)}, \quad \text{with} \quad D_2 = \sum_{i \notin \{I_1\}} \frac{P_i(1 - P_i)}{1 - (n-1)P_i}$$

⋮
⋮
⋮

- At last step ($k = n$), select one unit with probability

$$P_i^{(n)} = \frac{P_i(1 - P_i)}{D_n(1 - P_i)}, \quad \text{with} \quad D_n = \sum_{i \notin \{I_1, \dots, I_{n-1}\}} \frac{P_i(1 - P_i)}{1 - P_i}$$

A.3.3 Joint inclusion probabilities

For this selection procedure, joint probabilities of inclusion are as follows:

$$\pi_{IJ}(n) = P_I(n) \frac{\pi_J(n-1)}{1-P_I} + P_J(n) \frac{\pi_I(n-1)}{1-P_J} + \sum_{K \neq I, J}^N P_K(n) \pi_{IJ}^{(K)}(n-1) \quad (\text{A.7})$$

where $\pi_j(n-1) = (n-1)P_j$ is the unconditional probability of inclusion of the J th unit in a sample of $(n-1)$ units and $\frac{\pi_J(n-1)}{1-P_I} = \frac{(n-1)P_J}{1-P_I}$ is the conditional probability of inclusion of the J th unit in the last $(n-1)$ draws, given that the I th unit has been selected at first draw (see equation (A.1)). $\pi_{IJ}^{(K)}$ is the joint probability of inclusion of the I th and J th units in the remaining $(n-1)$ units, given that the K th unit was selected first. Equation (A.7) can be rewritten as:

$$\pi_{IJ}(n) = (n-1)[U_I(n)P_J + U_J(n)P_I] + \sum_{K \neq I, J}^N P_K(n) \pi_{IJ}^{(K)}(n-1) \quad (\text{A.8})$$

Formula (A.8) can be applied recursively, until when only a single unit remains to be drawn, at which point the final term disappears.

However, all possible samples of $(n-2)$ from the $(N-2)$ units other than the I th and the J th must be considered when calculating π_{IJ} . This sets a serious limit to the usefulness of the method for $n > 2$.

A.3.4 Proofs

- Proof of equation (A.3):

$$nP_I = P_I(n) + \sum_{J \neq I}^N P_J(n) \frac{(n-1)P_I}{1-P_J};$$

$$nP_I = \underbrace{P_I(n)}_{U_I(n)(1-P_I)} + \sum_{J \neq I}^N \frac{P_J(n)}{1-P_J} (n-1)P_I;$$

$$nP_I = U_I(n)(1-P_I) + (n-1)P_I \sum_{J \neq I}^N \frac{P_J(n)}{1-P_J};$$

$$nP_I = U_I(n)(1-P_I) + (n-1)P_I \sum_{J \neq I}^N U_J(n);$$

$$nP_I = U_I(n)(1-P_I) + (n-1)P_I[U(n) - U_I(n)];$$

$$nP_I = U_I(n)(1-P_I) + (n-1)P_I U(n) - (n-1)P_I U_I(n);$$

$$nP_I = U_I(n)[(1-P_I) - P_I(n-1)] + (n-1)P_I U(n);$$

$$U_I(n) = \frac{nP_I - (n-1)P_I U(n)}{1-P_I - nP_I + P_I};$$

$$U_I(n) = \frac{P_I[n - (n-1)U(n)]}{1 - nP_I}$$

- Proof of equation (A.6):

$$\begin{aligned}
[\alpha]^{-1} &= \frac{1 + (n-1)T(n)}{n} = \frac{nT(n) - T(n) + 1}{n} = \\
&= \frac{1}{n} \left[n \sum_{I=1}^N \frac{P_I}{1 - nP_I} - \sum_{I=1}^N \frac{P_I}{1 - nP_I} + \underbrace{1}_{\sum P_I} \right] = \\
&= \frac{1}{n} \left[n \sum_{I=1}^N \frac{P_I}{1 - nP_I} - \sum_{I=1}^N \frac{P_I}{1 - nP_I} + \sum_{I=1}^N P_I \right] = \\
&= \frac{1}{n} \left[\sum_{I=1}^N \frac{nP_I - P_I + P_I(1 - nP_I)}{1 - nP_I} \right] = \\
&= \frac{1}{n} \left[\sum_{I=1}^N \frac{nP_I - P_I + P_I - nP_I^2}{1 - nP_I} \right] = \\
&= \frac{1}{n} \left[\sum_{I=1}^N \frac{nP_I - nP_I^2}{1 - nP_I} \right] = \frac{1}{n} \left[\sum_{I=1}^N \frac{nP_I(1 - P_I)}{1 - nP_I} \right] = \\
&= \sum_{I=1}^N \frac{P_I(1 - P_I)}{1 - nP_I}.
\end{aligned}$$

A.4 RAO-SAMPFORD SAMPLING

Rao-Sampford sampling (Rao, 1965; Sampford, 1967) is an UPS rejective procedure with fixed size, that is, units are selected with unequal probabilities and with replacement and, if any unit is selected more than once, the entire sample is rejected and the procedure starts over. As any rejective design, it can be computationally very demanding, but it has some advantages (Haziza et al., 2008):

- Inclusion probabilities π_i are exactly proportional to the size variable;
- Joint-inclusion probabilities π_{ij} are strictly positive, which is necessary for the existence of an unbiased variance estimator of the Horvitz-Thompson total;
- The condition $\pi_i\pi_j - \pi_{ij} > 0$ is satisfied, which ensure that the Sen-Yates-Grundy variance is nonnegative.

After computing probabilities $p_i = x_i / \sum_{i \in \mathcal{U}} x_i$, $\forall i \in \mathcal{U}$, the sampling procedure selects unit i at the first step with probability p_i , and the successive k units, $k = 0, \dots, n-1$, with probability $p_i / (1 - np_i)$ and with replacement. If any unit appears in the sample more than once, the sample is discarded and the procedure starts over.

A.5 RANDOMISED SYSTEMATIC SAMPLING

Randomised Systematic sampling is a variant of the Systematic design where population units are randomly ordered before sampling (step 1 of the sampling procedure shown below). Systematic sampling, which is represented by steps 2–4 of the following procedure, has often lower variance than other designs included the SRS, but it can produce biased variance estimates as many joint-inclusion probabilities can be null.

Given a set of inclusion probabilities π_i , the Randomised Systematic sampling procedure is:

1. Sort the list of population units in random order;
2. For each $k = 0, \dots, N$, compute values

$$W_k = \sum_{i=1}^k \pi_i$$

clearly, $W_0 = 0$ and $W_N = n$;

3. Generate a random value u from an Uniform distribution on $(0, 1]$;
4. Select all units that satisfy

$$W_{k-1} \leq u + j \leq W_k, \quad j = 0, \dots, n - 1$$

Due to randomisation of the list, this design has high-entropy, but joint-inclusion probabilities are too complex to compute. However, Hartley and Rao (1962) derived an approximation for π_{ij} of order $O(N^{-4})$ (see section 2.6).

RESULTS FROM THE SIMULATION STUDY ON
VARIANCE APPROXIMATIONS

In this appendix are included the tables that report the results of the extensive simulation described in chapter 3, which studies the behaviour of the approximate variance estimators and of variance estimates obtained through the Monte Carlo method.

Tables B.1 to B.6 give an overview of the results by presenting some summary statistics of the distribution of RB, RS and MSE over all scenarios, while tables B.7 to B.12 show these summaries divided by sampling design.

The results of the simulation study are reported in details by tables B.13 to B.62, which show the Monte Carlo Relative Bias (RB), Relative Stability (RS) and Mean Square Error (MSE) of the estimators considered, by N , $CV(X)$, $CV(Y)$, sampling design and sampling fraction.

In the tables, RB values greater than 5 and RS values greater than 1 are in bold. As the Horvitz–Thompson and Sen–Yates–Grundy variance estimators are design–unbiased, their RB was omitted. Furthermore, in some scenarios, under Tillé’s elimination procedure some null joint–inclusion probabilities were produced, leading to inadmissible values for the relative stability, which were omitted.

The nomenclature used for the approximate variance estimators is consistent with the notation used in chapter 2, while v_{MC1} and v_{MC2} indicate the estimators obtained by using Monte Carlo π_{ij} approximations with the Horvitz–Thompson and Sen–Yates–Grundy estimators, respectively.

Table B.1: Summary statistics for the distribution of Monte Carlo Relative Bias of different variance estimators for the MU273 population.

Estimator	min	5%	25%	50%	75%	95%	max
v _{HT}	—	—	—	—	—	—	—
v _{SYG}	—	—	—	—	—	—	—
v _{MC1}	0.01	0.08	0.27	0.57	1.16	2.66	17.14
v _{MC2}	0	0.05	0.23	0.5	0.93	3.8	26.3
v _{D1}	0	0.07	0.33	0.73	1.56	4.89	12.37
v _{D2}	0	0.06	0.27	0.67	1.43	4.95	12.58
v _{D3}	0.01	0.07	0.32	0.72	1.5	6.48	18.45
v _{FP}	0	0.05	0.3	0.62	1.24	5.03	12.83
v _{Be}	0	0.07	0.29	0.59	1.25	4.97	12.64
v _{Ti}	0	0.06	0.29	0.7	1.52	6.53	18.43
v _{MT1}	0	0.1	0.36	0.78	1.67	4.95	12.29
v _{MT2}	0.01	0.07	0.32	0.71	1.5	6.47	18.45
v _{MT3}	0	0.06	0.27	0.67	1.43	4.95	12.58
v _{MT4}	0.01	0.09	0.48	0.99	1.82	4.76	12.09
v _{MT5}	0.01	0.08	0.45	1	1.88	4.75	12.02
v _{B1}	0	0.06	0.3	0.72	1.47	6.33	18.22
v _{B2}	0	0.07	0.37	0.84	1.71	6.1	17.88
v _{B3}	0	0.03	0.28	0.69	1.6	6.66	18.56
v _{B4}	0.01	0.27	0.9	1.45	2.13	5.06	16.73

Table B.2: Summary statistics for the distribution of Monte Carlo Relative Bias of different variance estimators for the Sukhatme population.

Estimator	min	5%	25%	50%	75%	95%	max
v _{HT}	—	—	—	—	—	—	—
v _{SYG}	—	—	—	—	—	—	—
v _{MC1}	0	0.08	0.36	0.73	1.19	2.48	5.1
v _{MC2}	0	0.07	0.28	0.69	1.3	4.04	7.34
v _{D1}	0.02	0.22	1.2	2.92	6.15	10.33	11.5
v _{D2}	0.01	0.24	1.24	2.96	5.13	7.48	8.8
v _{D3}	0.01	0.17	1.1	2.29	3.56	6.06	12.4
v _{FP}	0.01	0.06	0.3	0.82	2.18	6.94	10.77
v _{Be}	0	0.1	0.47	0.98	1.9	5.67	8.64
v _{Ti}	0	0.07	0.32	0.79	1.43	4.3	11.95
v _{MT1}	0	0.25	1.48	3.79	6.82	11.38	12.89
v _{MT2}	0.03	0.17	1.05	2.34	3.6	6.12	12.35
v _{MT3}	0	0.24	1.22	2.97	5.22	7.64	9.05
v _{MT4}	0.02	0.41	1.61	4.76	9.74	13.99	14.97
v _{MT5}	0.01	0.36	1.75	4.22	9.42	14.62	16.13
v _{B1}	0	0.16	1.16	2.85	4.54	6.55	11.47
v _{B2}	0.01	0.21	1.6	4.57	7.24	11.09	12.39
v _{B3}	0	0.1	0.43	0.88	1.95	5.17	12.6
v _{B4}	3.04	8.75	12.55	14.99	18.32	24.55	29.14

Table B.3: Summary statistics for the distribution of Monte Carlo Relative Stability of different variance estimators for the MU273 population.

Estimator	min	5%	25%	50%	75%	95%	max
v _{HT}	0.87	0.96	1	1.01	2.06	73.94	654.26
v _{SYG}	1	1	1	1	1	1	1
v _{MC1}	0.86	0.97	1.01	1.11	3.61	97.2	661.75
v _{MC2}	0.54	0.94	1	1	1.02	1.13	2.04
v _{D1}	0.85	0.93	0.97	0.98	1	1.15	1.78
v _{D2}	0.86	0.93	0.98	0.98	1	1.15	1.79
v _{D3}	0.92	0.96	0.98	0.99	1	1.21	2.37
v _{FP}	0.86	0.94	0.99	1	1	1.15	1.76
v _{Be}	0.9	0.97	0.99	1	1	1.17	1.68
v _{Ti}	0.96	0.99	1	1	1.01	1.2	2.16
v _{MT1}	0.85	0.93	0.97	0.98	1	1.15	1.79
v _{MT2}	0.92	0.96	0.98	0.99	1	1.21	2.37
v _{MT3}	0.86	0.93	0.98	0.98	1	1.15	1.79
v _{MT4}	0.85	0.93	0.96	0.97	0.99	1.14	1.76
v _{MT5}	0.85	0.92	0.96	0.97	0.99	1.14	1.78
v _{B1}	0.92	0.96	0.98	0.99	1	1.2	2.36
v _{B2}	0.9	0.95	0.97	0.98	1	1.2	2.4
v _{B3}	0.93	0.96	0.99	1	1.01	1.21	2.32
v _{B4}	0.88	0.94	0.96	0.97	0.98	1.14	2.01

Table B.4: Summary statistics for the distribution of Monte Carlo Relative Stability of different variance estimators for the Sukhatme population.

Estimator	min	5%	25%	50%	75%	95%	max
v _{HT}	0.48	0.93	0.98	1.03	1.38	22.97	58.43
v _{SYG}	1	1	1	1	1	1	1
v _{MC1}	0.43	0.92	0.98	1.04	1.31	22.88	53.99
v _{MC2}	0.54	0.96	0.99	1	1.01	1.08	1.23
v _{D1}	0.37	0.77	0.85	0.89	0.97	1.12	1.25
v _{D2}	0.37	0.82	0.88	0.91	0.97	1.13	1.27
v _{D3}	0.38	0.86	0.91	0.94	1	1.15	1.34
v _{FP}	0.36	0.85	0.96	0.99	1.01	1.11	1.26
v _{Be}	0.43	0.93	0.97	0.99	1	1.08	1.3
v _{Ti}	0.43	0.94	0.99	1	1.01	1.08	1.36
v _{MT1}	0.37	0.75	0.83	0.88	0.97	1.16	1.28
v _{MT2}	0.38	0.86	0.91	0.94	1	1.15	1.34
v _{MT3}	0.37	0.82	0.88	0.9	0.97	1.12	1.27
v _{MT4}	0.32	0.71	0.79	0.85	0.94	1.08	1.23
v _{MT5}	0.34	0.69	0.78	0.84	0.95	1.14	1.24
v _{B1}	0.37	0.83	0.89	0.92	0.98	1.14	1.33
v _{B2}	0.35	0.75	0.82	0.86	0.97	1.21	1.38
v _{B3}	0.4	0.91	0.96	0.98	1.01	1.1	1.33
v _{B4}	0.24	0.56	0.7	0.75	0.8	0.85	1.03

Table B.5: Summary statistics for the distribution of Monte Carlo MSE of different variance estimators for the MU273 population.

Estimator	min	5%	25%	50%	75%	95%	max
V _{HT}	3.61e+07	1.15e+08	6.24e+08	1.31e+10	2.74e+13	5.20e+16	4.96e+18
V _{SYG}	9.49e+05	5.51e+06	1.59e+08	1.23e+10	2.74e+13	5.21e+16	5.02e+18
V _{MC1}	3.64e+07	1.11e+08	6.45e+08	1.76e+10	3.43e+13	5.29e+16	5.09e+18
V _{MC2}	9.46e+05	5.77e+06	1.61e+08	1.40e+10	2.78e+13	4.95e+16	6.57e+18
V _{D1}	9.38e+05	6.23e+06	1.58e+08	1.38e+10	2.73e+13	5.09e+16	4.86e+18
V _{D2}	9.38e+05	6.25e+06	1.58e+08	1.39e+10	2.74e+13	5.09e+16	4.86e+18
V _{D3}	9.39e+05	6.67e+06	1.61e+08	1.41e+10	2.76e+13	5.09e+16	4.87e+18
V _{FP}	9.39e+05	6.26e+06	1.59e+08	1.42e+10	2.79e+13	5.18e+16	4.96e+18
V _{Be}	9.48e+05	6.65e+06	1.62e+08	1.40e+10	2.77e+13	5.19e+16	4.95e+18
V _{Ti}	9.49e+05	7.04e+06	1.64e+08	1.41e+10	2.78e+13	5.20e+16	4.95e+18
V _{MT1}	9.38e+05	6.23e+06	1.58e+08	1.37e+10	2.72e+13	5.08e+16	4.86e+18
V _{MT2}	9.39e+05	6.67e+06	1.61e+08	1.41e+10	2.75e+13	5.09e+16	4.87e+18
V _{MT3}	9.38e+05	6.25e+06	1.58e+08	1.39e+10	2.74e+13	5.09e+16	4.87e+18
V _{MT4}	9.36e+05	6.22e+06	1.58e+08	1.36e+10	2.71e+13	5.00e+16	4.78e+18
V _{MT5}	9.37e+05	6.21e+06	1.58e+08	1.35e+10	2.70e+13	5.02e+16	4.80e+18
V _{B1}	9.38e+05	6.64e+06	1.61e+08	1.39e+10	2.74e+13	5.09e+16	4.87e+18
V _{B2}	9.37e+05	6.60e+06	1.60e+08	1.37e+10	2.71e+13	5.02e+16	4.80e+18
V _{B3}	9.39e+05	6.68e+06	1.62e+08	1.42e+10	2.78e+13	5.16e+16	4.93e+18
V _{B4}	9.20e+05	6.43e+06	1.58e+08	1.37e+10	2.70e+13	4.89e+16	4.68e+18

Table B.6: Summary statistics for the distribution of Monte Carlo MSE of different variance estimators for the Sukhatme population.

Estimator	min	5%	25%	50%	75%	95%	max
V _{HT}	5.27e+10	1.23e+11	7.08e+11	1.54e+13	1.85e+15	5.40e+18	2.71e+20
V _{SYG}	4.53e+09	1.73e+10	4.31e+11	1.59e+13	1.85e+15	5.25e+18	2.81e+20
V _{MC1}	5.26e+10	1.21e+11	6.68e+11	1.56e+13	1.85e+15	5.32e+18	2.63e+20
V _{MC2}	4.73e+09	1.74e+10	4.19e+11	1.56e+13	1.81e+15	5.39e+18	2.32e+20
V _{D1}	5.05e+09	1.74e+10	4.05e+11	1.38e+13	1.63e+15	4.40e+18	2.39e+20
V _{D2}	5.16e+09	1.77e+10	4.17e+11	1.39e+13	1.66e+15	4.65e+18	2.41e+20
V _{D3}	5.38e+09	1.82e+10	4.40e+11	1.40e+13	1.69e+15	4.87e+18	2.44e+20
V _{FP}	4.78e+09	1.87e+10	4.57e+11	1.40e+13	1.67e+15	5.40e+18	2.37e+20
V _{Be}	4.90e+09	1.89e+10	4.51e+11	1.47e+13	1.84e+15	5.20e+18	2.80e+20
V _{Ti}	4.76e+09	1.93e+10	4.65e+11	1.48e+13	1.86e+15	5.40e+18	2.83e+20
V _{MT1}	5.19e+09	1.71e+10	3.95e+11	1.39e+13	1.58e+15	4.24e+18	2.39e+20
V _{MT2}	5.35e+09	1.82e+10	4.39e+11	1.40e+13	1.69e+15	4.86e+18	2.45e+20
V _{MT3}	5.14e+09	1.77e+10	4.16e+11	1.39e+13	1.66e+15	4.64e+18	2.43e+20
V _{MT4}	4.90e+09	1.69e+10	3.83e+11	1.33e+13	1.51e+15	4.06e+18	2.09e+20
V _{MT5}	5.04e+09	1.67e+10	3.75e+11	1.36e+13	1.51e+15	3.93e+18	2.21e+20
V _{B1}	5.25e+09	1.79e+10	4.28e+11	1.39e+13	1.68e+15	4.72e+18	2.41e+20
V _{B2}	5.83e+09	1.72e+10	3.94e+11	1.36e+13	1.56e+15	4.24e+18	2.24e+20
V _{B3}	4.76e+09	1.86e+10	4.63e+11	1.43e+13	1.77e+15	5.21e+18	2.60e+20
V _{B4}	3.68e+09	1.58e+10	3.57e+11	1.02e+13	1.28e+15	3.71e+18	1.57e+20

Table B.7: Summary statistics for the distribution of Monte Carlo Relative Bias of different variance estimators for the MU273 population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max
Brewer	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0.02	0.09	0.21	0.4	1.18	2.78	17.14
	VMC ₂	0	0.06	0.21	0.4	0.9	1.78	9.35
	VD ₁	0	0.06	0.3	0.57	0.95	1.57	2.13
	VD ₂	0	0.03	0.19	0.49	0.9	1.38	1.81
	VD ₃	0.01	0.06	0.32	0.61	1.11	3.01	3.58
	VFP	0	0.02	0.22	0.45	0.83	1.14	1.6
	VBe	0.01	0.07	0.22	0.41	0.75	1.19	1.7
	VTi	0.02	0.06	0.25	0.64	1.12	2.89	3.52
	VMT ₁	0.01	0.07	0.32	0.56	1.15	1.82	2.39
	VMT ₂	0.01	0.06	0.32	0.61	1.11	3.01	3.58
	VMT ₃	0	0.03	0.19	0.49	0.9	1.38	1.81
	VMT ₄	0.01	0.07	0.41	0.78	1.37	2.22	2.49
	VMT ₅	0.01	0.08	0.39	0.76	1.55	2.33	2.75
	VB ₁	0.01	0.03	0.23	0.62	1.08	2.73	3.46
VB ₂	0.01	0.09	0.3	0.65	1.45	2.72	3.36	
VB ₃	0	0.02	0.25	0.63	1.04	3.08	3.56	
VB ₄	0.04	0.23	0.75	1.14	1.81	2.28	3.66	
Tillé	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0.02	0.09	0.35	0.72	1.02	2.68	7.29
	VMC ₂	0.01	0.04	0.25	0.57	0.89	6.64	25.05
	VD ₁	0.01	0.08	0.6	1.68	4.21	8.48	12.37
	VD ₂	0.01	0.09	0.58	1.76	4.27	8.6	12.58
	VD ₃	0.01	0.07	0.65	2.05	5.89	12.18	18.45
	VFP	0	0.06	0.6	2.15	4.63	8.8	12.83
	VBe	0	0.11	0.56	1.96	4.44	8.7	12.64
	VTi	0	0.11	0.7	2.27	5.79	12	18.43
	VMT ₁	0	0.13	0.58	1.53	4.3	8.45	12.29
	VMT ₂	0.02	0.07	0.65	2.05	5.89	12.18	18.45
	VMT ₃	0.01	0.08	0.58	1.76	4.27	8.6	12.58
	VMT ₄	0.01	0.12	0.59	1.62	4.1	8.23	12.09
	VMT ₅	0.03	0.11	0.54	1.44	4.16	8.2	12.02
	VB ₁	0.01	0.09	0.61	2.03	5.65	12.05	18.22
VB ₂	0.03	0.1	0.63	1.83	5.13	12.05	17.88	
VB ₃	0.01	0.07	0.64	2.29	5.82	12.06	18.56	
VB ₄	0.04	0.3	1.07	2.25	4.22	10.37	16.73	
CPS	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0.01	0.08	0.28	0.57	1.11	2.5	3.71
	VMC ₂	0.01	0.06	0.2	0.5	1.05	11.87	26.3
	VD ₁	0	0.03	0.2	0.47	0.74	1.69	2.08
	VD ₂	0	0.04	0.2	0.41	0.69	1.43	1.7
	VD ₃	0.01	0.06	0.27	0.63	1.2	3.52	4.96
	VFP	0	0.05	0.21	0.43	0.72	1.31	1.59
	VBe	0	0.03	0.2	0.44	0.69	1.33	1.61
	VTi	0.01	0.04	0.34	0.63	1.23	3.28	5
	VMT ₁	0.02	0.09	0.26	0.52	0.88	1.83	2.37
	VMT ₂	0.01	0.06	0.27	0.63	1.2	3.52	4.96
	VMT ₃	0	0.04	0.2	0.41	0.69	1.44	1.7
	VMT ₄	0.02	0.1	0.38	0.65	1.23	2.13	2.67
	VMT ₅	0.01	0.07	0.32	0.7	1.33	2.29	2.95
	VB ₁	0.01	0.06	0.26	0.66	1.2	3.27	4.76
VB ₂	0.02	0.07	0.39	0.8	1.34	3.05	4.45	
VB ₃	0.01	0.05	0.29	0.64	1.23	3.5	5.06	
VB ₄	0.01	0.1	0.88	1.35	1.88	2.57	3.44	
Rand. Sys.	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0.03	0.07	0.28	0.73	1.38	2.64	12.08
	VMC ₂	0	0.08	0.27	0.54	0.92	3.11	10.18
	VD ₁	0.01	0.12	0.5	1.08	1.88	3.61	5.14
	VD ₂	0.02	0.14	0.42	1.04	1.79	3.14	4.96
	VD ₃	0.02	0.08	0.25	0.57	0.92	1.58	2.3
	VFP	0.04	0.11	0.35	0.71	1.29	2.69	4.75
	VBe	0.02	0.1	0.37	0.84	1.39	2.74	4.82
	VTi	0	0.05	0.23	0.45	0.81	1.24	2.37
	VMT ₁	0.05	0.12	0.57	1.19	2.09	3.64	5.2
	VMT ₂	0.02	0.08	0.25	0.57	0.92	1.58	2.3
	VMT ₃	0.02	0.14	0.42	1.04	1.79	3.13	4.96
	VMT ₄	0.01	0.08	0.68	1.43	2.31	3.82	5.37
	VMT ₅	0.02	0.08	0.65	1.46	2.44	3.96	5.43
	VB ₁	0	0.07	0.25	0.58	1	1.88	2.3
VB ₂	0	0.06	0.31	0.72	1.57	2.47	2.99	
VB ₃	0	0.03	0.21	0.44	0.82	1.2	2.35	
VB ₄	0.12	0.44	0.99	1.45	1.99	2.85	3.16	

Table B.8: Summary statistics for the distribution of Monte Carlo Relative Bias of different variance estimators for the Sukhatme population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max
Brewer	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0.01	0.09	0.46	0.73	1.16	2.54	3.72
	VMC ₂	0.01	0.15	0.42	0.76	1.18	4.76	7.24
	VD ₁	0.02	0.28	1.08	3.24	6.34	9.84	11.33
	VD ₂	0.04	0.29	1.05	2.46	4.92	7.18	7.53
	VD ₃	0.04	0.15	1.06	2.01	3.34	5.01	5.9
	VFP	0.02	0.04	0.28	0.62	1.07	5.41	7.71
	VBe	0.01	0.11	0.52	0.78	1.39	2.62	6.24
	VTi	0	0.06	0.38	0.68	1.15	2	2.36
	VM T ₁	0.1	0.3	1.41	3.72	6.76	11.31	12.84
	VM T ₂	0.03	0.16	1.04	2.08	3.4	5.06	5.86
	VM T ₃	0.02	0.32	1.06	2.49	4.99	7.35	7.84
	VM T ₄	0.07	0.23	1.58	5.28	9.79	13.37	14.72
	VM T ₅	0.07	0.35	1.33	4.67	9.83	14.62	16.06
	VB ₁	0.01	0.16	1	2.38	4.04	6.37	6.57
VB ₂	0.05	0.22	1.53	4.72	6.78	11.09	11.31	
VB ₃	0	0.11	0.45	0.85	1.69	3.05	3.69	
VB ₄	9.75	11.15	12.56	15.14	17.89	24.91	27.41	
Tillé	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0.03	0.07	0.37	0.76	1.31	2.36	2.55
	VMC ₂	0.03	0.07	0.39	1.05	1.84	4.03	6.44
	VD ₁	0.02	0.22	1.57	2.65	5.63	9.1	11.34
	VD ₂	0.32	0.61	1.68	2.96	5.08	6.57	8.42
	VD ₃	0.1	0.37	1.75	2.94	4.36	8.47	12.4
	VFP	0.05	0.23	1.4	2.25	4.31	7.29	10.77
	VBe	0.04	0.23	1.03	2.06	3.71	6.83	8.64
	VTi	0.06	0.12	0.75	1.59	3.21	8.42	11.95
	VM T ₁	0.08	0.45	1.58	3.54	5.94	10.35	12.8
	VM T ₂	0.13	0.42	1.84	2.94	4.4	8.5	12.35
	VM T ₃	0.29	0.56	1.7	2.98	5.12	6.72	8.38
	VM T ₄	0.04	0.45	1.35	3.27	7.96	12.57	14.72
	VM T ₅	0.05	1.15	2.32	3.46	7.94	13.69	16.02
	VB ₁	0.02	0.61	2.12	3.33	4.93	7.58	11.47
VB ₂	0.14	0.73	2.65	3.96	7.15	10.17	11.6	
VB ₃	0.04	0.1	0.59	1.78	3.23	9.19	12.6	
VB ₄	3.04	5.61	9.73	13.6	17.75	24.48	27.71	
CPS	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0	0.08	0.3	0.72	1.17	1.86	5.1
	VMC ₂	0	0.08	0.25	0.44	1.17	3.11	3.86
	VD ₁	0.03	0.3	1.11	2.9	6.28	9.48	10.63
	VD ₂	0.02	0.15	1.02	2.8	4.85	7.83	8.06
	VD ₃	0.04	0.31	0.86	1.84	3.25	6.04	7.14
	VFP	0.01	0.03	0.23	0.5	1.16	5.15	9.06
	VBe	0.01	0.05	0.45	0.77	1.34	2.62	3
	VTi	0	0.1	0.27	0.66	1.1	2.2	2.56
	VM T ₁	0	0.15	1.37	3.77	7.35	10.97	12.21
	VM T ₂	0.03	0.31	0.86	1.85	3.3	6.11	7.15
	VM T ₃	0	0.12	0.99	2.79	4.9	7.83	8.12
	VM T ₄	0.15	0.6	1.92	5.36	9.76	14.04	14.36
	VM T ₅	0.03	0.4	1.65	5.18	9.72	14.29	15.45
	VB ₁	0	0.14	0.89	2.27	3.97	7.4	7.85
VB ₂	0.1	0.26	1.05	4.5	6.86	11.06	11.9	
VB ₃	0.02	0.13	0.43	0.74	1.35	2.91	5.28	
VB ₄	10.13	11.35	12.97	15.07	18.37	25.21	29.14	
Rand. Sys.	VHT	—	—	—	—	—	—	—
	VSYG	—	—	—	—	—	—	—
	VMC ₁	0.02	0.07	0.27	0.71	1.28	2.1	3.65
	VMC ₂	0	0.03	0.23	0.55	1.26	4.28	7.34
	VD ₁	0.03	0.14	1.26	3.69	6.58	10.69	11.5
	VD ₂	0.01	0.21	0.97	3.2	5.73	8.24	8.8
	VD ₃	0.01	0.07	0.66	2.54	3.73	5.53	7.1
	VFP	0.01	0.08	0.26	0.6	1.81	5.3	6.98
	VBe	0	0.09	0.31	0.79	1.34	4.51	5.97
	VTi	0.01	0.04	0.21	0.52	1.07	2.02	3.09
	VM T ₁	0.02	0.15	1.57	4.19	7.35	12.28	12.89
	VM T ₂	0.03	0.09	0.68	2.42	3.77	5.34	7.1
	VM T ₃	0.06	0.22	1.01	3.24	5.75	8.43	9.05
	VM T ₄	0.02	0.5	1.58	5.26	10.45	14.09	14.97
	VM T ₅	0.01	0.26	1.63	4.79	9.91	15.52	16.13
	VB ₁	0.01	0.16	0.57	3.17	4.98	6.1	7.37
VB ₂	0.01	0.14	1.02	5.06	8.27	11.41	12.39	
VB ₃	0.02	0.06	0.34	0.94	1.61	2.74	3.06	
VB ₄	10.35	11.2	13.22	15.54	18.08	24.49	27.48	

Table B.9: Summary statistics for the distribution of Monte Carlo Relative Stability of different variance estimators for the MU273 population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max	
Brewer	v _{HT}	0.92	0.96	1	1.01	2.17	72.32	644.61	
	v _{SYG}	1	1	1	1	1	1	1	
	v _{MC1}	0.86	0.97	1	1.08	3.2	90.13	642.14	
	v _{MC2}	0.82	0.93	0.99	1	1.03	1.28	1.56	
	v _{D1}	0.85	0.89	0.97	0.98	0.99	1	1.06	
	v _{D2}	0.86	0.89	0.97	0.98	0.99	1	1.06	
	v _{D3}	0.94	0.96	0.97	0.99	1	1.01	1.15	
	v _{FP}	0.86	0.9	0.98	0.99	1	1.01	1.06	
	v _{Be}	0.9	0.94	0.99	0.99	1	1	1.04	
	v _{Ti}	0.99	0.99	1	1	1	1.01	1.01	
	v _{MT1}	0.85	0.89	0.96	0.97	0.99	1.01	1.01	1.07
	v _{MT2}	0.94	0.96	0.97	0.99	1	1.01	1.15	
	v _{MT3}	0.86	0.89	0.97	0.98	0.99	1	1.06	
	v _{MT4}	0.85	0.89	0.96	0.97	0.98	1	1.06	
	v _{MT5}	0.85	0.88	0.95	0.97	0.98	1	1.06	
	v _{B1}	0.94	0.96	0.97	0.99	0.99	1.01	1.14	
v _{B2}	0.93	0.95	0.96	0.97	0.99	1.01	1.15		
v _{B3}	0.94	0.97	0.98	0.99	1	1.01	1.13		
v _{B4}	0.88	0.93	0.95	0.96	0.97	0.99	1.08		
Tillé	v _{HT}	0.87	0.96	1	1.05	2.21	72.69	603.09	
	v _{SYG}	1	1	1	1	1	1	1	
	v _{MC1}	0.87	0.97	1.02	1.15	5	102.02	609.86	
	v _{MC2}	0.54	0.96	1	1	1.02	1.1	1.17	
	v _{D1}	0.94	0.96	0.99	1.01	1.11	1.38	1.78	
	v _{D2}	0.94	0.97	0.99	1.01	1.11	1.38	1.79	
	v _{D3}	0.94	0.97	0.99	1.01	1.13	1.67	2.37	
	v _{FP}	0.95	0.98	1	1.02	1.12	1.38	1.76	
	v _{Be}	0.97	0.99	1	1.01	1.13	1.39	1.68	
	v _{Ti}	0.97	0.99	1	1.02	1.15	1.68	2.16	
	v _{MT1}	0.94	0.96	0.99	1	1.1	1.38	1.79	
	v _{MT2}	0.94	0.97	0.99	1.01	1.13	1.67	2.37	
	v _{MT3}	0.94	0.97	0.99	1.01	1.11	1.38	1.79	
	v _{MT4}	0.93	0.96	0.98	1	1.1	1.37	1.76	
	v _{MT5}	0.93	0.95	0.98	1	1.1	1.37	1.78	
	v _{B1}	0.94	0.96	0.99	1.01	1.13	1.66	2.36	
v _{B2}	0.93	0.95	0.98	1	1.12	1.65	2.4		
v _{B3}	0.95	0.97	1	1.02	1.14	1.66	2.32		
v _{B4}	0.92	0.94	0.97	0.99	1.1	1.48	2.01		
CPS	v _{HT}	0.92	0.96	1	1.01	1.88	70.62	635.98	
	v _{SYG}	1	1	1	1	1	1	1	
	v _{MC1}	0.92	0.96	1	1.09	3.4	88.02	642.73	
	v _{MC2}	0.78	0.82	1	1	1.01	1.07	1.17	
	v _{D1}	0.92	0.95	0.97	0.98	0.99	1	1.11	
	v _{D2}	0.92	0.95	0.98	0.98	0.99	1.01	1.12	
	v _{D3}	0.92	0.97	0.99	0.99	1	1.04	1.34	
	v _{FP}	0.92	0.95	0.98	1	1	1.01	1.11	
	v _{Be}	0.98	0.99	1	1	1	1	1	
	v _{Ti}	0.96	1	1	1	1.01	1.08	1.18	
	v _{MT1}	0.92	0.94	0.96	0.98	0.99	1.01	1.12	
	v _{MT2}	0.92	0.97	0.99	0.99	1	1.04	1.34	
	v _{MT3}	0.92	0.95	0.98	0.98	0.99	1.01	1.12	
	v _{MT4}	0.91	0.94	0.96	0.97	0.99	1	1.11	
	v _{MT5}	0.91	0.94	0.95	0.97	0.99	1	1.11	
	v _{B1}	0.92	0.97	0.98	0.99	1	1.04	1.33	
v _{B2}	0.9	0.96	0.97	0.98	1	1.04	1.35		
v _{B3}	0.93	0.98	0.99	1	1.01	1.04	1.32		
v _{B4}	0.9	0.94	0.96	0.97	0.97	1.01	1.21		
Rand. Sys.	v _{HT}	0.91	0.96	1	1	1.86	69.98	654.26	
	v _{SYG}	1	1	1	1	1	1	1	
	v _{MC1}	0.9	0.97	1.01	1.1	3.62	97.14	661.75	
	v _{MC2}	0.74	0.99	1	1.01	1.03	1.22	2.04	
	v _{D1}	0.9	0.93	0.96	0.98	0.99	1	1.09	
	v _{D2}	0.9	0.93	0.97	0.98	0.99	1	1.09	
	v _{D3}	0.93	0.95	0.98	0.99	1	1.01	1.13	
	v _{FP}	0.9	0.93	0.98	0.99	1	1	1.09	
	v _{Be}	0.95	0.97	0.99	1	1	1	1.06	
	v _{Ti}	1	1	1	1	1	1	1	
	v _{MT1}	0.9	0.92	0.96	0.97	0.99	1	1.09	
	v _{MT2}	0.93	0.95	0.98	0.99	1	1.01	1.13	
	v _{MT3}	0.9	0.93	0.97	0.98	0.99	1	1.09	
	v _{MT4}	0.9	0.92	0.96	0.97	0.98	1	1.09	
	v _{MT5}	0.89	0.92	0.95	0.96	0.98	1	1.09	
	v _{B1}	0.92	0.95	0.98	0.98	0.99	1.01	1.13	
v _{B2}	0.92	0.94	0.96	0.97	0.99	1.01	1.13		
v _{B3}	0.93	0.96	0.99	1	1	1.01	1.13		
v _{B4}	0.92	0.93	0.95	0.96	0.97	0.99	1.12		

Table B.10: Summary statistics for the distribution of Monte Carlo Relative Stability of different variance estimators for the Sukhatme population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max
Brewer	VHT	0.9	0.94	0.98	1.03	1.52	23.22	58.43
	VSYG	1	1	1	1	1	1	1
	VMC ₁	0.89	0.92	0.98	1.02	1.23	21.17	53.31
	VMC ₂	0.77	0.94	0.99	1.01	1.03	1.09	1.23
	VD ₁	0.81	0.82	0.85	0.88	0.95	1.17	1.21
	VD ₂	0.82	0.83	0.88	0.9	0.96	1.18	1.23
	VD ₃	0.83	0.84	0.91	0.94	1	1.19	1.26
	VFP	0.8	0.82	0.97	0.99	1.03	1.13	1.2
	VBe	0.93	0.93	0.97	0.98	1	1.1	1.12
	VTi	0.93	0.94	0.99	1	1.02	1.06	1.1
	VMT ₁	0.8	0.81	0.83	0.85	0.95	1.24	1.28
	VMT ₂	0.83	0.84	0.91	0.94	1	1.19	1.25
	VMT ₃	0.82	0.83	0.88	0.89	0.96	1.18	1.22
	VMT ₄	0.71	0.74	0.79	0.84	0.93	1.14	1.19
	VMT ₅	0.75	0.75	0.78	0.82	0.93	1.22	1.24
	VB ₁	0.82	0.83	0.89	0.91	0.98	1.19	1.24
VB ₂	0.77	0.78	0.81	0.84	0.96	1.27	1.38	
VB ₃	0.87	0.88	0.96	0.98	1.01	1.11	1.16	
VB ₄	0.52	0.56	0.71	0.76	0.8	0.84	0.86	
Tillé	VHT	0.48	0.85	1.03	1.07	1.4	23.08	53.8
	VSYG	1	1	1	1	1	1	1
	VMC ₁	0.43	0.85	1.01	1.08	1.44	23.07	53.99
	VMC ₂	0.54	0.96	0.99	1	1.01	1.03	1.06
	VD ₁	0.37	0.74	0.86	0.92	1.03	1.14	1.25
	VD ₂	0.37	0.74	0.89	0.94	1.04	1.15	1.27
	VD ₃	0.38	0.77	0.92	0.99	1.06	1.18	1.34
	VFP	0.36	0.82	0.95	1.01	1.08	1.17	1.26
	VBe	0.43	0.87	0.96	1	1.04	1.16	1.3
	VTi	0.43	0.88	0.97	1.01	1.07	1.18	1.36
	VMT ₁	0.37	0.73	0.83	0.9	1.04	1.15	1.26
	VMT ₂	0.38	0.76	0.92	0.99	1.06	1.17	1.34
	VMT ₃	0.37	0.75	0.89	0.94	1.04	1.15	1.27
	VMT ₄	0.32	0.65	0.81	0.86	0.99	1.12	1.23
	VMT ₅	0.34	0.68	0.8	0.85	1.01	1.12	1.24
	VB ₁	0.37	0.74	0.9	0.96	1.06	1.17	1.33
VB ₂	0.35	0.69	0.84	0.89	1.04	1.19	1.32	
VB ₃	0.4	0.81	0.96	1.01	1.06	1.18	1.33	
VB ₄	0.24	0.57	0.69	0.76	0.81	0.97	1.03	
CPS	VHT	0.88	0.93	0.98	1	1.22	20.86	50.62
	VSYG	1	1	1	1	1	1	1
	VMC ₁	0.89	0.94	0.98	1.02	1.24	20.9	50.74
	VMC ₂	0.96	0.98	1	1	1.01	1.11	1.17
	VD ₁	0.76	0.77	0.85	0.89	0.96	1.09	1.14
	VD ₂	0.82	0.82	0.87	0.9	0.97	1.1	1.15
	VD ₃	0.88	0.89	0.91	0.93	0.99	1.12	1.17
	VFP	0.86	0.87	0.96	0.99	1	1.03	1.09
	VBe	0.93	0.95	0.98	0.99	1	1.02	1.04
	VTi	0.96	0.97	1	1	1	1.02	1.04
	VMT ₁	0.73	0.75	0.82	0.88	0.97	1.15	1.19
	VMT ₂	0.88	0.89	0.91	0.93	0.99	1.12	1.17
	VMT ₃	0.82	0.82	0.87	0.9	0.97	1.1	1.15
	VMT ₄	0.7	0.71	0.78	0.84	0.94	1.06	1.12
	VMT ₅	0.68	0.69	0.78	0.84	0.95	1.12	1.16
	VB ₁	0.83	0.84	0.89	0.91	0.98	1.11	1.16
VB ₂	0.74	0.75	0.82	0.85	0.97	1.19	1.29	
VB ₃	0.93	0.94	0.95	0.98	1	1.04	1.06	
VB ₄	0.55	0.6	0.68	0.75	0.79	0.85	0.89	
Rand. Sys.	VHT	0.88	0.93	0.97	1	1.22	18.26	45.85
	VSYG	1	1	1	1	1	1	1
	VMC ₁	0.88	0.93	0.97	1	1.26	19.52	48.08
	VMC ₂	0.82	0.96	0.99	1	1	1.03	1.09
	VD ₁	0.77	0.81	0.85	0.89	0.94	1.08	1.13
	VD ₂	0.83	0.83	0.87	0.9	0.95	1.09	1.14
	VD ₃	0.86	0.87	0.91	0.93	0.98	1.1	1.17
	VFP	0.84	0.88	0.95	0.98	0.99	1.07	1.09
	VBe	0.94	0.95	0.98	0.99	1	1.01	1.03
	VTi	0.97	0.98	1	1	1	1.01	1.01
	VMT ₁	0.74	0.78	0.83	0.87	0.94	1.14	1.17
	VMT ₂	0.87	0.87	0.91	0.93	0.98	1.1	1.16
	VMT ₃	0.83	0.84	0.87	0.9	0.95	1.09	1.14
	VMT ₄	0.71	0.74	0.78	0.85	0.92	1.04	1.11
	VMT ₅	0.69	0.72	0.78	0.84	0.93	1.12	1.15
	VB ₁	0.84	0.85	0.89	0.91	0.96	1.1	1.15
VB ₂	0.75	0.76	0.81	0.85	0.96	1.19	1.28	
VB ₃	0.92	0.93	0.96	0.97	0.99	1.03	1.05	
VB ₄	0.56	0.59	0.7	0.74	0.78	0.83	0.85	

Table B.11: Summary statistics for the distribution of Monte Carlo MSE of different variance estimators for the MU₂₇₃ population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max
Brewer	V _{HT}	5.16e+07	1.33e+08	6.45e+08	1.39e+10	2.78e+13	4.55e+16	4.96e+18
	V _{SYG}	9.50e+05	7.16e+06	1.65e+08	1.38e+10	2.78e+13	4.60e+16	5.02e+18
	V _{MC1}	4.36e+07	1.22e+08	6.36e+08	1.64e+10	3.44e+13	4.60e+16	5.09e+18
	V _{MC2}	9.46e+05	6.84e+06	1.63e+08	1.37e+10	2.70e+13	6.08e+16	6.57e+18
	V _{D1}	9.39e+05	6.37e+06	1.58e+08	1.35e+10	2.73e+13	4.46e+16	4.86e+18
	V _{D2}	9.40e+05	6.38e+06	1.58e+08	1.36e+10	2.74e+13	4.47e+16	4.86e+18
	V _{D3}	9.41e+05	6.88e+06	1.62e+08	1.38e+10	2.76e+13	4.48e+16	4.87e+18
	V _{FP}	9.41e+05	6.40e+06	1.59e+08	1.39e+10	2.79e+13	4.55e+16	4.96e+18
	V _{Be}	9.50e+05	6.73e+06	1.64e+08	1.38e+10	2.77e+13	4.57e+16	4.95e+18
	V _{Ti}	9.51e+05	7.19e+06	1.65e+08	1.39e+10	2.78e+13	4.57e+16	4.95e+18
	V _{MT1}	9.40e+05	6.37e+06	1.58e+08	1.34e+10	2.71e+13	4.44e+16	4.86e+18
	V _{MT2}	9.41e+05	6.88e+06	1.62e+08	1.38e+10	2.75e+13	4.48e+16	4.87e+18
	V _{MT3}	9.40e+05	6.38e+06	1.58e+08	1.36e+10	2.74e+13	4.47e+16	4.87e+18
	V _{MT4}	9.38e+05	6.35e+06	1.57e+08	1.33e+10	2.71e+13	4.41e+16	4.78e+18
	V _{MT5}	9.38e+05	6.35e+06	1.57e+08	1.32e+10	2.69e+13	4.39e+16	4.80e+18
	V _{B1}	9.40e+05	6.85e+06	1.62e+08	1.36e+10	2.74e+13	4.47e+16	4.87e+18
V _{B2}	9.39e+05	6.82e+06	1.61e+08	1.34e+10	2.70e+13	4.40e+16	4.80e+18	
V _{B3}	9.41e+05	6.89e+06	1.63e+08	1.39e+10	2.78e+13	4.55e+16	4.93e+18	
V _{B4}	9.23e+05	6.61e+06	1.59e+08	1.34e+10	2.70e+13	4.37e+16	4.68e+18	
Tillé	V _{HT}	3.61e+07	7.94e+07	5.56e+08	8.26e+09	2.63e+13	5.01e+16	3.59e+18
	V _{SYG}	9.51e+05	5.28e+06	1.16e+08	6.96e+09	2.59e+13	4.98e+16	3.60e+18
	V _{MC1}	3.64e+07	8.91e+07	6.14e+08	1.90e+10	3.43e+13	5.01e+16	3.67e+18
	V _{MC2}	9.49e+05	5.39e+06	1.34e+08	1.22e+10	3.09e+13	4.20e+16	4.00e+18
	V _{D1}	9.39e+05	7.17e+06	1.59e+08	1.23e+10	3.03e+13	4.79e+16	3.53e+18
	V _{D2}	9.40e+05	7.21e+06	1.60e+08	1.23e+10	3.04e+13	4.79e+16	3.53e+18
	V _{D3}	9.40e+05	8.45e+06	1.76e+08	1.24e+10	3.05e+13	4.80e+16	3.53e+18
	V _{FP}	9.40e+05	7.22e+06	1.60e+08	1.25e+10	3.09e+13	4.88e+16	3.60e+18
	V _{Be}	9.49e+05	7.50e+06	1.62e+08	1.25e+10	3.07e+13	4.85e+16	3.60e+18
	V _{Ti}	9.50e+05	8.70e+06	1.65e+08	1.25e+10	3.08e+13	4.85e+16	3.61e+18
	V _{MT1}	9.39e+05	7.17e+06	1.61e+08	1.22e+10	3.00e+13	4.78e+16	3.53e+18
	V _{MT2}	9.40e+05	8.45e+06	1.76e+08	1.24e+10	3.05e+13	4.80e+16	3.53e+18
	V _{MT3}	9.40e+05	7.21e+06	1.60e+08	1.23e+10	3.04e+13	4.79e+16	3.53e+18
	V _{MT4}	9.37e+05	7.11e+06	1.58e+08	1.22e+10	3.00e+13	4.74e+16	3.47e+18
	V _{MT5}	9.38e+05	7.12e+06	1.60e+08	1.21e+10	2.98e+13	4.74e+16	3.48e+18
	V _{B1}	9.39e+05	8.40e+06	1.76e+08	1.23e+10	3.04e+13	4.80e+16	3.53e+18
V _{B2}	9.38e+05	8.35e+06	1.77e+08	1.21e+10	3.00e+13	4.75e+16	3.49e+18	
V _{B3}	9.40e+05	8.45e+06	1.74e+08	1.25e+10	3.09e+13	4.85e+16	3.58e+18	
V _{B4}	9.20e+05	7.78e+06	1.59e+08	1.21e+10	3.00e+13	4.66e+16	3.40e+18	
CPS	V _{HT}	4.23e+07	1.17e+08	6.37e+08	1.46e+10	2.70e+13	4.93e+16	3.89e+18
	V _{SYG}	9.50e+05	6.67e+06	1.62e+08	1.44e+10	2.70e+13	4.93e+16	3.89e+18
	V _{MC1}	4.25e+07	1.18e+08	6.55e+08	1.64e+10	2.77e+13	5.02e+16	3.89e+18
	V _{MC2}	9.50e+05	6.69e+06	1.62e+08	1.44e+10	2.79e+13	4.18e+16	3.16e+18
	V _{D1}	9.39e+05	6.31e+06	1.57e+08	1.41e+10	2.68e+13	4.83e+16	3.81e+18
	V _{D2}	9.40e+05	6.32e+06	1.57e+08	1.42e+10	2.68e+13	4.83e+16	3.81e+18
	V _{D3}	9.40e+05	6.79e+06	1.60e+08	1.44e+10	2.68e+13	4.84e+16	3.82e+18
	V _{FP}	9.40e+05	6.33e+06	1.58e+08	1.45e+10	2.71e+13	4.93e+16	3.89e+18
	V _{Be}	9.49e+05	6.65e+06	1.62e+08	1.42e+10	2.70e+13	4.93e+16	3.89e+18
	V _{Ti}	9.50e+05	7.10e+06	1.62e+08	1.44e+10	2.70e+13	4.93e+16	3.89e+18
	V _{MT1}	9.40e+05	6.30e+06	1.57e+08	1.40e+10	2.68e+13	4.82e+16	3.81e+18
	V _{MT2}	9.41e+05	6.79e+06	1.60e+08	1.44e+10	2.68e+13	4.84e+16	3.82e+18
	V _{MT3}	9.40e+05	6.32e+06	1.57e+08	1.42e+10	2.68e+13	4.83e+16	3.81e+18
	V _{MT4}	9.37e+05	6.29e+06	1.57e+08	1.39e+10	2.66e+13	4.75e+16	3.75e+18
	V _{MT5}	9.38e+05	6.29e+06	1.56e+08	1.38e+10	2.67e+13	4.76e+16	3.76e+18
	V _{B1}	9.40e+05	6.77e+06	1.60e+08	1.42e+10	2.68e+13	4.83e+16	3.81e+18
V _{B2}	9.39e+05	6.74e+06	1.59e+08	1.40e+10	2.67e+13	4.77e+16	3.77e+18	
V _{B3}	9.41e+05	6.80e+06	1.61e+08	1.45e+10	2.70e+13	4.90e+16	3.86e+18	
V _{B4}	9.22e+05	6.54e+06	1.57e+08	1.39e+10	2.61e+13	4.66e+16	3.67e+18	
Rand. Sys.	V _{HT}	4.88e+07	1.32e+08	6.63e+08	1.64e+10	2.68e+13	5.40e+16	4.16e+18
	V _{SYG}	9.49e+05	6.98e+06	1.66e+08	1.50e+10	2.70e+13	5.44e+16	4.19e+18
	V _{MC1}	4.99e+07	1.32e+08	6.66e+08	1.88e+10	2.71e+13	5.35e+16	4.01e+18
	V _{MC2}	9.50e+05	7.00e+06	1.64e+08	1.50e+10	2.77e+13	5.56e+16	4.46e+18
	V _{D1}	9.38e+05	6.41e+06	1.58e+08	1.50e+10	2.63e+13	5.30e+16	4.09e+18
	V _{D2}	9.38e+05	6.42e+06	1.59e+08	1.50e+10	2.64e+13	5.31e+16	4.09e+18
	V _{D3}	9.39e+05	6.64e+06	1.63e+08	1.50e+10	2.66e+13	5.31e+16	4.09e+18
	V _{FP}	9.39e+05	6.43e+06	1.59e+08	1.50e+10	2.69e+13	5.40e+16	4.16e+18
	V _{Be}	9.48e+05	6.78e+06	1.66e+08	1.50e+10	2.70e+13	5.43e+16	4.19e+18
	V _{Ti}	9.49e+05	6.98e+06	1.66e+08	1.50e+10	2.70e+13	5.44e+16	4.19e+18
	V _{MT1}	9.38e+05	6.41e+06	1.58e+08	1.51e+10	2.61e+13	5.29e+16	4.08e+18
	V _{MT2}	9.39e+05	6.64e+06	1.63e+08	1.51e+10	2.66e+13	5.31e+16	4.09e+18
	V _{MT3}	9.38e+05	6.42e+06	1.59e+08	1.50e+10	2.64e+13	5.31e+16	4.09e+18
	V _{MT4}	9.36e+05	6.41e+06	1.58e+08	1.50e+10	2.61e+13	5.22e+16	4.01e+18
	V _{MT5}	9.37e+05	6.40e+06	1.58e+08	1.50e+10	2.59e+13	5.23e+16	4.03e+18
	V _{B1}	9.38e+05	6.63e+06	1.63e+08	1.50e+10	2.64e+13	5.31e+16	4.09e+18
V _{B2}	9.37e+05	6.60e+06	1.62e+08	1.50e+10	2.61e+13	5.23e+16	4.03e+18	
V _{B3}	9.39e+05	6.65e+06	1.64e+08	1.50e+10	2.68e+13	5.39e+16	4.14e+18	
V _{B4}	9.21e+05	6.52e+06	1.60e+08	1.46e+10	2.60e+13	5.12e+16	3.94e+18	

Table B.12: Summary statistics for the distribution of Monte Carlo MSE of different variance estimators for the Sukhatme population, conditioned on sampling design.

Sampling	Estimator	min	5%	25%	50%	75%	95%	max
Brewer	VHT	6.93e+10	1.53e+11	7.30e+11	1.56e+13	1.74e+15	4.49e+18	1.96e+20
	VSYG	4.53e+09	1.92e+10	4.78e+11	1.56e+13	1.78e+15	4.36e+18	2.08e+20
	VMC1	5.87e+10	1.34e+11	6.54e+11	1.58e+13	1.76e+15	4.96e+18	1.92e+20
	VMC2	4.73e+09	1.92e+10	4.67e+11	1.54e+13	1.78e+15	5.29e+18	1.61e+20
	VD1	5.05e+09	1.74e+10	4.30e+11	1.41e+13	1.53e+15	3.71e+18	1.68e+20
	VD2	5.16e+09	1.77e+10	4.37e+11	1.42e+13	1.57e+15	3.82e+18	1.71e+20
	VD3	5.38e+09	1.82e+10	4.60e+11	1.43e+13	1.62e+15	4.01e+18	1.73e+20
	VFP	4.78e+09	1.87e+10	4.64e+11	1.42e+13	1.67e+15	4.49e+18	1.67e+20
	VBe	4.90e+09	1.89e+10	4.64e+11	1.51e+13	1.74e+15	4.26e+18	1.93e+20
	VT1	4.76e+09	1.93e+10	4.82e+11	1.52e+13	1.76e+15	4.36e+18	1.95e+20
	VMT1	5.19e+09	1.71e+10	4.21e+11	1.42e+13	1.49e+15	3.63e+18	1.70e+20
	VMT2	5.35e+09	1.82e+10	4.59e+11	1.43e+13	1.61e+15	4.01e+18	1.73e+20
	VMT3	5.14e+09	1.77e+10	4.36e+11	1.42e+13	1.56e+15	3.82e+18	1.71e+20
	VMT4	4.90e+09	1.69e+10	4.17e+11	1.32e+13	1.45e+15	3.40e+18	1.48e+20
	VMT5	5.04e+09	1.67e+10	4.07e+11	1.37e+13	1.43e+15	3.36e+18	1.58e+20
	VB1	5.25e+09	1.79e+10	4.53e+11	1.42e+13	1.58e+15	3.89e+18	1.71e+20
VB2	5.85e+09	1.72e+10	4.27e+11	1.38e+13	1.47e+15	3.51e+18	1.60e+20	
VB3	4.76e+09	1.86e+10	4.75e+11	1.47e+13	1.70e+15	4.29e+18	1.82e+20	
VB4	3.84e+09	1.58e+10	3.83e+11	1.04e+13	1.30e+15	3.05e+18	1.07e+20	
Tillé	VHT	5.27e+10	1.17e+11	7.60e+11	1.62e+13	1.68e+15	4.28e+18	1.82e+20
	VSYG	4.99e+09	1.73e+10	4.31e+11	1.66e+13	1.61e+15	3.91e+18	2.18e+20
	VMC1	5.26e+10	1.17e+11	7.64e+11	1.62e+13	1.68e+15	4.44e+18	1.85e+20
	VMC2	5.00e+09	1.74e+10	4.19e+11	1.67e+13	1.60e+15	4.04e+18	1.84e+20
	VD1	5.07e+09	1.78e+10	4.04e+11	1.37e+13	1.45e+15	3.55e+18	1.61e+20
	VD2	5.18e+09	1.81e+10	4.16e+11	1.38e+13	1.48e+15	3.66e+18	1.62e+20
	VD3	5.38e+09	1.86e+10	4.40e+11	1.39e+13	1.53e+15	3.83e+18	1.64e+20
	VFP	4.82e+09	1.91e+10	4.68e+11	1.40e+13	1.60e+15	4.28e+18	1.59e+20
	VBe	4.94e+09	1.91e+10	4.50e+11	1.46e+13	1.62e+15	4.06e+18	1.89e+20
	VT1	4.81e+09	1.95e+10	4.64e+11	1.47e+13	1.64e+15	4.17e+18	1.91e+20
	VMT1	5.20e+09	1.75e+10	3.94e+11	1.38e+13	1.41e+15	3.47e+18	1.60e+20
	VMT2	5.35e+09	1.86e+10	4.38e+11	1.39e+13	1.52e+15	3.83e+18	1.65e+20
	VMT3	5.15e+09	1.81e+10	4.15e+11	1.38e+13	1.48e+15	3.65e+18	1.63e+20
	VMT4	4.92e+09	1.72e+10	3.82e+11	1.29e+13	1.37e+15	3.25e+18	1.41e+20
	VMT5	5.05e+09	1.70e+10	3.74e+11	1.33e+13	1.35e+15	3.21e+18	1.48e+20
	VB1	5.26e+09	1.83e+10	4.27e+11	1.38e+13	1.50e+15	3.71e+18	1.62e+20
VB2	5.83e+09	1.76e+10	3.94e+11	1.34e+13	1.39e+15	3.35e+18	1.50e+20	
VB3	4.80e+09	1.90e+10	4.62e+11	1.43e+13	1.61e+15	4.10e+18	1.75e+20	
VB4	3.99e+09	1.58e+10	3.55e+11	1.01e+13	1.24e+15	2.92e+18	1.06e+20	
CPS	VHT	5.81e+10	1.35e+11	6.61e+11	1.48e+13	1.85e+15	5.40e+18	1.57e+20
	VSYG	4.79e+09	1.92e+10	4.62e+11	1.50e+13	1.87e+15	5.44e+18	1.51e+20
	VMC1	5.82e+10	1.35e+11	6.66e+11	1.50e+13	1.85e+15	5.57e+18	1.56e+20
	VMC2	4.74e+09	1.92e+10	4.71e+11	1.50e+13	1.85e+15	5.77e+18	1.67e+20
	VD1	5.10e+09	1.74e+10	4.23e+11	1.40e+13	1.63e+15	4.45e+18	1.32e+20
	VD2	5.22e+09	1.78e+10	4.30e+11	1.41e+13	1.66e+15	4.66e+18	1.34e+20
	VD3	5.43e+09	1.82e+10	4.53e+11	1.42e+13	1.70e+15	4.87e+18	1.36e+20
	VFP	4.83e+09	1.87e+10	4.56e+11	1.39e+13	1.70e+15	5.40e+18	1.31e+20
	VBe	4.96e+09	1.89e+10	4.57e+11	1.50e+13	1.84e+15	5.32e+18	1.49e+20
	VT1	4.81e+09	1.93e+10	4.75e+11	1.50e+13	1.86e+15	5.41e+18	1.51e+20
	VMT1	5.24e+09	1.72e+10	4.14e+11	1.41e+13	1.58e+15	4.29e+18	1.35e+20
	VMT2	5.40e+09	1.82e+10	4.52e+11	1.42e+13	1.70e+15	4.86e+18	1.36e+20
	VMT3	5.19e+09	1.77e+10	4.29e+11	1.41e+13	1.66e+15	4.65e+18	1.34e+20
	VMT4	4.94e+09	1.69e+10	4.10e+11	1.31e+13	1.51e+15	4.10e+18	1.16e+20
	VMT5	5.08e+09	1.67e+10	4.03e+11	1.36e+13	1.51e+15	3.98e+18	1.25e+20
	VB1	5.30e+09	1.79e+10	4.46e+11	1.41e+13	1.68e+15	4.72e+18	1.34e+20
VB2	5.91e+09	1.73e+10	4.26e+11	1.36e+13	1.56e+15	4.26e+18	1.27e+20	
VB3	4.81e+09	1.86e+10	4.67e+11	1.45e+13	1.78e+15	5.21e+18	1.42e+20	
VB4	3.86e+09	1.58e+10	3.77e+11	1.02e+13	1.28e+15	3.71e+18	8.24e+19	
Rand. Sys.	VHT	5.96e+10	1.31e+11	6.53e+11	1.58e+13	1.86e+15	4.31e+18	2.71e+20
	VSYG	4.94e+09	1.95e+10	4.79e+11	1.57e+13	1.85e+15	4.45e+18	2.81e+20
	VMC1	6.11e+10	1.38e+11	6.65e+11	1.60e+13	1.86e+15	4.21e+18	2.63e+20
	VMC2	4.76e+09	1.95e+10	4.82e+11	1.55e+13	1.85e+15	4.57e+18	2.32e+20
	VD1	5.10e+09	1.75e+10	4.08e+11	1.46e+13	1.62e+15	3.77e+18	2.39e+20
	VD2	5.22e+09	1.78e+10	4.19e+11	1.48e+13	1.66e+15	3.83e+18	2.41e+20
	VD3	5.45e+09	1.83e+10	4.41e+11	1.49e+13	1.71e+15	3.89e+18	2.44e+20
	VFP	4.81e+09	1.88e+10	4.68e+11	1.46e+13	1.79e+15	4.06e+18	2.37e+20
	VBe	4.94e+09	1.91e+10	4.70e+11	1.56e+13	1.84e+15	4.41e+18	2.80e+20
	VT1	4.78e+09	1.95e+10	4.79e+11	1.56e+13	1.85e+15	4.47e+18	2.83e+20
	VMT1	5.25e+09	1.72e+10	3.97e+11	1.48e+13	1.58e+15	3.77e+18	2.39e+20
	VMT2	5.42e+09	1.82e+10	4.40e+11	1.49e+13	1.71e+15	3.90e+18	2.45e+20
	VMT3	5.20e+09	1.78e+10	4.18e+11	1.47e+13	1.66e+15	3.84e+18	2.43e+20
	VMT4	4.93e+09	1.70e+10	3.86e+11	1.37e+13	1.54e+15	3.32e+18	2.09e+20
	VMT5	5.08e+09	1.67e+10	3.77e+11	1.42e+13	1.51e+15	3.49e+18	2.21e+20
	VB1	5.32e+09	1.79e+10	4.28e+11	1.48e+13	1.68e+15	3.83e+18	2.41e+20
VB2	5.95e+09	1.73e+10	3.95e+11	1.43e+13	1.56e+15	3.55e+18	2.24e+20	
VB3	4.79e+09	1.87e+10	4.64e+11	1.52e+13	1.80e+15	4.13e+18	2.60e+20	
VB4	3.68e+09	1.58e+10	3.58e+11	1.02e+13	1.33e+15	2.95e+18	1.57e+20	

Table B.13: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f. MU273 population (N = 273), CV(Y) = 1.29, CV(X) = 1.27

Table with columns for estimator, sampling, and MSE for f=1%, 5%, 10%, 15%, and 20% across Brewer, Tillé, CPS, and Rand. Sys. designs.

Table B.14: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f. MU273 population (N = 273), CV(Y) = 1.29, CV(X) = 1

Table with columns: sampling, estimator, f=1%, f=5%, f=10%, f=15%, f=20% and sub-columns for RB, RS, MSE. Rows are grouped by estimator (Brewer, Tillé, CPS, Rand. Sys.) and sampling fraction (f). Each group contains multiple estimator types (VHT, VSYG, VMC1, VMC2, VD1, VD2, VD3, VFP, VBe, VTi, VMT1-5, VB1-4, VB4).

Table B.16: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f. MU273 population (N = 273), CV(Y) = 1.29, CV(X) = 0.67

Table with columns for sampling, estimator, and performance metrics (RB, RS, MSE) across five sampling fractions (f=1%, f=5%, f=10%, f=15%, f=20%) for three designs: Brewer, Tillé, and Rand. Sys.

Table B.24: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f. MU273 population (N = 273), CV(Y) = 0.82, CV(X) = 1

Table with columns for sampling, estimator, and sampling fractions (f=1%, f=5%, f=10%, f=15%, f=20%). Rows are grouped by design (Brewer, Tillé, CPS, Rand. Sys.) and estimator (VHT, VSYG, VMC1, VMC2, VD1, VD2, VD3, VFP, VBe, VT1, VMT1, VMT2, VMT3, VMT4, VMT5, VB1, VB2, VB3, VB4).

Table B.30: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f. MU273 population (N = 273), CV(Y) = 0.67, CV(X) = 0.81

Table with columns for sampling, estimator, and MSE/RB/RS for five sampling fractions (f=1%, f=5%, f=10%, f=15%, f=20%). Rows are grouped by design (Brewer, Tillé, CPS, Rand. Sys.) and estimator (VHT, VSYG, VMC1, VMC2, VD1, VD2, VD3, VFP, VBe, VT1, VMT1, VMT2, VMT3, VMT4, VMT5, VB1, VB2, VB3, VB4). Numerical values are provided for each cell.

Table B.37: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f. MU273 population (N = 273), CV(Y) = 0.5, CV(X) = 0.5

Table with columns: sampling, estimator, and sub-columns for f=1%, f=5%, f=10%, f=15%, and f=20% (each with RB, RS, MSE). Rows are grouped by design: Brewer, Tillé, CPS, and Rand. Sys. Each design includes estimators VHT, VSYG, VM C1, VM C2, VD1, VD2, VD3, VFP, VBe, VT1, VM T1, VM T2, VM T3, VM T4, VM T5, VB1, VB2, VB3, and VB4.

Table B.45: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f. Sukhatme population (N = 34), CV(Y) = 0.95, CV(X) = 0.77

Table with columns for sampling, estimator, and MSE/RS/RB for f=5%, f=10%, f=15%, and f=20%. Rows are grouped by design (Brewer, Tillé, CPS, Rand. Sys.) and estimator type (VHT, VSYG, VMC1, VMC2, VD1, VD2, VD3, VFP, VBe, VTr, VMT1-5, VB1-5).

Table B.49: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f. Sukhatme population (N = 34), CV(Y) = 0.76, CV(X) = 0.96

Table with 14 columns: sampling, estimator, and three groups of metrics (RB, RS, MSE) for sampling fractions f=5%, f=10%, f=15%, and f=20%. The rows are categorized by design type: Brewer, Tillé, CPS, and Rand. Sys.

Table B.51: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f . Sukhatme population ($N = 34$), $CV(Y) = 0.76$, $CV(X) = 0.67$

sampling	estimator	$f=5\%$			$f=10\%$			$f=15\%$			$f=20\%$		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	VHT	—	2.29	4.7e+12	—	4.16	6e+11	—	6.68	2.2e+11	—	7.84	1.5e+11
	VSYG	—	1	2.1e+12	—	1	1.4e+11	—	1	3.3e+10	—	1	2e+10
	VMC1	1.09	1.55	3.2e+12	-0.75	3.24	4.7e+11	-2.06	5.7	1.9e+11	-0.61	6.9	1.4e+11
	VMC2	1.27	1.01	2.1e+12	-0.78	1.01	1.5e+11	-1.16	1	3.3e+10	0.15	1	2e+10
	VD1	2.22	1.03	2.1e+12	-0.2	0.99	1.4e+11	-0.8	0.93	3.1e+10	0.24	0.9	1.8e+10
	VD2	2.4	1.04	2.1e+12	0.16	0.99	1.4e+11	-0.18	0.95	3.1e+10	1.03	0.92	1.8e+10
	VD3	2.57	1.04	2.1e+12	0.62	1.01	1.5e+11	0.61	0.96	3.2e+10	2.03	0.94	1.8e+10
	VFP	2.09	1.03	2.1e+12	-0.97	1.01	1.5e+11	-1.12	0.99	3.3e+10	0.22	0.97	1.9e+10
	VBe	1.42	1.03	2.1e+12	-0.6	1.01	1.5e+11	-0.68	0.99	3.3e+10	0.64	0.98	1.9e+10
	VT1	1.21	1.02	2.1e+12	-1.06	1	1.4e+11	-1.44	1	3.3e+10	-0.3	1	2e+10
	VMT1	3.96	1.07	2.2e+12	1.25	1	1.4e+11	0.1	0.92	3.1e+10	0.84	0.89	1.7e+10
	VMT2	2.63	1.04	2.1e+12	0.63	1	1.4e+11	0.61	0.96	3.2e+10	2.04	0.94	1.8e+10
	VMT3	2.45	1.04	2.1e+12	0.17	0.99	1.4e+11	-0.18	0.94	3.1e+10	1.04	0.92	1.8e+10
	VMT4	-0.31	0.98	2e+12	-1.86	0.95	1.4e+11	-2.29	0.91	3e+10	-1.22	0.88	1.7e+10
	VMT5	2.49	1.04	2.1e+12	-0.18	0.97	1.4e+11	-1.3	0.9	3e+10	-0.56	0.86	1.7e+10
	VB1	2.4	1.04	2.1e+12	0.25	1	1.4e+11	-0.01	0.95	3.2e+10	1.23	0.93	1.8e+10
VB2	2.88	1.05	2.2e+12	1.14	0.99	1.4e+11	1.12	0.92	3.1e+10	2.5	0.89	1.7e+10	
VB3	1.91	1.03	2.1e+12	-0.63	1.01	1.5e+11	-1.14	0.98	3.3e+10	-0.04	0.97	1.9e+10	
VB4	-17.59	0.7	1.4e+12	-14.94	0.79	1.1e+11	-15.8	0.82	2.7e+10	-15.59	0.82	1.6e+10	
Tillé	VHT	—	1.73	3.6e+12	—	3.51	4.9e+11	—	5.7	1.8e+11	—	7.03	1.2e+11
	VSYG	—	1	2.1e+12	—	1	1.4e+11	—	1	3.1e+10	—	1	1.7e+10
	VMC1	1.16	1.79	3.7e+12	1.42	3.53	4.9e+11	-1.01	5.72	1.8e+11	-0.3	7.05	1.2e+11
	VMC2	1.59	1	2.1e+12	1.01	1	1.4e+11	0.2	1	3.1e+10	0.13	1.01	1.8e+10
	VD1	2.58	1.03	2.1e+12	2.58	1.03	1.4e+11	1.93	1.02	3.2e+10	2.02	1.03	1.8e+10
	VD2	2.76	1.03	2.1e+12	2.95	1.04	1.5e+11	2.56	1.03	3.2e+10	2.84	1.05	1.8e+10
	VD3	2.94	1.04	2.1e+12	3.44	1.06	1.5e+11	3.39	1.05	3.3e+10	3.88	1.08	1.9e+10
	VFP	2.45	1.02	2.1e+12	1.85	1.07	1.5e+11	1.65	1.08	3.4e+10	2.09	1.11	2e+10
	VBe	1.8	1.02	2.1e+12	2.18	1.05	1.5e+11	2.09	1.08	3.4e+10	2.4	1.11	1.9e+10
	VT1	1.59	1.01	2.1e+12	1.74	1.05	1.5e+11	1.41	1.09	3.4e+10	1.59	1.14	2e+10
	VMT1	4.32	1.06	2.2e+12	4.05	1.05	1.5e+11	2.82	1.01	3.1e+10	2.61	1.01	1.8e+10
	VMT2	3	1.04	2.1e+12	3.44	1.05	1.5e+11	3.39	1.05	3.3e+10	3.87	1.08	1.9e+10
	VMT3	2.82	1.03	2.1e+12	2.96	1.04	1.5e+11	2.56	1.03	3.2e+10	2.84	1.05	1.8e+10
	VMT4	0.04	0.98	2e+12	0.87	1	1.4e+11	0.4	0.99	3.1e+10	0.53	1	1.8e+10
	VMT5	2.85	1.03	2.1e+12	2.59	1.01	1.4e+11	1.39	0.98	3.1e+10	1.19	0.98	1.7e+10
	VB1	2.76	1.03	2.1e+12	3.06	1.05	1.5e+11	2.75	1.04	3.2e+10	3.05	1.06	1.9e+10
VB2	3.24	1.04	2.1e+12	3.94	1.04	1.5e+11	3.86	1.01	3.2e+10	4.28	1.02	1.8e+10	
VB3	2.28	1.02	2.1e+12	2.18	1.06	1.5e+11	1.64	1.07	3.4e+10	1.83	1.11	1.9e+10	
VB4	-17.27	0.69	1.4e+12	-12.5	0.83	1.2e+11	-13.37	0.88	2.8e+10	-13.92	0.92	1.6e+10	
CPS	VHT	—	1.48	3.1e+12	—	3.11	4.7e+11	—	5.78	1.9e+11	—	6.97	1.4e+11
	VSYG	—	1	2.1e+12	—	1	1.5e+11	—	1	3.3e+10	—	1	2e+10
	VMC1	-1.21	1.52	3.1e+12	-0.83	3.13	4.7e+11	1.26	5.81	1.9e+11	-0.5	6.98	1.4e+11
	VMC2	-0.95	1	2.1e+12	-1.08	1	1.5e+11	0.43	1	3.3e+10	-0.44	1	2e+10
	VD1	0.14	1.01	2.1e+12	-0.42	0.98	1.5e+11	0.87	0.94	3.1e+10	-0.4	0.91	1.8e+10
	VD2	0.32	1.02	2.1e+12	-0.06	0.99	1.5e+11	1.49	0.95	3.2e+10	0.39	0.92	1.8e+10
	VD3	0.49	1.02	2.1e+12	0.41	1	1.5e+11	2.3	0.97	3.2e+10	1.38	0.95	1.9e+10
	VFP	0.01	1.01	2.1e+12	-1.15	1.01	1.5e+11	0.55	1	3.3e+10	-0.4	0.98	1.9e+10
	VBe	-0.62	1	2.1e+12	-0.77	1	1.5e+11	0.89	0.99	3.3e+10	-0.05	0.98	1.9e+10
	VT1	-0.82	1	2.1e+12	-1.22	1	1.5e+11	0.11	1	3.3e+10	-0.99	1.01	2e+10
	VMT1	1.84	1.05	2.2e+12	1.01	0.99	1.5e+11	1.78	0.93	3.1e+10	0.2	0.89	1.7e+10
	VMT2	0.55	1.02	2.1e+12	0.42	1	1.5e+11	2.3	0.97	3.2e+10	1.39	0.94	1.8e+10
	VMT3	0.37	1.02	2.1e+12	-0.04	0.98	1.5e+11	1.5	0.95	3.2e+10	0.4	0.92	1.8e+10
	VMT4	-2.34	0.96	2e+12	-2.07	0.95	1.4e+11	-0.65	0.91	3e+10	-1.85	0.88	1.7e+10
	VMT5	0.4	1.02	2.1e+12	-0.41	0.96	1.4e+11	0.36	0.9	3e+10	-1.19	0.87	1.7e+10
	VB1	0.32	1.02	2.1e+12	0.05	0.99	1.5e+11	1.67	0.96	3.2e+10	0.59	0.93	1.8e+10
VB2	0.78	1.03	2.1e+12	0.91	0.98	1.5e+11	2.83	0.93	3.1e+10	1.84	0.89	1.7e+10	
VB3	-0.15	1.01	2.1e+12	-0.81	1	1.5e+11	0.51	0.99	3.3e+10	-0.67	0.97	1.9e+10	
VB4	-19.24	0.69	1.4e+12	-15.07	0.79	1.2e+11	-14.4	0.81	2.7e+10	-16.12	0.83	1.6e+10	
Rand. Sys.	VHT	—	1.46	3e+12	—	2.99	4.5e+11	—	5.36	1.8e+11	—	6.64	1.3e+11
	VSYG	—	1	2.1e+12	—	1	1.5e+11	—	1	3.4e+10	—	1	2e+10
	VMC1	1.62	1.54	3.2e+12	0.06	3.15	4.8e+11	0.94	5.69	1.9e+11	0.59	6.96	1.4e+11
	VMC2	1.18	1	2.1e+12	0.94	0.99	1.5e+11	-0.23	1	3.4e+10	0.27	1	2e+10
	VD1	2.13	1.01	2.1e+12	1.67	0.98	1.5e+11	0.46	0.92	3.1e+10	0.98	0.89	1.8e+10
	VD2	2.31	1.02	2.1e+12	2.04	0.98	1.5e+11	1.08	0.93	3.2e+10	1.77	0.91	1.8e+10
	VD3	2.49	1.02	2.1e+12	2.51	1	1.5e+11	1.89	0.95	3.2e+10	2.76	0.93	1.9e+10
	VFP	2	1.01	2.1e+12	0.96	1.01	1.5e+11	0.11	0.98	3.3e+10	0.96	0.97	1.9e+10
	VBe	1.29	1	2.1e+12	1.33	1	1.5e+11	0.56	0.99	3.4e+10	1.38	0.98	1.9e+10
	VT1	1.07	1	2.1e+12	0.88	1	1.5e+11	-0.24	1	3.4e+10	0.36	1	2e+10
	VMT1	3.9	1.05	2.2e+12	3.11	0.98	1.5e+11	1.37	0.91	3.1e+10	1.58	0.88	1.7e+10
	VMT2	2.55	1.02	2.1e+12	2.52	0.99	1.5e+11	1.9	0.95	3.2e+10	2.77	0.93	1.9e+10
	VMT3	2.37	1.02	2.1e+12	2.05	0.98	1.5e+11	1.09	0.93	3.2e+10	1.78	0.91	1.8e+10
	VMT4	-0.39	0.96	2e+12	-0.02	0.94	1.4e+11	-1.05	0.89	3e+10	-0.5	0.87	1.7e+10
	VMT5	2.43	1.02	2.1e+12	1.66	0.96	1.4e+11	-0.04	0.88	3e+10	0.16	0.85	1.7e+10
	VB1	2.31	1.02	2.1e+12	2.14	0.99	1.5e+11	1.26	0.94	3.2e+10	1.96	0.92	1.8e+10
VB2	2.82	1.03	2.1e+12	3	0.98	1.5e+11	2.43	0.91	3.1e+10	3.24	0.88	1.7e+10	
VB3	1.8	1.01	2.1e+12	1.29	1	1.5e+11	0.1	0.97	3.3e+10	0.67	0.96	1.9e+10	
VB4	-17.72	0.68	1.4e+12	-13.24	0.78	1.2e+11	-14.77	0.8	2.7e+10	-15	0.81	1.6e+10	

Table B.52: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f. Sukhatme population (N = 34), CV(Y) = 0.76, CV(X) = 0.5

sampling	estimator	f=5%			f=10%			f=15%			f=20%		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	VHT	—	1.05	8.3e+12	—	1.29	7.1e+11	—	1.65	2.1e+11	—	1.88	1.3e+11
	VSYG	—	1	7.9e+12	—	1	5.5e+11	—	1	1.3e+11	—	1	7.1e+10
	VMC1	0.94	0.94	7.4e+12	-0.69	1.11	6.1e+11	-0.84	1.42	1.8e+11	0.39	1.6	1.1e+11
	VMC2	0.86	1.01	7.9e+12	-0.64	0.99	5.5e+11	-0.5	0.99	1.2e+11	0.86	0.99	7e+10
	VD1	0.71	0.99	7.8e+12	-0.88	0.97	5.3e+11	-0.88	0.95	1.2e+11	0.42	0.93	6.6e+10
	VD2	0.81	1	7.8e+12	-0.7	0.98	5.4e+11	-0.57	0.95	1.2e+11	0.82	0.94	6.7e+10
	VD3	0.91	1	7.9e+12	-0.37	0.99	5.4e+11	0.15	0.98	1.2e+11	1.86	0.97	6.9e+10
	VFP	0.63	0.99	7.8e+12	-0.76	0.99	5.5e+11	-0.62	0.98	1.2e+11	0.75	0.96	6.9e+10
	VBe	0.74	1	7.9e+12	-0.72	0.99	5.4e+11	-0.6	0.98	1.2e+11	0.78	0.97	6.9e+10
	VT1	0.72	1	7.9e+12	-0.65	1	5.5e+11	-0.27	1	1.3e+11	1.33	1.01	7.2e+10
	VMT1	1.46	1	7.9e+12	-0.3	0.97	5.3e+11	-0.49	0.94	1.2e+11	0.72	0.92	6.6e+10
	VMT2	0.94	1	7.9e+12	-0.36	0.98	5.4e+11	0.16	0.98	1.2e+11	1.87	0.97	6.9e+10
	VMT3	0.84	0.99	7.9e+12	-0.68	0.97	5.3e+11	-0.56	0.95	1.2e+11	0.83	0.94	6.7e+10
	VMT4	-0.71	0.96	7.6e+12	-1.81	0.95	5.2e+11	-1.72	0.93	1.2e+11	-0.41	0.91	6.5e+10
	VMT5	0.65	0.99	7.8e+12	-1.09	0.96	5.2e+11	-1.27	0.93	1.2e+11	-0.07	0.91	6.5e+10
VB1	0.81	1	7.8e+12	-0.56	0.98	5.4e+11	-0.16	0.97	1.2e+11	1.46	0.96	6.9e+10	
VB2	0.83	0.99	7.8e+12	-0.53	0.97	5.3e+11	-0.13	0.95	1.2e+11	1.49	0.94	6.7e+10	
VB3	0.8	1	7.9e+12	-0.59	0.99	5.5e+11	-0.18	0.99	1.2e+11	1.43	0.99	7e+10	
VB4	-15.13	0.73	5.7e+12	-11.97	0.82	4.5e+11	-11.42	0.84	1.1e+11	-10.23	0.83	5.9e+10	
Tillé	VHT	—	1.09	8.3e+12	—	1.24	6.1e+11	—	1.5	1.5e+11	—	1.65	9.3e+10
	VSYG	—	1	7.6e+12	—	1	4.9e+11	—	1	1e+11	—	1	5.6e+10
	VMC1	1.68	1.1	8.4e+12	-0.67	1.25	6.2e+11	-0.98	1.51	1.6e+11	-0.08	1.65	9.3e+10
	VMC2	1.68	1	7.7e+12	-0.69	1	5e+11	-0.69	1	1e+11	-0.07	1	5.6e+10
	VD1	2.83	1.05	8.1e+12	1.99	1.1	5.4e+11	3.69	1.15	1.2e+11	5.34	1.18	6.6e+10
	VD2	2.93	1.06	8.1e+12	2.18	1.11	5.5e+11	4.01	1.16	1.2e+11	5.76	1.19	6.7e+10
	VD3	3.04	1.06	8.1e+12	2.51	1.12	5.5e+11	4.77	1.19	1.2e+11	6.86	1.24	7e+10
	VFP	2.75	1.05	8e+12	2.17	1.13	5.6e+11	3.97	1.19	1.2e+11	5.71	1.22	6.9e+10
	VBe	2.85	1.06	8.1e+12	2.14	1.12	5.5e+11	3.98	1.19	1.2e+11	5.67	1.23	6.9e+10
	VT1	2.83	1.06	8.1e+12	2.22	1.13	5.6e+11	4.34	1.22	1.3e+11	6.28	1.28	7.2e+10
	VMT1	3.59	1.07	8.1e+12	2.57	1.1	5.4e+11	4.09	1.14	1.2e+11	5.65	1.17	6.6e+10
	VMT2	3.06	1.06	8.1e+12	2.52	1.12	5.5e+11	4.78	1.19	1.2e+11	6.87	1.24	7e+10
	VMT3	2.96	1.06	8.1e+12	2.19	1.1	5.5e+11	4.02	1.16	1.2e+11	5.77	1.19	6.7e+10
	VMT4	1.38	1.02	7.8e+12	1.03	1.08	5.3e+11	2.81	1.13	1.2e+11	4.47	1.16	6.5e+10
	VMT5	2.77	1.05	8e+12	1.76	1.08	5.3e+11	3.27	1.12	1.2e+11	4.82	1.15	6.5e+10
VB1	2.93	1.06	8.1e+12	2.32	1.11	5.5e+11	4.45	1.18	1.2e+11	6.44	1.22	6.9e+10	
VB2	2.95	1.05	8.1e+12	2.32	1.1	5.4e+11	4.46	1.15	1.2e+11	6.46	1.19	6.7e+10	
VB3	2.91	1.06	8.1e+12	2.31	1.13	5.6e+11	4.43	1.21	1.2e+11	6.42	1.26	7.1e+10	
VB4	-13.35	0.77	5.9e+12	-9.38	0.92	4.5e+11	-7.31	0.99	1e+11	-5.8	1.02	5.7e+10	
CPS	VHT	—	0.93	7.6e+12	—	1.12	6.2e+11	—	1.43	1.8e+11	—	1.65	1.1e+11
	VSYG	—	1	8.2e+12	—	1	5.5e+11	—	1	1.3e+11	—	1	6.9e+10
	VMC1	0.49	0.94	7.7e+12	-1.19	1.12	6.2e+11	0.63	1.43	1.8e+11	-1.79	1.65	1.1e+11
	VMC2	0.52	1	8.2e+12	-1.24	1	5.5e+11	-0.21	1	1.2e+11	-2.13	1	6.9e+10
	VD1	0.54	0.99	8.1e+12	-1.36	0.98	5.4e+11	-0.29	0.96	1.2e+11	-2.49	0.95	6.5e+10
	VD2	0.65	0.99	8.1e+12	-1.18	0.98	5.4e+11	0.02	0.97	1.2e+11	-2.11	0.96	6.6e+10
	VD3	0.75	1	8.1e+12	-0.86	0.99	5.5e+11	0.75	0.99	1.2e+11	-1.1	0.99	6.8e+10
	VFP	0.46	0.99	8.1e+12	-1.21	1	5.5e+11	-0.01	0.99	1.2e+11	-2.17	0.98	6.8e+10
	VBe	0.6	1	8.2e+12	-1.21	1	5.5e+11	-0.1	0.99	1.2e+11	-2.2	0.99	6.8e+10
	VT1	0.58	1	8.2e+12	-1.13	1	5.6e+11	0.22	1.02	1.3e+11	-1.67	1.02	7.1e+10
	VMT1	1.28	1	8.2e+12	-0.79	0.98	5.4e+11	0.1	0.95	1.2e+11	-2.2	0.94	6.5e+10
	VMT2	0.77	1	8.1e+12	-0.85	0.99	5.5e+11	0.76	0.99	1.2e+11	-1.09	0.99	6.8e+10
	VMT3	0.67	0.99	8.1e+12	-1.17	0.98	5.4e+11	0.03	0.97	1.2e+11	-2.1	0.96	6.6e+10
	VMT4	-0.87	0.96	7.9e+12	-2.29	0.96	5.3e+11	-1.14	0.94	1.2e+11	-3.3	0.93	6.5e+10
	VMT5	0.47	0.99	8e+12	-1.58	0.96	5.3e+11	-0.68	0.94	1.2e+11	-2.97	0.93	6.4e+10
VB1	0.65	0.99	8.1e+12	-1.04	0.99	5.5e+11	0.43	0.98	1.2e+11	-1.49	0.98	6.8e+10	
VB2	0.65	0.99	8.1e+12	-1.02	0.97	5.4e+11	0.47	0.96	1.2e+11	-1.46	0.96	6.6e+10	
VB3	0.64	1	8.1e+12	-1.06	1	5.5e+11	0.4	1.01	1.3e+11	-1.53	1	6.9e+10	
VB4	-15.24	0.73	5.9e+12	-12.37	0.82	4.6e+11	-10.91	0.85	1.1e+11	-12.85	0.87	6e+10	
Rand. Sys.	VHT	—	0.93	7.3e+12	—	1.1	6.1e+11	—	1.41	1.8e+11	—	1.63	1.2e+11
	VSYG	—	1	7.9e+12	—	1	5.5e+11	—	1	1.3e+11	—	1	7.1e+10
	VMC1	-0.08	0.93	7.3e+12	-0.48	1.13	6.3e+11	0.02	1.47	1.9e+11	-0.19	1.7	1.2e+11
	VMC2	-0.19	1	7.8e+12	0	1	5.5e+11	0.18	1	1.3e+11	0.09	1	7.1e+10
	VD1	-0.19	0.99	7.8e+12	-0.43	0.97	5.4e+11	-0.53	0.94	1.2e+11	-0.83	0.92	6.6e+10
	VD2	-0.09	0.99	7.8e+12	-0.25	0.97	5.4e+11	-0.22	0.95	1.2e+11	-0.44	0.93	6.6e+10
	VD3	0.01	0.99	7.8e+12	0.07	0.98	5.5e+11	0.51	0.97	1.2e+11	0.58	0.96	6.9e+10
	VFP	-0.26	0.99	7.8e+12	-0.28	0.99	5.5e+11	-0.28	0.97	1.2e+11	-0.5	0.95	6.8e+10
	VBe	-0.14	1	7.8e+12	-0.24	0.99	5.5e+11	-0.25	0.97	1.2e+11	-0.49	0.96	6.9e+10
	VT1	-0.16	1	7.8e+12	-0.16	1	5.5e+11	0.07	1	1.3e+11	0.04	1	7.1e+10
	VMT1	0.54	1	7.9e+12	0.14	0.97	5.4e+11	-0.14	0.93	1.2e+11	-0.53	0.91	6.5e+10
	VMT2	0.04	0.99	7.8e+12	0.09	0.98	5.4e+11	0.52	0.97	1.2e+11	0.59	0.96	6.8e+10
	VMT3	-0.06	0.99	7.8e+12	-0.23	0.97	5.4e+11	-0.21	0.95	1.2e+11	-0.43	0.93	6.6e+10
	VMT4	-1.59	0.96	7.6e+12	-1.37	0.95	5.3e+11	-1.37	0.92	1.2e+11	-1.65	0.91	6.5e+10
	VMT5	-0.25	0.98	7.7e+12	-0.65	0.95	5.3e+11	-0.92	0.92	1.2e+11	-1.31	0.9	6.4e+10
VB1	-0.09	0.99	7.8e+12	-0.11	0.98	5.4e+11	0.2	0.96	1.2e+11	0.19	0.95	6.8e+10	
VB2	-0.08	0.99	7.8e+12	-0.1	0.97	5.4e+11	0.23	0.94	1.2e+11	0.22	0.93	6.6e+10	
VB3	-0.1	1	7.8e+12	-0.12	0.99	5.5e+11	0.16	0.98	1.2e+11	0.15	0.98	7e+10	
VB4	-15.87	0.73	5.7e+12	-11.53	0.81	4.5e+11	-11.12	0.83	1.1e+11	-11.36	0.83	5.9e+10	

Table B.53: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f . Sukhatme population ($N = 34$), $CV(Y) = 0.67$, $CV(X) = 1.1$

sampling	estimator	f=5%			f=10%			f=15%			f=20%		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	V _{HT}	—	0.94	4.1e+19	—	1.03	4.4e+18	—	1.04	1.1e+18	—	1.13	8e+17
	V _{SYG}	—	1	4.4e+19	—	1	4.3e+18	—	1	1.1e+18	—	1	7e+17
	V _{MC1}	-0.47	0.92	4e+19	3.66	1.15	4.9e+18	-0.58	1.01	1.1e+18	0.36	1.12	7.9e+17
	V _{MC2}	-6.77	0.78	3.4e+19	4.74	1.23	5.3e+18	-0.85	1.03	1.1e+18	-3.16	1.03	7.2e+17
	V _{D1}	-7.22	0.81	3.5e+19	-8.4	0.85	3.7e+18	-9.84	0.84	9.2e+17	-11.33	0.87	6.1e+17
	V _{D2}	-6.57	0.82	3.6e+19	-7.08	0.88	3.8e+18	-7.18	0.89	9.8e+17	-7.53	0.95	6.7e+17
	V _{D3}	-5.9	0.83	3.6e+19	-4.97	0.92	4e+18	-3.2	0.96	1.1e+18	-1.78	1.07	7.5e+17
	V _{FP}	-7.7	0.8	3.5e+19	0.23	1.03	4.4e+18	0.15	1.04	1.1e+18	0.43	1.13	8e+17
	V _{Be}	-0.96	0.93	4.1e+19	-2.32	0.98	4.2e+18	-2.53	1	1.1e+18	-6.14	0.97	6.9e+17
	V _{T1}	-0.56	0.94	4.1e+19	-1.08	1	4.3e+18	-0.38	1.03	1.1e+18	-1.97	1.06	7.4e+17
	V _{MT1}	-6.56	0.82	3.6e+19	-9.28	0.83	3.6e+18	-11.31	0.81	8.9e+17	-12.84	0.84	5.9e+17
	V _{MT2}	-5.86	0.83	3.6e+19	-5.01	0.92	4e+18	-3.39	0.96	1.1e+18	-2.13	1.06	7.5e+17
	V _{MT3}	-6.52	0.82	3.6e+19	-7.12	0.88	3.8e+18	-7.34	0.89	9.7e+17	-7.84	0.94	6.6e+17
	V _{MT4}	-13.13	0.71	3.1e+19	-12.37	0.78	3.4e+18	-13.37	0.77	8.5e+17	-14.71	0.81	5.7e+17
V _{MT5}	-10.13	0.76	3.3e+19	-12.71	0.77	3.3e+18	-14.61	0.75	8.2e+17	-16.06	0.78	5.5e+17	
V _{B1}	-6.57	0.82	3.6e+19	-6.32	0.89	3.8e+18	-5.99	0.91	1e+18	-5.84	0.98	6.9e+17	
V _{B2}	-9.44	0.77	3.4e+19	-10.82	0.81	3.5e+18	-11.1	0.81	8.9e+17	-11.12	0.87	6.1e+17	
V _{B3}	-3.69	0.87	3.8e+19	-1.81	0.98	4.2e+18	-0.88	1.01	1.1e+18	-0.56	1.1	7.7e+17	
V _{B4}	-26.29	0.52	2.3e+19	-17.45	0.7	3e+18	-15.41	0.74	8.2e+17	-14.9	0.81	5.7e+17	
Tillé	V _{HT}	—	0.84	3.8e+19	—	1.11	4.1e+18	—	0.94	1.6e+18	—	1.09	7.8e+17
	V _{SYG}	—	1	4.6e+19	—	1	3.7e+18	—	1	1.7e+18	—	1	7.2e+17
	V _{MC1}	-0.6	0.85	3.9e+19	2.53	1.16	4.3e+18	0.43	0.95	1.6e+18	1.75	1.11	7.9e+17
	V _{MC2}	-1.31	0.85	3.9e+19	-3.71	1.03	3.8e+18	1.08	1.05	1.8e+18	-6	1.01	7.2e+17
	V _{D1}	-6.12	0.74	3.4e+19	-7.14	0.92	3.4e+18	-8.72	0.76	1.3e+18	-10.34	0.83	6e+17
	V _{D2}	-5.59	0.74	3.4e+19	-5.77	0.95	3.5e+18	-6.04	0.8	1.3e+18	-6.39	0.91	6.5e+17
	V _{D3}	-5.06	0.75	3.4e+19	-3.68	0.99	3.7e+18	-2.11	0.87	1.5e+18	-0.37	1.03	7.4e+17
	V _{FP}	-6.5	0.73	3.3e+19	1.61	1.11	4.1e+18	1.5	0.94	1.6e+18	1.93	1.09	7.8e+17
	V _{Be}	1.09	0.87	4e+19	-1.21	1.06	3.9e+18	-2.03	0.88	1.5e+18	-6.31	0.9	6.5e+17
	V _{T1}	1.54	0.88	4e+19	0.1	1.08	4e+18	0.4	0.92	1.5e+18	-1.55	0.99	7.1e+17
	V _{MT1}	-5.86	0.74	3.4e+19	-8.03	0.9	3.3e+18	-10.21	0.73	1.2e+18	-11.85	0.8	5.8e+17
	V _{MT2}	-4.87	0.76	3.5e+19	-3.76	0.99	3.7e+18	-2.28	0.87	1.5e+18	-0.84	1.02	7.3e+17
	V _{MT3}	-5.4	0.75	3.4e+19	-5.84	0.95	3.5e+18	-6.2	0.8	1.3e+18	-6.81	0.9	6.4e+17
	V _{MT4}	-12.09	0.65	3e+19	-11.16	0.85	3.1e+18	-12.3	0.7	1.2e+18	-13.76	0.77	5.5e+17
V _{MT5}	-9.46	0.68	3.1e+19	-11.51	0.84	3.1e+18	-13.56	0.68	1.1e+18	-15.1	0.75	5.3e+17	
V _{B1}	-5.59	0.74	3.4e+19	-5.08	0.96	3.6e+18	-4.91	0.82	1.4e+18	-4.6	0.94	6.7e+17	
V _{B2}	-8.88	0.69	3.1e+19	-9.64	0.87	3.2e+18	-10.1	0.73	1.2e+18	-9.94	0.84	6e+17	
V _{B3}	-2.3	0.8	3.7e+19	-0.52	1.06	3.9e+18	0.29	0.91	1.5e+18	0.74	1.05	7.5e+17	
V _{B4}	-24.52	0.49	2.2e+19	-16.37	0.75	2.8e+18	-14.37	0.67	1.1e+18	-13.8	0.78	5.5e+17	
CPS	V _{HT}	—	1.04	3.3e+19	—	1	5.4e+18	—	0.95	1.1e+18	—	0.96	9e+17
	V _{SYG}	—	1	3.2e+19	—	1	5.4e+18	—	1	1.2e+18	—	1	9.4e+17
	V _{MC1}	-1.02	1.04	3.3e+19	1.15	1.07	5.8e+18	1.03	0.97	1.2e+18	3.21	1.04	9.8e+17
	V _{MC2}	-0.43	1.11	3.5e+19	3.04	1.17	6.4e+18	1.47	1	1.2e+18	3.74	1.08	1e+18
	V _{D1}	-8.53	0.88	2.8e+19	-9.38	0.83	4.5e+18	-9.48	0.77	9.3e+17	-10.63	0.76	7.2e+17
	V _{D2}	-7.83	0.89	2.8e+19	-8.06	0.86	4.7e+18	-6.92	0.82	9.9e+17	-7.07	0.82	7.7e+17
	V _{D3}	-7.12	0.9	2.9e+19	-6.04	0.9	4.9e+18	-3.15	0.88	1.1e+18	-1.88	0.91	8.6e+17
	V _{FP}	-9.04	0.87	2.7e+19	-1.16	1	5.4e+18	0.18	0.95	1.1e+18	0.36	0.96	9e+17
	V _{Be}	-2.96	0.99	3.1e+19	-2.62	0.98	5.3e+18	-1.15	0.95	1.1e+18	-2.45	0.93	8.8e+17
	V _{T1}	-2.52	1	3.2e+19	-1.59	1	5.4e+18	0.63	0.97	1.2e+18	0.32	0.96	9.1e+17
	V _{MT1}	-7.59	0.9	2.8e+19	-10.33	0.81	4.4e+18	-10.97	0.75	9e+17	-12.21	0.73	6.9e+17
	V _{MT2}	-7.13	0.9	2.9e+19	-6.11	0.89	4.9e+18	-3.22	0.88	1.1e+18	-1.95	0.91	8.6e+17
	V _{MT3}	-7.83	0.89	2.8e+19	-8.12	0.86	4.7e+18	-6.97	0.82	9.8e+17	-7.11	0.82	7.7e+17
	V _{MT4}	-14.35	0.77	2.4e+19	-13.31	0.76	4.1e+18	-13.02	0.72	8.6e+17	-14.04	0.7	6.6e+17
V _{MT5}	-11.12	0.83	2.6e+19	-13.72	0.75	4.1e+18	-14.29	0.69	8.3e+17	-15.45	0.68	6.4e+17	
V _{B1}	-7.83	0.89	2.8e+19	-7.4	0.87	4.7e+18	-5.83	0.83	1e+18	-5.67	0.84	7.9e+17	
V _{B2}	-10.4	0.84	2.7e+19	-11.9	0.79	4.3e+18	-10.96	0.74	8.9e+17	-10.99	0.75	7e+17	
V _{B3}	-5.25	0.94	3e+19	-2.91	0.96	5.2e+18	-0.69	0.93	1.1e+18	-0.34	0.94	8.9e+17	
V _{B4}	-27.94	0.55	1.7e+19	-18.3	0.68	3.7e+18	-15.22	0.68	8.2e+17	-14.63	0.7	6.5e+17	
Rand. Sys.	V _{HT}	—	0.96	5.7e+19	—	0.95	3.9e+18	—	1.07	1.2e+18	—	0.97	8.3e+17
	V _{SYG}	—	1	5.9e+19	—	1	4.1e+18	—	1	1.2e+18	—	1	8.6e+17
	V _{MC1}	-1.31	0.93	5.5e+19	0.85	0.98	4e+18	1.4	1.13	1.3e+18	0.46	0.98	8.4e+17
	V _{MC2}	-6.96	0.83	4.9e+19	4.26	1.09	4.5e+18	-1.9	1.03	1.2e+18	1.3	1	8.6e+17
	V _{D1}	-6.58	0.85	5e+19	-7.4	0.81	3.4e+18	-11.39	0.84	9.8e+17	-10.68	0.78	6.7e+17
	V _{D2}	-6.05	0.86	5.1e+19	-6.28	0.83	3.4e+18	-8.68	0.9	1e+18	-7.36	0.83	7.2e+17
	V _{D3}	-5.53	0.87	5.1e+19	-4.6	0.86	3.6e+18	-4.46	0.98	1.1e+18	-2.57	0.91	7.9e+17
	V _{FP}	-6.96	0.84	5e+19	0.09	0.95	3.9e+18	-1.04	1.07	1.2e+18	-0.33	0.97	8.3e+17
	V _{Be}	0.65	1	5.9e+19	1.24	1	4.1e+18	-5.92	0.94	1.1e+18	-0.42	0.99	8.5e+17
	V _{T1}	1.09	1	5.9e+19	2.02	1.01	4.2e+18	-3.09	1	1.2e+18	1.39	1.01	8.6e+17
	V _{MT1}	-6.35	0.85	5e+19	-8.46	0.79	3.3e+18	-12.78	0.82	9.5e+17	-12.28	0.75	6.4e+17
	V _{MT2}	-5.34	0.87	5.1e+19	-4.44	0.87	3.6e+18	-4.75	0.98	1.1e+18	-2.39	0.92	7.9e+17
	V _{MT3}	-5.87	0.86	5.1e+19	-6.11	0.84	3.5e+18	-8.94	0.89	1e+18	-7.16	0.84	7.2e+17
	V _{MT4}	-12.52	0.74	4.4e+19	-11.41	0.75	3.1e+18	-14.86	0.78	9e+17	-14.09	0.72	6.2e+17
V _{MT5}	-9.93	0.79	4.6e+19	-11.92	0.73	3e+18	-16.03	0.76	8.8e+17	-15.52	0.69	6e+17	
V _{B1}	-6.05	0.86	5.1e+19	-5.75	0.84	3.5e+18	-7.31	0.92	1.1e+18	-6.09	0.85	7.3e+17	
V _{B2}	-9.35	0.8	4.7e+19	-10.41	0.76	3.1e+18	-12.33	0.82	9.6e+17	-11.4	0.75	6.5e+17	
V _{B3}	-2.76	0.92	5.4e+19	-1.09	0.93	3.8e+18	-2.3	1.03	1.2e+18	-0.78	0.95	8.2e+17	
V _{B4}	-24.83	0.56	3.3e+19	-16.65	0.66	2.7e+18	-16.66	0.75	8.7e+17	-14.97	0.7	6e+17	

Table B.54: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f . Sukhatme population ($N = 34$), $CV(Y) = 0.67$, $CV(X) = 0.96$

sampling	estimator	f=5%			f=10%			f=15%			f=20%		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	VHT	—	0.94	1.9e+15	—	0.97	2.2e+14	—	1	7.1e+13	—	1.01	4.2e+13
	VSYG	—	1	2e+15	—	1	2.3e+14	—	1	7.1e+13	—	1	4.2e+13
	VMC ₁	-0.36	0.97	1.9e+15	-1	0.96	2.2e+14	-1.03	0.98	7e+13	0.56	0.99	4.1e+13
	VMC ₂	-0.4	0.95	1.9e+15	-0.86	0.97	2.2e+14	-0.6	1	7.1e+13	0.67	1.03	4.3e+13
	VD ₁	-4.47	0.85	1.7e+15	-5.52	0.86	2e+14	-6.77	0.85	6.1e+13	-7.01	0.84	3.5e+13
	VD ₂	-3.97	0.85	1.7e+15	-4.66	0.88	2e+14	-5.23	0.88	6.3e+13	-4.96	0.88	3.6e+13
	VD ₃	-3.47	0.86	1.7e+15	-2.95	0.91	2.1e+14	-1.37	0.94	6.7e+13	0.7	0.96	4e+13
	VFP	-4.84	0.84	1.7e+15	-0.5	0.99	2.3e+14	-0.54	1	7.1e+13	-0.02	1	4.1e+13
	VBe	-0.56	0.94	1.9e+15	-1.02	0.97	2.2e+14	-1.52	0.98	7e+13	-1.4	0.98	4.1e+13
	VT ₁	-0.39	0.95	1.9e+15	-0.17	0.99	2.3e+14	0.85	1.01	7.2e+13	2.32	1.03	4.3e+13
	VMT ₁	-3.48	0.85	1.7e+15	-5.73	0.84	1.9e+14	-7.61	0.82	5.9e+13	-8.04	0.81	3.4e+13
	VMT ₂	-3.51	0.86	1.7e+15	-3.03	0.9	2.1e+14	-1.46	0.94	6.7e+13	0.61	0.96	4e+13
	VMT ₃	-4.01	0.85	1.7e+15	-4.73	0.88	2e+14	-5.31	0.88	6.2e+13	-5.04	0.88	3.6e+13
	VMT ₄	-9.28	0.76	1.5e+15	-8.75	0.8	1.8e+14	-9.64	0.8	5.7e+13	-9.8	0.79	3.3e+13
	VMT ₅	-6.32	0.8	1.6e+15	-8.49	0.8	1.8e+14	-10.28	0.78	5.5e+13	-10.68	0.77	3.2e+13
	VB ₁	-3.97	0.85	1.7e+15	-3.84	0.89	2e+14	-2.97	0.91	6.5e+13	-1.48	0.92	3.8e+13
VB ₂	-5.77	0.81	1.6e+15	-6.69	0.82	1.9e+14	-6.4	0.83	5.9e+13	-5.13	0.83	3.5e+13	
VB ₃	-2.17	0.9	1.8e+15	-0.98	0.96	2.2e+14	0.46	0.99	7.1e+13	2.16	1.01	4.2e+13	
VB ₄	-23.65	0.57	1.1e+15	-16.3	0.71	1.6e+14	-14.29	0.75	5.4e+13	-12.89	0.77	3.2e+13	
Tillé	VHT	—	1.07	2e+15	—	1.06	2.5e+14	—	1.07	6.5e+13	—	1.07	4.3e+13
	VSYG	—	1	1.8e+15	—	1	2.3e+14	—	1	6.1e+13	—	1	4e+13
	VMC ₁	-0.96	1.06	2e+15	0.98	1.08	2.5e+14	1.07	1.06	6.5e+13	1.45	1.08	4.3e+13
	VMC ₂	-1.18	0.97	1.8e+15	0.97	1.02	2.4e+14	0.61	1	6.1e+13	0.66	1.02	4.1e+13
	VD ₁	-2.13	0.9	1.7e+15	0.74	0.91	2.1e+14	1.93	0.91	5.6e+13	1.47	0.9	3.6e+13
	VD ₂	-1.62	0.91	1.7e+15	1.68	0.93	2.2e+14	3.65	0.94	5.8e+13	3.76	0.95	3.8e+13
	VD ₃	-1.11	0.92	1.7e+15	3.52	0.96	2.2e+14	8.03	1.01	6.2e+13	10.01	1.04	4.2e+13
	VFP	-2.51	0.9	1.7e+15	6.41	1.05	2.4e+14	9	1.07	6.6e+13	9.48	1.08	4.3e+13
	VBe	1.84	1.01	1.9e+15	5.18	1.02	2.4e+14	6.57	1.02	6.3e+13	6.41	1.02	4.1e+13
	VT ₁	2.02	1.01	1.9e+15	6.15	1.04	2.4e+14	9.48	1.06	6.5e+13	10.85	1.08	4.4e+13
	VMT ₁	-1.1	0.91	1.7e+15	0.49	0.89	2.1e+14	1.04	0.88	5.4e+13	0.33	0.87	3.5e+13
	VMT ₂	-1.15	0.92	1.7e+15	3.42	0.96	2.2e+14	7.89	1	6.2e+13	9.86	1.03	4.1e+13
	VMT ₃	-1.66	0.91	1.7e+15	1.59	0.93	2.1e+14	3.52	0.94	5.7e+13	3.62	0.94	3.8e+13
	VMT ₄	-7.06	0.81	1.5e+15	-2.7	0.85	2e+14	-1.22	0.85	5.2e+13	-1.57	0.85	3.4e+13
	VMT ₅	-4.01	0.86	1.6e+15	-2.45	0.84	1.9e+14	-1.89	0.83	5.1e+13	-2.55	0.82	3.3e+13
	VB ₁	-1.62	0.91	1.7e+15	2.56	0.94	2.2e+14	6.23	0.97	6e+13	7.58	0.99	4e+13
VB ₂	-3.45	0.87	1.6e+15	-0.53	0.87	2e+14	2.51	0.89	5.4e+13	3.56	0.9	3.6e+13	
VB ₃	0.21	0.96	1.8e+15	5.66	1.02	2.4e+14	9.95	1.07	6.5e+13	11.59	1.09	4.4e+13	
VB ₄	-21.82	0.61	1.1e+15	-10.63	0.75	1.7e+14	-6.26	0.8	4.9e+13	-4.77	0.83	3.3e+13	
CPS	VHT	—	0.99	1.9e+15	—	0.99	2.4e+14	—	0.98	6.9e+13	—	0.98	4.2e+13
	VSYG	—	1	1.9e+15	—	1	2.4e+14	—	1	7.1e+13	—	1	4.2e+13
	VMC ₁	0.12	0.99	1.9e+15	-0.19	0.98	2.4e+14	0.25	0.98	6.9e+13	0.14	0.97	4.1e+13
	VMC ₂	0.03	0.98	1.8e+15	-0.41	0.99	2.4e+14	-0.18	1	7e+13	0.06	0.99	4.2e+13
	VD ₁	-3.64	0.89	1.7e+15	-5.21	0.87	2.1e+14	-6.31	0.84	6e+13	-6.92	0.83	3.5e+13
	VD ₂	-3.16	0.9	1.7e+15	-4.34	0.89	2.1e+14	-4.76	0.87	6.2e+13	-4.85	0.87	3.7e+13
	VD ₃	-2.68	0.9	1.7e+15	-2.63	0.92	2.2e+14	-0.82	0.93	6.6e+13	0.85	0.95	4e+13
	VFP	-4	0.88	1.6e+15	-0.1	1	2.4e+14	-0.06	0.99	7e+13	0.13	0.99	4.2e+13
	VBe	0.45	0.99	1.9e+15	-0.68	0.99	2.4e+14	-1.17	0.97	6.9e+13	-1.51	0.96	4.1e+13
	VT ₁	0.62	1	1.9e+15	0.17	1	2.4e+14	1.26	1.01	7.1e+13	2.3	1.01	4.3e+13
	VMT ₁	-2.72	0.89	1.7e+15	-5.44	0.85	2e+14	-7.15	0.82	5.8e+13	-7.96	0.8	3.4e+13
	VMT ₂	-2.7	0.9	1.7e+15	-2.71	0.92	2.2e+14	-0.93	0.93	6.6e+13	0.73	0.95	4e+13
	VMT ₃	-3.18	0.9	1.7e+15	-4.41	0.89	2.1e+14	-4.85	0.87	6.2e+13	-4.95	0.87	3.7e+13
	VMT ₄	-8.5	0.8	1.5e+15	-8.45	0.81	1.9e+14	-9.2	0.79	5.6e+13	-9.72	0.78	3.3e+13
	VMT ₅	-5.58	0.84	1.6e+15	-8.2	0.8	1.9e+14	-9.84	0.77	5.5e+13	-10.6	0.76	3.2e+13
	VB ₁	-3.16	0.9	1.7e+15	-3.52	0.9	2.2e+14	-2.45	0.9	6.4e+13	-1.36	0.91	3.9e+13
VB ₂	-5.05	0.85	1.6e+15	-6.41	0.83	2e+14	-5.88	0.82	5.8e+13	-5	0.83	3.5e+13	
VB ₃	-1.28	0.94	1.8e+15	-0.62	0.97	2.3e+14	0.98	0.99	7e+13	2.29	1	4.2e+13	
VB ₄	-22.84	0.6	1.1e+15	-15.96	0.72	1.7e+14	-13.87	0.75	5.3e+13	-12.79	0.76	3.2e+13	
Rand. Sys.	VHT	—	1.02	2.1e+15	—	0.98	2.4e+14	—	0.98	6.5e+13	—	0.97	4.2e+13
	VSYG	—	1	2e+15	—	1	2.4e+14	—	1	6.6e+13	—	1	4.3e+13
	VMC ₁	0.46	1.02	2.1e+15	-0.96	0.97	2.4e+14	-0.82	0.97	6.4e+13	-0.52	0.98	4.2e+13
	VMC ₂	0.15	1	2e+15	-0.84	0.98	2.4e+14	-0.83	0.99	6.6e+13	-0.43	1	4.3e+13
	VD ₁	-3.94	0.89	1.8e+15	-6.06	0.86	2.1e+14	-8.45	0.83	5.5e+13	-10.33	0.81	3.5e+13
	VD ₂	-3.44	0.9	1.8e+15	-5.21	0.88	2.1e+14	-6.96	0.86	5.7e+13	-8.37	0.85	3.6e+13
	VD ₃	-2.94	0.91	1.8e+15	-3.54	0.91	2.2e+14	-3.13	0.92	6.1e+13	-3	0.93	4e+13
	VFP	-4.31	0.89	1.8e+15	-1.06	0.99	2.4e+14	-2.47	0.97	6.4e+13	-3.66	0.96	4.1e+13
	VBe	0.03	0.99	2e+15	-1.42	0.98	2.4e+14	-3.16	0.97	6.4e+13	-4.51	0.96	4.1e+13
	VT ₁	0.2	1	2e+15	-0.59	1	2.4e+14	-0.87	1	6.6e+13	-1.11	1	4.3e+13
	VMT ₁	-2.97	0.9	1.8e+15	-6.31	0.85	2.1e+14	-9.25	0.81	5.3e+13	-11.33	0.79	3.4e+13
	VMT ₂	-2.98	0.91	1.8e+15	-3.6	0.91	2.2e+14	-3.21	0.92	6.1e+13	-3.06	0.93	4e+13
	VMT ₃	-3.48	0.9	1.8e+15	-5.27	0.88	2.1e+14	-7.02	0.86	5.7e+13	-8.43	0.85	3.6e+13
	VMT ₄	-8.78	0.81	1.6e+15	-9.27	0.81	2e+14	-11.28	0.78	5.2e+13	-13.02	0.77	3.3e+13
	VMT ₅	-5.82	0.85	1.7e+15	-9.04	0.8	1.9e+14	-11.88	0.76	5e+13	-13.88	0.74	3.2e+13
	VB ₁	-3.44	0.9	1.8e+15	-4.41	0.89	2.2e+14	-4.69	0.89	5.9e+13	-5.08	0.89	3.8e+13
VB ₂	-5.27	0.86	1.7e+15	-7.29	0.82	2e+14	-8.02	0.81	5.4e+13	-8.59	0.81	3.4e+13	
VB ₃	-1.62	0.95	1.9e+15	-1.52	0.97	2.4e+14	-1.37	0.97	6.4e+13	-1.56	0.97	4.2e+13	
VB ₄	-23.19	0.6	1.2e+15	-16.71	0.71	1.7e+14	-15.93	0.74	4.9e+13	-16.05	0.74	3.2e+13	

Table B.55: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f . Sukhatme population ($N = 34$), $CV(Y) = 0.67$, $CV(X) = 0.77$

sampling	estimator	$f=5\%$			$f=10\%$			$f=15\%$			$f=20\%$		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	VHT	—	1.02	2.1e+14	—	1	3.3e+13	—	1.04	9.1e+12	—	1.03	5.4e+12
	VSYG	—	1	2e+14	—	1	3.3e+13	—	1	8.7e+12	—	1	5.3e+12
	VMC1	1.8	1.01	2.1e+14	0.32	1.03	3.4e+13	0.19	1.07	9.3e+12	-1.56	1.03	5.4e+12
	VMC2	1.83	1	2.1e+14	0.16	1.02	3.4e+13	0.11	1.02	8.8e+12	-0.28	1.01	5.3e+12
	VD1	1.41	0.89	1.8e+14	-1.76	0.88	2.9e+13	-2.2	0.88	7.7e+12	-3.16	0.87	4.6e+12
	VD2	1.76	0.9	1.8e+14	-1.21	0.89	2.9e+13	-1.28	0.9	7.8e+12	-2.01	0.89	4.7e+12
	VD3	2.12	0.9	1.8e+14	-0.14	0.91	3e+13	1.13	0.94	8.1e+12	1.35	0.93	4.9e+12
	VFP	1.15	0.88	1.8e+14	-0.45	0.98	3.2e+13	-0.2	1	8.7e+12	-0.75	0.99	5.2e+12
	VBe	1.6	0.96	2e+14	-0.37	0.97	3.2e+13	-0.18	0.98	8.5e+12	-0.34	0.99	5.2e+12
	Vt1	1.4	0.96	2e+14	-0.45	0.98	3.2e+13	0.21	1	8.7e+12	0.43	1	5.3e+12
	VMT1	3.16	0.89	1.8e+14	-1.09	0.85	2.8e+13	-2.26	0.85	7.4e+12	-3.6	0.83	4.4e+12
	VMT2	2.07	0.9	1.8e+14	-0.2	0.91	3e+13	1.05	0.93	8.1e+12	1.28	0.93	4.9e+12
	VMT3	1.72	0.89	1.8e+14	-1.26	0.89	2.9e+13	-1.35	0.9	7.8e+12	-2.07	0.89	4.7e+12
	VMT4	-1.89	0.83	1.7e+14	-3.91	0.84	2.8e+13	-4.14	0.85	7.4e+12	-5.03	0.83	4.4e+12
	VMT5	1.23	0.86	1.8e+14	-2.93	0.82	2.7e+13	-4.07	0.82	7.1e+12	-5.37	0.8	4.2e+12
	VB1	1.76	0.9	1.8e+14	-0.7	0.9	2.9e+13	0.18	0.92	8e+12	0.16	0.91	4.8e+12
VB2	1.77	0.86	1.8e+14	-1.12	0.84	2.7e+13	-0.21	0.85	7.4e+12	-0.22	0.84	4.4e+12	
VB3	1.75	0.93	1.9e+14	-0.27	0.96	3.2e+13	0.58	0.99	8.6e+12	0.53	0.98	5.2e+12	
VB4	-19	0.63	1.3e+14	-14.9	0.75	2.5e+13	-14.37	0.79	6.9e+12	-14.96	0.79	4.2e+12	
Tillé	VHT	—	1.13	2.5e+14	—	1.05	2.8e+13	—	1.18	9.2e+12	—	1.17	5.3e+12
	VSYG	—	1	2.2e+14	—	1	2.7e+13	—	1	7.8e+12	—	1	4.5e+12
	VMC1	-0.45	1.14	2.5e+14	-0.51	1.05	2.8e+13	1.12	1.18	9.2e+12	0.62	1.17	5.3e+12
	VMC2	-1.88	0.98	2.2e+14	-1.33	0.97	2.6e+13	-0.91	1	7.8e+12	-1.83	1	4.5e+12
	VD1	-0.02	0.96	2.1e+14	0.02	0.9	2.4e+13	1.05	0.95	7.4e+12	0.03	0.94	4.2e+12
	VD2	0.32	0.96	2.2e+14	0.57	0.91	2.5e+13	2	0.97	7.5e+12	1.23	0.96	4.3e+12
	VD3	0.66	0.97	2.2e+14	1.67	0.93	2.5e+13	4.53	1.01	7.8e+12	4.83	1.01	4.6e+12
	VFP	-0.27	0.95	2.1e+14	1.02	1	2.7e+13	3.11	1.08	8.4e+12	2.54	1.08	4.9e+12
	VBe	0.54	1.04	2.3e+14	1.34	1	2.7e+13	2.92	1.05	8.2e+12	2.24	1.04	4.7e+12
	Vt1	0.38	1.04	2.3e+14	1.22	1.01	2.7e+13	3.44	1.07	8.3e+12	3.31	1.07	4.8e+12
	VMT1	1.53	0.95	2.1e+14	0.84	0.87	2.4e+13	1.02	0.91	7.1e+12	-0.36	0.9	4.1e+12
	VMT2	0.62	0.97	2.2e+14	1.61	0.93	2.5e+13	4.46	1	7.8e+12	4.75	1.01	4.6e+12
	VMT3	0.29	0.96	2.1e+14	0.52	0.91	2.5e+13	1.92	0.97	7.5e+12	1.16	0.96	4.3e+12
	VMT4	-3.27	0.9	2e+14	-2.18	0.86	2.3e+13	-0.95	0.91	7.1e+12	-1.9	0.9	4.1e+12
	VMT5	-0.37	0.92	2.1e+14	-1.04	0.84	2.3e+13	-0.85	0.88	6.9e+12	-2.19	0.87	3.9e+12
	VB1	0.32	0.96	2.2e+14	1.1	0.92	2.5e+13	3.56	0.99	7.7e+12	3.58	0.99	4.5e+12
VB2	0.14	0.93	2.1e+14	0.86	0.85	2.3e+13	3.2	0.91	7.1e+12	3.29	0.91	4.1e+12	
VB3	0.49	1	2.2e+14	1.35	0.98	2.7e+13	3.91	1.06	8.3e+12	3.87	1.07	4.8e+12	
VB4	-19.71	0.69	1.5e+14	-13.72	0.77	2.1e+13	-11.59	0.85	6.6e+12	-12.28	0.86	3.9e+12	
CPS	VHT	—	0.98	2.7e+14	—	1.04	2.6e+13	—	1.02	8.6e+12	—	1.02	5.9e+12
	VSYG	—	1	2.7e+14	—	1	2.5e+13	—	1	8.4e+12	—	1	5.7e+12
	VMC1	-0.83	1	2.7e+14	1.16	1.06	2.7e+13	0.43	0.98	8.2e+12	0.4	1.03	5.9e+12
	VMC2	-1.23	0.99	2.7e+14	0.16	1.01	2.5e+13	0.04	0.96	8.1e+12	-0.25	1	5.7e+12
	VD1	-1.75	0.9	2.5e+14	-0.5	0.9	2.3e+13	-1.11	0.87	7.3e+12	-2.99	0.87	5e+12
	VD2	-1.45	0.9	2.5e+14	0.06	0.91	2.3e+13	-0.2	0.89	7.5e+12	-1.84	0.89	5.1e+12
	VD3	-1.15	0.91	2.5e+14	1.16	0.93	2.3e+13	2.21	0.92	7.8e+12	1.52	0.93	5.3e+12
	VFP	-1.98	0.89	2.4e+14	0.49	1.01	2.6e+13	0.76	0.99	8.3e+12	-0.39	0.99	5.7e+12
	VBe	-0.71	1	2.7e+14	0.51	0.99	2.5e+13	1.11	0.99	8.3e+12	-0.24	0.98	5.6e+12
	Vt1	-0.86	1	2.7e+14	0.38	1	2.5e+13	1.42	1	8.4e+12	0.58	1	5.7e+12
	VMT1	-0.46	0.89	2.4e+14	0.38	0.88	2.2e+13	-1.15	0.84	7.1e+12	-3.48	0.83	4.8e+12
	VMT2	-1.15	0.91	2.5e+14	1.1	0.93	2.3e+13	2.15	0.92	7.8e+12	1.46	0.93	5.3e+12
	VMT3	-1.45	0.9	2.5e+14	0	0.91	2.3e+13	-0.26	0.89	7.5e+12	-1.9	0.89	5.1e+12
	VMT4	-4.95	0.84	2.3e+14	-2.68	0.86	2.2e+13	-3.07	0.84	7.1e+12	-4.86	0.83	4.8e+12
	VMT5	-2.32	0.85	2.3e+14	-1.49	0.84	2.1e+13	-2.97	0.81	6.8e+12	-5.25	0.8	4.6e+12
	VB1	-1.45	0.9	2.5e+14	0.6	0.92	2.3e+13	1.27	0.91	7.6e+12	0.33	0.91	5.2e+12
VB2	-1.85	0.86	2.4e+14	0.44	0.86	2.2e+13	0.94	0.84	7e+12	-0.16	0.84	4.8e+12	
VB3	-1.05	0.95	2.6e+14	0.75	0.99	2.5e+13	1.61	0.98	8.2e+12	0.82	0.98	5.6e+12	
VB4	-20.58	0.66	1.8e+14	-14.32	0.77	1.9e+13	-13.57	0.78	6.6e+12	-14.55	0.79	4.5e+12	
Rand. Sys.	VHT	—	1.05	2e+14	—	1.03	2.9e+13	—	1	9e+12	—	1.03	5.7e+12
	VSYG	—	1	1.9e+14	—	1	2.8e+13	—	1	9e+12	—	1	5.6e+12
	VMC1	3.65	1.12	2.1e+14	-0.79	1.02	2.9e+13	-0.7	0.98	8.9e+12	1.45	1.03	5.8e+12
	VMC2	2.23	1.02	1.9e+14	-0.32	0.98	2.8e+13	-0.71	0.98	8.9e+12	0.89	1	5.6e+12
	VD1	1.92	0.92	1.7e+14	-1.39	0.9	2.5e+13	-3.28	0.87	7.8e+12	-2.79	0.87	4.8e+12
	VD2	2.27	0.93	1.8e+14	-0.84	0.91	2.6e+13	-2.4	0.88	8e+12	-1.64	0.89	4.9e+12
	VD3	2.62	0.93	1.8e+14	0.24	0.93	2.6e+13	-0.1	0.92	8.2e+12	1.73	0.93	5.2e+12
	VFP	1.66	0.91	1.7e+14	-0.28	1.01	2.9e+13	-1.39	0.98	8.8e+12	-0.26	0.99	5.5e+12
	VBe	2.11	1	1.9e+14	-0.13	0.99	2.8e+13	-0.74	0.99	8.9e+12	-0.09	0.98	5.5e+12
	Vt1	1.89	1	1.9e+14	-0.23	1	2.8e+13	-0.5	1	9e+12	0.67	1	5.6e+12
	VMT1	3.68	0.92	1.7e+14	-0.62	0.87	2.5e+13	-3.37	0.83	7.5e+12	-3.26	0.83	4.6e+12
	VMT2	2.58	0.93	1.8e+14	0.18	0.93	2.6e+13	-0.14	0.91	8.2e+12	1.66	0.93	5.2e+12
	VMT3	2.23	0.92	1.8e+14	-0.89	0.91	2.6e+13	-2.44	0.88	8e+12	-1.7	0.89	4.9e+12
	VMT4	-1.4	0.86	1.6e+14	-3.55	0.86	2.4e+13	-5.19	0.83	7.5e+12	-4.66	0.83	4.6e+12
	VMT5	1.74	0.88	1.7e+14	-2.47	0.84	2.4e+13	-5.15	0.8	7.2e+12	-5.04	0.8	4.5e+12
	VB1	2.27	0.93	1.8e+14	-0.32	0.92	2.6e+13	-1	0.9	8.1e+12	0.53	0.91	5.1e+12
VB2	2.29	0.89	1.7e+14	-0.61	0.86	2.4e+13	-1.43	0.83	7.5e+12	0.1	0.84	4.7e+12	
VB3	2.25	0.96	1.8e+14	-0.02	0.98	2.8e+13	-0.57	0.97	8.7e+12	0.97	0.98	5.5e+12	
VB4	-18.62	0.66	1.3e+14	-14.83	0.77	2.2e+13	-15.29	0.78	7e+12	-14.5	0.79	4.4e+12	

Table B.56: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f. Sukhatme population (N = 34), CV(Y) = 0.67, CV(X) = 0.67

sampling	estimator	f=5%			f=10%			f=15%			f=20%		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	VHT	—	4.95	6.7e+12	—	12.85	9.6e+11	—	35.19	3.7e+11	—	57.79	2.6e+11
	VSYG	—	1	1.4e+12	—	1	7.5e+10	—	1	1.1e+10	—	1	4.5e+09
	VMC1	1.15	2.99	4.1e+12	-0.75	10.01	7.5e+11	-2.49	31.06	3.3e+11	-1.64	53.31	2.4e+11
	VMC2	1.38	1.1	1.5e+12	-0.87	1.07	8e+10	-1	1.06	1.1e+10	-0.5	1.05	4.7e+09
	VD1	3.74	1.2	1.6e+12	1.45	1.19	8.9e+10	1.19	1.16	1.2e+10	1.71	1.12	5e+09
	VD2	3.95	1.21	1.6e+12	1.84	1.2	9e+10	1.84	1.18	1.3e+10	2.53	1.14	5.2e+09
	VD3	4.15	1.22	1.6e+12	2.43	1.22	9.1e+10	2.95	1.22	1.3e+10	3.98	1.19	5.4e+09
	VFP	3.58	1.2	1.6e+12	-0.99	1.1	8.2e+10	-1	1.08	1.1e+10	-0.29	1.06	4.8e+09
	VB _e	1.58	1.11	1.5e+12	-0.21	1.1	8.2e+10	0.1	1.1	1.2e+10	0.92	1.08	4.9e+09
	VT _i	1.19	1.1	1.5e+12	-1.07	1.07	8e+10	-1.35	1.06	1.1e+10	-0.93	1.05	4.8e+09
	VMT1	6.13	1.28	1.7e+12	3.72	1.27	9.5e+10	2.78	1.21	1.3e+10	2.89	1.15	5.2e+09
	VMT2	4.19	1.22	1.6e+12	2.42	1.22	9.1e+10	2.93	1.21	1.3e+10	3.96	1.18	5.3e+09
	VMT3	3.98	1.21	1.6e+12	1.83	1.2	9e+10	1.82	1.17	1.2e+10	2.52	1.13	5.1e+09
	VMT4	1.17	1.14	1.6e+12	-0.24	1.15	8.6e+10	-0.33	1.12	1.2e+10	0.22	1.08	4.9e+09
	VMT5	4.63	1.24	1.7e+12	2.26	1.23	9.2e+10	1.34	1.17	1.2e+10	1.46	1.11	5e+09
	VB ₁	3.95	1.21	1.6e+12	2.04	1.21	9e+10	2.29	1.2	1.3e+10	3.14	1.16	5.3e+09
	VB ₂	5.13	1.26	1.7e+12	4.18	1.3	9.7e+10	5.17	1.32	1.4e+10	6.48	1.29	5.9e+09
VB ₃	2.77	1.16	1.6e+12	-0.1	1.12	8.4e+10	-0.58	1.09	1.2e+10	-0.19	1.05	4.8e+09	
VB ₄	-17.96	0.73	9.9e+11	-15.74	0.81	6e+10	-17.32	0.83	8.8e+09	-18.36	0.85	3.8e+09	
Tillé	VHT	—	3.75	5.6e+12	—	11.45	9.4e+11	—	33.27	3.9e+11	—	53.81	2.7e+11
	VSYG	—	1	1.5e+12	—	1	8.2e+10	—	1	1.2e+10	—	1	5e+09
	VMC1	0.49	3.98	5.9e+12	1.93	11.49	9.5e+11	-2.54	33.35	3.9e+11	-0.99	53.99	2.7e+11
	VMC2	1.25	1	1.5e+12	0.83	1	8.2e+10	-0.07	1	1.2e+10	-0.38	1	5e+09
	VD1	3.15	1.08	1.6e+12	2.73	1.08	8.9e+10	1.02	1.05	1.2e+10	0.2	1.02	5.1e+09
	VD2	3.35	1.08	1.6e+12	3.13	1.09	9e+10	1.68	1.07	1.2e+10	1.02	1.04	5.2e+09
	VD3	3.56	1.09	1.6e+12	3.74	1.11	9.2e+10	2.79	1.1	1.3e+10	2.46	1.08	5.4e+09
	VFP	2.99	1.07	1.6e+12	0.24	1	8.2e+10	-1.17	0.98	1.1e+10	-1.74	0.97	4.8e+09
	VB _e	1.04	1	1.5e+12	1.09	1	8.3e+10	0.06	1	1.2e+10	-0.48	0.99	4.9e+09
	VT _i	0.66	0.98	1.5e+12	0.25	0.97	8e+10	-1.3	0.96	1.1e+10	-2.15	0.97	4.8e+09
	VMT1	5.5	1.15	1.7e+12	5.02	1.16	9.5e+10	2.6	1.1	1.3e+10	1.36	1.04	5.2e+09
	VMT2	3.59	1.09	1.6e+12	3.72	1.11	9.1e+10	2.77	1.1	1.3e+10	2.44	1.07	5.4e+09
	VMT3	3.38	1.08	1.6e+12	3.11	1.09	9e+10	1.65	1.07	1.2e+10	1	1.03	5.1e+09
	VMT4	0.59	1.02	1.5e+12	1.02	1.05	8.6e+10	-0.49	1.02	1.2e+10	-1.26	0.99	4.9e+09
	VMT5	4.01	1.11	1.7e+12	3.55	1.12	9.2e+10	1.17	1.07	1.2e+10	-0.05	1.01	5.1e+09
	VB ₁	3.35	1.08	1.6e+12	3.33	1.1	9.1e+10	2.13	1.08	1.3e+10	1.63	1.06	5.3e+09
	VB ₂	4.5	1.13	1.7e+12	5.49	1.19	9.8e+10	4.99	1.19	1.4e+10	4.88	1.17	5.8e+09
VB ₃	2.2	1.04	1.5e+12	1.17	1.02	8.4e+10	-0.72	0.99	1.1e+10	-1.61	0.96	4.8e+09	
VB ₄	-18.38	0.65	9.7e+11	-14.66	0.73	6e+10	-17.41	0.76	8.8e+09	-19.47	0.8	4e+09	
CPS	VHT	—	2.61	3.8e+12	—	9.37	7.5e+11	—	30.01	3.3e+11	—	50.62	2.4e+11
	VSYG	—	1	1.5e+12	—	1	8e+10	—	1	1.1e+10	—	1	4.8e+09
	VMC1	-1.26	2.74	4e+12	-1.22	9.45	7.5e+11	1.6	30.14	3.3e+11	-0.29	50.74	2.4e+11
	VMC2	-0.86	1	1.5e+12	-1.62	1	8e+10	0.25	0.99	1.1e+10	-0.16	0.99	4.7e+09
	VD1	1.71	1.09	1.6e+12	0.86	1.11	8.9e+10	2.64	1.1	1.2e+10	2.02	1.06	5.1e+09
	VD2	1.92	1.1	1.6e+12	1.25	1.13	9e+10	3.29	1.12	1.2e+10	2.85	1.09	5.2e+09
	VD3	2.12	1.11	1.6e+12	1.84	1.14	9.1e+10	4.41	1.16	1.3e+10	4.29	1.13	5.4e+09
	VFP	1.56	1.09	1.6e+12	-1.57	1.03	8.2e+10	0.41	1.02	1.1e+10	0.03	1.01	4.8e+09
	VB _e	-0.36	1.01	1.5e+12	-0.81	1.03	8.2e+10	1.46	1.04	1.2e+10	1.21	1.03	5e+09
	VT _i	-0.73	1	1.5e+12	-1.66	1	7.9e+10	-0.05	1	1.1e+10	-0.66	1.01	4.8e+09
	VMT1	4.03	1.16	1.7e+12	3.12	1.19	9.5e+10	4.25	1.15	1.3e+10	3.2	1.09	5.2e+09
	VMT2	2.15	1.11	1.6e+12	1.83	1.14	9.1e+10	4.4	1.15	1.3e+10	4.28	1.13	5.4e+09
	VMT3	1.95	1.1	1.6e+12	1.24	1.12	8.9e+10	3.28	1.12	1.2e+10	2.83	1.08	5.2e+09
	VMT4	-0.8	1.04	1.5e+12	-0.81	1.08	8.6e+10	1.1	1.07	1.2e+10	0.53	1.03	4.9e+09
	VMT5	2.56	1.13	1.7e+12	1.67	1.15	9.2e+10	2.8	1.12	1.2e+10	1.76	1.06	5.1e+09
	VB ₁	1.92	1.1	1.6e+12	1.45	1.13	9e+10	3.75	1.14	1.3e+10	3.45	1.11	5.3e+09
	VB ₂	3.05	1.15	1.7e+12	3.58	1.22	9.7e+10	6.68	1.26	1.4e+10	6.79	1.23	5.9e+09
VB ₃	0.79	1.06	1.5e+12	-0.68	1.05	8.4e+10	0.82	1.03	1.1e+10	0.12	1	4.8e+09	
VB ₄	-19.5	0.67	9.8e+11	-16.24	0.76	6e+10	-16.17	0.77	8.5e+09	-18.1	0.81	3.9e+09	
Rand. Sys.	VHT	—	2.52	3.8e+12	—	8.71	7.1e+11	—	27.17	3.1e+11	—	45.75	2.3e+11
	VSYG	—	1	1.5e+12	—	1	8.1e+10	—	1	1.1e+10	—	1	4.9e+09
	VMC1	1.6	2.74	4.1e+12	-1.02	9.17	7.4e+11	1.62	28.69	3.2e+11	1.83	48.08	2.4e+11
	VMC2	1.03	0.99	1.5e+12	0.37	0.99	8e+10	-0.34	0.97	1.1e+10	1.3	0.96	4.8e+09
	VD1	3.3	1.09	1.6e+12	2.77	1.1	8.9e+10	2.31	1.07	1.2e+10	4.25	1.03	5.1e+09
	VD2	3.51	1.1	1.6e+12	3.16	1.11	9e+10	2.97	1.09	1.2e+10	5.09	1.06	5.2e+09
	VD3	3.71	1.1	1.6e+12	3.76	1.13	9.1e+10	4.08	1.13	1.3e+10	6.57	1.1	5.5e+09
	VFP	3.15	1.09	1.6e+12	0.34	1.01	8.2e+10	0.08	1	1.1e+10	2.19	0.97	4.8e+09
	VB _e	1.13	1.01	1.5e+12	1.12	1.01	8.2e+10	1.13	1.01	1.1e+10	3.38	1	4.9e+09
	VT _i	0.74	0.99	1.5e+12	0.27	0.98	8e+10	-0.39	0.97	1.1e+10	1.4	0.97	4.8e+09
	VMT1	5.69	1.16	1.7e+12	5.04	1.17	9.5e+10	3.92	1.12	1.3e+10	5.46	1.06	5.3e+09
	VMT2	3.75	1.1	1.6e+12	3.75	1.12	9.1e+10	4.07	1.12	1.3e+10	6.55	1.1	5.4e+09
	VMT3	3.54	1.1	1.6e+12	3.15	1.11	9e+10	2.95	1.09	1.2e+10	5.08	1.05	5.2e+09
	VMT4	0.75	1.04	1.5e+12	1.06	1.06	8.6e+10	0.78	1.04	1.2e+10	2.73	1	4.9e+09
	VMT5	4.2	1.13	1.7e+12	3.57	1.14	9.2e+10	2.47	1.09	1.2e+10	3.99	1.03	5.1e+09
	VB ₁	3.51	1.1	1.6e+12	3.36	1.12	9e+10	3.42	1.11	1.3e+10	5.71	1.08	5.3e+09
	VB ₂	4.69	1.14	1.7e+12	5.49	1.2	9.8e+10	6.36	1.22	1.4e+10	9.14	1.2	6e+09
VB ₃	2.32	1.05	1.6e+12	1.23	1.03	8.4e+10	0.49	1.01	1.1e+10	2.28	0.97	4.8e+09	
VB ₄	-18.33	0.66	9.8e+11	-14.58	0.74	6e+10	-16.46	0.76	8.5e+09	-16.36	0.75	3.7e+09	

Table B.57: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f. Sukhatme population (N = 34), CV(Y) = 0.67, CV(X) = 0.5

sampling	estimator	f=5%			f=10%			f=15%			f=20%		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	V _{HT}	—	1.61	6e+12	—	2.82	6.6e+11	—	4.75	2.4e+11	—	5.95	1.7e+11
	V _{SYG}	—	1	3.7e+12	—	1	2.3e+11	—	1	5e+10	—	1	2.8e+10
	V _{MC1}	1.09	1.19	4.4e+12	-0.78	2.16	5.1e+11	-1.44	3.89	2e+11	-0.09	4.98	1.4e+11
	V _{MC2}	0.97	1.02	3.8e+12	-0.72	1	2.3e+11	-0.86	0.99	5e+10	0.7	0.99	2.8e+10
	V _{D1}	1.29	1.03	3.8e+12	-0.52	1.01	2.4e+11	-0.75	0.98	4.9e+10	0.76	0.96	2.7e+10
	V _{D2}	1.39	1.03	3.8e+12	-0.34	1.01	2.4e+11	-0.44	0.98	4.9e+10	1.16	0.97	2.7e+10
	V _{D3}	1.49	1.04	3.9e+12	-0.04	1.02	2.4e+11	0.16	1	5e+10	2	0.99	2.8e+10
	V _{Fp}	1.22	1.03	3.8e+12	-0.88	1.01	2.4e+11	-0.96	1	5e+10	0.63	0.98	2.7e+10
	V _{Be}	0.89	1.02	3.8e+12	-0.73	1.01	2.4e+11	-0.79	0.99	5e+10	0.82	0.98	2.7e+10
	V _{T1}	0.82	1.02	3.8e+12	-0.82	1.01	2.4e+11	-0.8	1.01	5.1e+10	0.9	1.01	2.8e+10
	V _{MT1}	2.24	1.05	3.9e+12	0.3	1.02	2.4e+11	-0.16	0.98	4.9e+10	1.23	0.95	2.7e+10
	V _{MT2}	1.52	1.04	3.9e+12	-0.03	1.02	2.4e+11	0.17	1	5e+10	2	0.99	2.8e+10
	V _{MT3}	1.42	1.03	3.8e+12	-0.32	1.01	2.4e+11	-0.43	0.98	4.9e+10	1.17	0.96	2.7e+10
	V _{MT4}	-0.13	1	3.7e+12	-1.46	0.99	2.3e+11	-1.59	0.96	4.8e+10	-0.07	0.94	2.6e+10
	V _{MT5}	1.43	1.04	3.9e+12	-0.49	1	2.3e+11	-0.95	0.96	4.8e+10	0.44	0.94	2.6e+10
	V _{B1}	1.39	1.03	3.8e+12	-0.23	1.01	2.4e+11	-0.15	1	5e+10	1.59	0.98	2.7e+10
	V _{B2}	1.63	1.04	3.9e+12	0.17	1.02	2.4e+11	0.33	0.99	5e+10	2.13	0.97	2.7e+10
V _{B3}	1.15	1.03	3.8e+12	-0.63	1.01	2.4e+11	-0.63	1.01	5e+10	1.06	1	2.8e+10	
V _{B4}	-15.15	0.74	2.7e+12	-12.33	0.83	1.9e+11	-12.27	0.86	4.3e+10	-11.1	0.85	2.4e+10	
Tillé	V _{HT}	—	1.36	5.1e+12	—	2.33	5.1e+11	—	3.95	1.7e+11	—	4.95	1.1e+11
	V _{SYG}	—	1	3.7e+12	—	1	2.2e+11	—	1	4.3e+10	—	1	2.3e+10
	V _{MC1}	2.08	1.42	5.3e+12	-0.68	2.35	5.1e+11	-1.4	3.97	1.7e+11	-0.36	4.97	1.1e+11
	V _{MC2}	2.14	1.01	3.8e+12	-0.73	1.01	2.2e+11	-0.83	1	4.3e+10	-0.42	1	2.3e+10
	V _{D1}	3.22	1.05	3.9e+12	1.47	1.09	2.4e+11	2.65	1.14	4.9e+10	3.81	1.17	2.7e+10
	V _{D2}	3.33	1.06	3.9e+12	1.66	1.1	2.4e+11	2.98	1.15	4.9e+10	4.22	1.18	2.7e+10
	V _{D3}	3.43	1.06	4e+12	1.96	1.11	2.4e+11	3.6	1.17	5e+10	5.08	1.22	2.8e+10
	V _{Fp}	3.15	1.05	3.9e+12	1.15	1.11	2.4e+11	2.45	1.16	5e+10	3.69	1.2	2.7e+10
	V _{Be}	2.81	1.04	3.9e+12	1.25	1.09	2.4e+11	2.63	1.16	5e+10	3.84	1.2	2.7e+10
	V _{T1}	2.74	1.04	3.9e+12	1.17	1.09	2.4e+11	2.64	1.17	5e+10	3.95	1.23	2.8e+10
	V _{MT1}	4.19	1.07	4e+12	2.29	1.11	2.4e+11	3.25	1.14	4.9e+10	4.29	1.17	2.7e+10
	V _{MT2}	3.46	1.06	4e+12	1.97	1.11	2.4e+11	3.61	1.17	5e+10	5.09	1.22	2.8e+10
	V _{MT3}	3.35	1.06	3.9e+12	1.67	1.1	2.4e+11	2.98	1.15	4.9e+10	4.22	1.18	2.7e+10
	V _{MT4}	1.77	1.02	3.8e+12	0.51	1.07	2.3e+11	1.78	1.12	4.8e+10	2.95	1.15	2.6e+10
	V _{MT5}	3.36	1.06	4e+12	1.48	1.09	2.4e+11	2.44	1.12	4.8e+10	3.47	1.14	2.6e+10
	V _{B1}	3.33	1.06	3.9e+12	1.77	1.1	2.4e+11	3.27	1.16	5e+10	4.67	1.2	2.7e+10
	V _{B2}	3.57	1.06	4e+12	2.15	1.1	2.4e+11	3.75	1.16	5e+10	5.21	1.19	2.7e+10
V _{B3}	3.08	1.05	3.9e+12	1.38	1.1	2.4e+11	2.79	1.17	5e+10	4.13	1.22	2.8e+10	
V _{B4}	-13.54	0.75	2.8e+12	-10.54	0.89	2e+11	-9.23	0.97	4.2e+10	-8.39	1	2.3e+10	
CPS	V _{HT}	—	1.12	4.3e+12	—	2.15	5.1e+11	—	3.86	1.9e+11	—	5.12	1.4e+11
	V _{SYG}	—	1	3.9e+12	—	1	2.4e+11	—	1	5e+10	—	1	2.7e+10
	V _{MC1}	0.27	1.15	4.5e+12	-1.3	2.16	5.1e+11	1.1	3.88	1.9e+11	-1.5	5.13	1.4e+11
	V _{MC2}	0.27	1	3.9e+12	-1.41	1	2.4e+11	-0.42	1	5e+10	-2.05	0.99	2.7e+10
	V _{D1}	0.74	1.01	3.9e+12	-1.01	1	2.4e+11	0.03	0.98	4.9e+10	-1.95	0.97	2.6e+10
	V _{D2}	0.84	1.01	3.9e+12	-0.83	1.01	2.4e+11	0.35	0.99	5e+10	-1.56	0.98	2.7e+10
	V _{D3}	0.94	1.02	3.9e+12	-0.54	1.01	2.4e+11	0.95	1.01	5.1e+10	-0.76	1	2.7e+10
	V _{Fp}	0.67	1.01	3.9e+12	-1.34	1.01	2.4e+11	-0.17	1.01	5.1e+10	-2.08	0.99	2.7e+10
	V _{Be}	0.38	1	3.9e+12	-1.23	1	2.4e+11	-0.09	1	5e+10	-1.94	0.99	2.7e+10
	V _{T1}	0.31	1	3.9e+12	-1.31	1	2.4e+11	-0.12	1.01	5.1e+10	-1.88	1.02	2.8e+10
	V _{MT1}	1.67	1.03	4e+12	-0.2	1.01	2.4e+11	0.63	0.98	5e+10	-1.49	0.96	2.6e+10
	V _{MT2}	0.97	1.02	3.9e+12	-0.52	1.01	2.4e+11	0.96	1.01	5.1e+10	-0.75	1	2.7e+10
	V _{MT3}	0.87	1.01	3.9e+12	-0.81	1.01	2.4e+11	0.36	0.99	5e+10	-1.55	0.98	2.6e+10
	V _{MT4}	-0.68	0.98	3.8e+12	-1.95	0.98	2.3e+11	-0.81	0.97	4.9e+10	-2.76	0.95	2.6e+10
	V _{MT5}	0.86	1.01	3.9e+12	-0.99	1	2.4e+11	-0.16	0.97	4.9e+10	-2.26	0.95	2.6e+10
	V _{B1}	0.84	1.01	3.9e+12	-0.72	1.01	2.4e+11	0.64	1	5e+10	-1.15	0.99	2.7e+10
	V _{B2}	1.06	1.02	4e+12	-0.34	1.01	2.4e+11	1.13	1	5e+10	-0.62	0.98	2.7e+10
V _{B3}	0.62	1.01	3.9e+12	-1.1	1.01	2.4e+11	0.14	1.01	5.1e+10	-1.68	1.01	2.7e+10	
V _{B4}	-15.57	0.72	2.8e+12	-12.74	0.83	2e+11	-11.6	0.86	4.3e+10	-13.51	0.89	2.4e+10	
Rand. Sys.	V _{HT}	—	1.12	4.2e+12	—	2.09	5e+11	—	3.8	1.9e+11	—	4.99	1.4e+11
	V _{SYG}	—	1	3.8e+12	—	1	2.4e+11	—	1	5.1e+10	—	1	2.8e+10
	V _{MC1}	0.19	1.16	4.4e+12	-0.84	2.19	5.2e+11	-0.27	4	2e+11	-0.32	5.24	1.5e+11
	V _{MC2}	0.01	1	3.8e+12	-0.02	1	2.4e+11	-0.03	1	5.1e+10	0.15	1	2.8e+10
	V _{D1}	0.47	1.01	3.8e+12	0.03	1	2.4e+11	-0.07	0.97	4.9e+10	0.05	0.94	2.6e+10
	V _{D2}	0.57	1.01	3.8e+12	0.21	1	2.4e+11	0.24	0.97	4.9e+10	0.45	0.95	2.7e+10
	V _{D3}	0.67	1.02	3.8e+12	0.51	1.01	2.4e+11	0.84	0.99	5e+10	1.27	0.98	2.7e+10
	V _{Fp}	0.4	1.01	3.8e+12	-0.3	1.01	2.4e+11	-0.3	0.99	5e+10	-0.08	0.97	2.7e+10
	V _{Be}	0.09	1	3.8e+12	-0.15	0.99	2.4e+11	-0.12	0.98	5e+10	0.1	0.97	2.7e+10
	V _{T1}	0.02	1	3.8e+12	-0.23	1	2.4e+11	-0.14	1	5e+10	0.15	1	2.8e+10
	V _{MT1}	1.41	1.03	3.9e+12	0.84	1.01	2.4e+11	0.52	0.97	4.9e+10	0.52	0.94	2.6e+10
	V _{MT2}	0.7	1.02	3.8e+12	0.52	1.01	2.4e+11	0.85	0.99	5e+10	1.27	0.98	2.7e+10
	V _{MT3}	0.6	1.01	3.8e+12	0.23	1	2.4e+11	0.25	0.97	4.9e+10	0.46	0.95	2.7e+10
	V _{MT4}	-0.94	0.98	3.7e+12	-0.92	0.98	2.3e+11	-0.92	0.95	4.8e+10	-0.77	0.93	2.6e+10
	V _{MT5}	0.6	1.01	3.8e+12	0.04	0.99	2.4e+11	-0.27	0.95	4.8e+10	-0.26	0.93	2.6e+10
	V _{B1}	0.57	1.01	3.8e+12	0.32	1	2.4e+11	0.53	0.98	5e+10	0.87	0.97	2.7e+10
	V _{B2}	0.8	1.02	3.8e+12	0.7	1	2.4e+11	1.02	0.98	5e+10	1.41	0.96	2.7e+10
V _{B3}	0.34	1.01	3.8e+12	-0.06	1	2.4e+11	0.04	0.99	5e+10	0.33	0.99	2.7e+10	
V _{B4}	-15.82	0.72	2.7e+12	-11.81	0.82	2e+11	-11.69	0.84	4.3e+10	-11.75	0.85	2.4e+10	

Table B.58: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f . Sukhatme population ($N = 34$), $CV(Y) = 0.5$, $CV(X) = 1.1$

sampling	estimator	f=5%			f=10%			f=15%			f=20%		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	VHT	—	0.94	2e+20	—	1.03	2.1e+19	—	1.04	5.5e+18	—	1.13	3.8e+18
	VSYG	—	1	2.1e+20	—	1	2.1e+19	—	1	5.2e+18	—	1	3.4e+18
	VMC ₁	-0.4	0.92	1.9e+20	3.59	1.15	2.4e+19	-0.64	1.01	5.3e+18	0.36	1.12	3.8e+18
	VMC ₂	-6.51	0.78	1.6e+20	4.66	1.22	2.5e+19	-0.88	1.03	5.4e+18	-3.09	1.03	3.4e+18
	VD ₁	-7.22	0.81	1.7e+20	-8.34	0.85	1.8e+19	-9.91	0.84	4.4e+18	-11.32	0.87	2.9e+18
	VD ₂	-6.56	0.82	1.7e+20	-7.02	0.88	1.8e+19	-7.25	0.89	4.7e+18	-7.51	0.95	3.2e+18
	VD ₃	-5.89	0.83	1.7e+20	-4.89	0.92	1.9e+19	-3.2	0.96	5e+18	-1.65	1.07	3.6e+18
	VFP	-7.71	0.8	1.7e+20	0.27	1.03	2.1e+19	0.05	1.04	5.4e+18	0.41	1.13	3.8e+18
	VBe	-0.96	0.93	1.9e+20	-2.24	0.98	2e+19	-2.61	1	5.2e+18	-6.1	0.97	3.3e+18
	VTi	-0.56	0.94	1.9e+20	-0.97	1	2e+19	-0.38	1.03	5.4e+18	-1.81	1.06	3.5e+18
	VMT ₁	-6.56	0.82	1.7e+20	-9.22	0.83	1.7e+19	-11.37	0.81	4.2e+18	-12.83	0.84	2.8e+18
	VMT ₂	-5.85	0.83	1.7e+20	-4.94	0.92	1.9e+19	-3.39	0.96	5e+18	-2	1.06	3.6e+18
	VMT ₃	-6.52	0.82	1.7e+20	-7.06	0.88	1.8e+19	-7.41	0.89	4.6e+18	-7.83	0.94	3.2e+18
	VMT ₄	-13.12	0.71	1.5e+20	-12.31	0.78	1.6e+19	-13.43	0.77	4.1e+18	-14.7	0.81	2.7e+18
	VMT ₅	-10.13	0.76	1.6e+20	-12.66	0.77	1.6e+19	-14.67	0.75	3.9e+18	-16.04	0.78	2.6e+18
VB ₁	-6.56	0.82	1.7e+20	-6.25	0.89	1.8e+19	-5.99	0.91	4.8e+18	-5.71	0.98	3.3e+18	
VB ₂	-9.43	0.77	1.6e+20	-10.76	0.81	1.7e+19	-11.1	0.81	4.2e+18	-11	0.87	2.9e+18	
VB ₃	-3.69	0.87	1.8e+20	-1.74	0.98	2e+19	-0.87	1.01	5.3e+18	-0.43	1.1	3.7e+18	
VB ₄	-26.28	0.52	1.1e+20	-17.38	0.7	1.4e+19	-15.4	0.74	3.9e+18	-14.78	0.81	2.7e+18	
Tillé	VHT	—	0.84	1.8e+20	—	1.11	2e+19	—	0.94	7.5e+18	—	1.09	3.7e+18
	VSYG	—	1	2.2e+20	—	1	1.8e+19	—	1	8e+18	—	1	3.4e+18
	VMC ₁	-0.69	0.85	1.9e+20	2.55	1.16	2e+19	0.3	0.95	7.6e+18	1.91	1.11	3.8e+18
	VMC ₂	-1.3	0.85	1.8e+20	-3.58	1.03	1.8e+19	1.02	1.05	8.3e+18	-5.75	1.01	3.4e+18
	VD ₁	-6.04	0.74	1.6e+20	-6.86	0.92	1.6e+19	-8.49	0.76	6.1e+18	-9.83	0.83	2.8e+18
	VD ₂	-5.51	0.74	1.6e+20	-5.49	0.95	1.7e+19	-5.81	0.8	6.4e+18	-5.86	0.91	3.1e+18
	VD ₃	-4.97	0.75	1.6e+20	-3.37	0.99	1.8e+19	-1.79	0.87	6.9e+18	0.29	1.03	3.5e+18
	VFP	-6.44	0.73	1.6e+20	1.91	1.11	2e+19	1.73	0.94	7.5e+18	2.46	1.09	3.7e+18
	VBe	1.14	0.87	1.9e+20	-0.91	1.06	1.9e+19	-1.79	0.89	7.1e+18	-5.72	0.91	3.1e+18
	VTi	1.59	0.88	1.9e+20	0.42	1.08	1.9e+19	0.72	0.92	7.4e+18	-0.84	1	3.4e+18
	VMT ₁	-5.77	0.74	1.6e+20	-7.76	0.9	1.6e+19	-9.98	0.73	5.8e+18	-11.34	0.81	2.7e+18
	VMT ₂	-4.79	0.76	1.6e+20	-3.45	0.99	1.7e+19	-1.97	0.87	6.9e+18	-0.17	1.02	3.5e+18
	VMT ₃	-5.33	0.75	1.6e+20	-5.56	0.95	1.7e+19	-5.96	0.8	6.4e+18	-6.28	0.9	3.1e+18
	VMT ₄	-12.02	0.65	1.4e+20	-10.9	0.85	1.5e+19	-12.08	0.7	5.6e+18	-13.27	0.77	2.6e+18
	VMT ₅	-9.38	0.68	1.5e+20	-11.25	0.84	1.5e+19	-13.34	0.68	5.4e+18	-14.62	0.75	2.5e+18
VB ₁	-5.51	0.74	1.6e+20	-4.78	0.96	1.7e+19	-4.6	0.82	6.6e+18	-3.95	0.94	3.2e+18	
VB ₂	-8.79	0.69	1.5e+20	-9.35	0.87	1.5e+19	-9.82	0.73	5.8e+18	-9.33	0.84	2.9e+18	
VB ₃	-2.23	0.8	1.7e+20	-0.2	1.06	1.9e+19	0.61	0.92	7.3e+18	1.43	1.06	3.6e+18	
VB ₄	-24.48	0.49	1.1e+20	-16.1	0.75	1.3e+19	-14.09	0.67	5.3e+18	-13.2	0.78	2.6e+18	
CPS	VHT	—	1.04	1.6e+20	—	1	2.6e+19	—	0.95	5.4e+18	—	0.96	4.3e+18
	VSYG	—	1	1.5e+20	—	1	2.6e+19	—	1	5.7e+18	—	1	4.5e+18
	VMC ₁	-0.98	1.04	1.6e+20	1.06	1.07	2.8e+19	0.94	0.97	5.6e+18	3.16	1.04	4.6e+18
	VMC ₂	-0.41	1.11	1.7e+20	2.91	1.17	3e+19	1.36	1	5.7e+18	3.69	1.08	4.8e+18
	VD ₁	-8.46	0.88	1.3e+20	-9.39	0.83	2.2e+19	-9.56	0.77	4.4e+18	-10.61	0.76	3.4e+18
	VD ₂	-7.76	0.89	1.3e+20	-8.06	0.86	2.2e+19	-7	0.82	4.7e+18	-7.04	0.82	3.7e+18
	VD ₃	-7.05	0.91	1.4e+20	-6.04	0.9	2.3e+19	-3.15	0.88	5.1e+18	-1.74	0.91	4.1e+18
	VFP	-8.98	0.87	1.3e+20	-1.18	1	2.6e+19	0.08	0.95	5.4e+18	0.36	0.96	4.3e+18
	VBe	-2.87	0.99	1.5e+20	-2.62	0.98	2.5e+19	-1.29	0.95	5.5e+18	-2.45	0.93	4.2e+18
	VTi	-2.43	1	1.5e+20	-1.57	1	2.6e+19	0.58	0.97	5.6e+18	0.44	0.96	4.3e+18
	VMT ₁	-7.53	0.9	1.3e+20	-10.34	0.81	2.1e+19	-11.04	0.75	4.3e+18	-12.18	0.73	3.3e+18
	VMT ₂	-7.06	0.9	1.4e+20	-6.11	0.89	2.3e+19	-3.23	0.88	5e+18	-1.81	0.91	4.1e+18
	VMT ₃	-7.77	0.89	1.3e+20	-8.12	0.86	2.2e+19	-7.06	0.82	4.7e+18	-7.09	0.82	3.7e+18
	VMT ₄	-14.29	0.77	1.2e+20	-13.31	0.76	2e+19	-13.1	0.72	4.1e+18	-14.02	0.7	3.2e+18
	VMT ₅	-11.07	0.83	1.2e+20	-13.73	0.75	1.9e+19	-14.36	0.69	4e+18	-15.42	0.68	3e+18
VB ₁	-7.76	0.89	1.3e+20	-7.4	0.87	2.2e+19	-5.83	0.83	4.8e+18	-5.53	0.84	3.8e+18	
VB ₂	-10.35	0.84	1.3e+20	-11.9	0.79	2e+19	-10.97	0.74	4.3e+18	-10.86	0.75	3.3e+18	
VB ₃	-5.17	0.94	1.4e+20	-2.9	0.96	2.5e+19	-0.69	0.93	5.3e+18	-0.2	0.94	4.2e+18	
VB ₄	-27.87	0.55	8.2e+19	-18.3	0.68	1.8e+19	-15.21	0.68	3.9e+18	-14.51	0.7	3.1e+18	
Rand. Sys.	VHT	—	0.96	2.7e+20	—	0.95	1.9e+19	—	1.07	5.9e+18	—	0.97	4e+18
	VSYG	—	1	2.8e+20	—	1	2e+19	—	1	5.5e+18	—	1	4.1e+18
	VMC ₁	-1.25	0.93	2.6e+20	0.85	0.98	1.9e+19	1.3	1.13	6.3e+18	0.46	0.98	4e+18
	VMC ₂	-6.77	0.83	2.3e+20	4.18	1.09	2.1e+19	-1.89	1.03	5.7e+18	1.28	1	4.1e+18
	VD ₁	-6.55	0.85	2.4e+20	-7.41	0.81	1.6e+19	-11.5	0.84	4.7e+18	-10.8	0.78	3.2e+18
	VD ₂	-6.03	0.86	2.4e+20	-6.29	0.83	1.6e+19	-8.8	0.9	5e+18	-7.47	0.83	3.4e+18
	VD ₃	-5.49	0.87	2.4e+20	-4.58	0.86	1.7e+19	-4.53	0.98	5.5e+18	-2.58	0.91	3.8e+18
	VFP	-6.94	0.84	2.4e+20	0.09	0.95	1.9e+19	-1.22	1.07	5.9e+18	-0.48	0.97	4e+18
	VBe	0.66	1	2.8e+20	1.2	1	2e+19	-5.97	0.94	5.2e+18	-0.58	0.99	4.1e+18
	VTi	1.1	1	2.8e+20	2.02	1.01	2e+19	-3.09	1	5.5e+18	1.36	1.01	4.1e+18
	VMT ₁	-6.31	0.85	2.4e+20	-8.46	0.79	1.6e+19	-12.89	0.82	4.5e+18	-12.39	0.75	3.1e+18
	VMT ₂	-5.31	0.87	2.5e+20	-4.42	0.87	1.7e+19	-4.81	0.98	5.4e+18	-2.4	0.92	3.8e+18
	VMT ₃	-5.84	0.86	2.4e+20	-6.12	0.84	1.6e+19	-9.05	0.89	4.9e+18	-7.29	0.84	3.4e+18
	VMT ₄	-12.5	0.74	2.1e+20	-11.42	0.75	1.5e+19	-14.97	0.78	4.3e+18	-14.2	0.78	2.9e+18
	VMT ₅	-9.9	0.79	2.2e+20	-11.93	0.73	1.4e+19	-16.13	0.76	4.2e+18	-15.62	0.69	2.8e+18
VB ₁	-6.03	0.86	2.4e+20	-5.73	0.84	1.7e+19	-7.37	0.92	5.1e+18	-6.1	0.85	3.5e+18	
VB ₂	-9.32	0.8	2.2e+20	-10.4	0.76	1.5e+19	-12.39	0.82	4.6e+18	-11.41	0.75	3.1e+18	
VB ₃	-2.73	0.92	2.6e+20	-1.07	0.93	1.8e+19	-2.36	1.03	5.7e+18	-0.79	0.95	3.9e+18	
VB ₄	-24.82	0											

Table B.59: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f . Sukhatme population ($N = 34$), $CV(Y) = 0.5$, $CV(X) = 0.96$

sampling	estimator	f=5%			f=10%			f=15%			f=20%		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	VHT	—	0.94	4.2e+16	—	0.98	4.4e+15	—	1	1.5e+15	—	1.01	9.1e+14
	VSYG	—	1	4.4e+16	—	1	4.5e+15	—	1	1.5e+15	—	1	9e+14
	VMC1	0.08	0.98	4.4e+16	-1.19	0.96	4.3e+15	-0.55	0.99	1.5e+15	0.4	1.01	9.1e+14
	VMC2	-0.09	0.96	4.3e+16	-1.08	0.96	4.3e+15	-0.21	1.01	1.6e+15	0.62	1.04	9.4e+14
	VD1	-5.01	0.84	3.7e+16	-6.16	0.86	3.8e+15	-6.98	0.85	1.3e+15	-7.61	0.84	7.5e+14
	VD2	-4.51	0.85	3.8e+16	-5.3	0.87	3.9e+15	-5.45	0.88	1.4e+15	-5.58	0.88	7.9e+14
	VD3	-4.02	0.86	3.8e+16	-3.63	0.9	4e+15	-1.73	0.93	1.4e+15	-0.2	0.95	8.6e+14
	VFP	-5.37	0.84	3.7e+16	-0.62	0.98	4.4e+15	-0.32	0.99	1.5e+15	-0.29	1	8.9e+14
	VBe	-0.55	0.94	4.2e+16	-1.27	0.97	4.3e+15	-1.39	0.98	1.5e+15	-1.73	0.99	8.8e+14
	Vt1	-0.31	0.94	4.2e+16	-0.29	0.99	4.4e+15	1.07	1.01	1.6e+15	2	1.03	9.2e+14
	VMT1	-4.28	0.85	3.8e+16	-6.61	0.84	3.7e+15	-7.95	0.82	1.3e+15	-8.72	0.81	7.3e+14
	VMT2	-4.06	0.86	3.8e+16	-3.71	0.9	4e+15	-1.82	0.93	1.4e+15	-0.3	0.95	8.5e+14
	VMT3	-4.55	0.85	3.8e+16	-5.38	0.87	3.9e+15	-5.53	0.87	1.3e+15	-5.65	0.87	7.8e+14
	VMT4	-9.79	0.76	3.4e+16	-9.37	0.8	3.6e+15	-9.85	0.8	1.2e+15	-10.38	0.79	7.1e+14
	VMT5	-7.1	0.8	3.5e+16	-9.34	0.79	3.5e+15	-10.61	0.78	1.2e+15	-11.34	0.77	6.9e+14
	VB1	-4.51	0.85	3.8e+16	-4.52	0.89	3.9e+15	-3.33	0.9	1.4e+15	-2.37	0.91	8.2e+14
VB2	-6.59	0.81	3.6e+16	-7.76	0.81	3.6e+15	-7.18	0.82	1.3e+15	-6.42	0.83	7.4e+14	
VB3	-2.44	0.9	4e+16	-1.28	0.96	4.3e+15	0.52	0.99	1.5e+15	1.68	1	9e+14	
VB4	-23.37	0.56	2.5e+16	-16.05	0.71	3.2e+15	-13.57	0.75	1.2e+15	-12.51	0.76	6.9e+14	
Tillé	VHT	—	1.05	4.2e+16	—	1.04	5.2e+15	—	1.07	1.4e+15	—	1.03	9.4e+14
	VSYG	—	1	4e+16	—	1	5e+15	—	1	1.3e+15	—	1	9.1e+14
	VMC1	-0.76	1.03	4.2e+16	1.12	1.06	5.2e+15	1.04	1.05	1.3e+15	1.65	1.04	9.5e+14
	VMC2	-1.3	0.97	3.9e+16	0.7	1.01	5e+15	-0.03	0.99	1.3e+15	0.68	1.01	9.2e+14
	VD1	-2.35	0.91	3.7e+16	0.63	0.92	4.6e+15	2.35	0.93	1.2e+15	2.24	0.9	8.2e+14
	VD2	-1.85	0.92	3.7e+16	1.57	0.94	4.6e+15	4.08	0.97	1.2e+15	4.55	0.95	8.6e+14
	VD3	-1.34	0.93	3.7e+16	3.37	0.97	4.8e+15	8.34	1.03	1.3e+15	10.53	1.03	9.4e+14
	VFP	-2.72	0.91	3.6e+16	6.93	1.06	5.3e+15	10.02	1.11	1.4e+15	10.77	1.09	9.9e+14
	VBe	2.22	1.02	4.1e+16	5.52	1.03	5.1e+15	7.21	1.04	1.3e+15	7.47	1.03	9.3e+14
	Vt1	2.48	1.03	4.1e+16	6.61	1.05	5.2e+15	10.29	1.09	1.4e+15	11.95	1.08	9.8e+14
	VMT1	-1.6	0.92	3.7e+16	0.11	0.9	4.5e+15	1.31	0.9	1.1e+15	0.99	0.87	7.9e+14
	VMT2	-1.38	0.93	3.7e+16	3.27	0.96	4.8e+15	8.19	1.03	1.3e+15	10.37	1.03	9.4e+14
	VMT3	-1.88	0.92	3.7e+16	1.48	0.93	4.6e+15	3.94	0.96	1.2e+15	4.41	0.94	8.6e+14
	VMT4	-7.27	0.82	3.3e+16	-2.8	0.86	4.3e+15	-0.81	0.88	1.1e+15	-0.83	0.85	7.7e+14
	VMT5	-4.49	0.87	3.5e+16	-2.81	0.85	4.2e+15	-1.62	0.85	1.1e+15	-1.91	0.82	7.5e+14
	VB1	-1.85	0.92	3.7e+16	2.41	0.95	4.7e+15	6.53	1	1.3e+15	8.08	0.99	9e+14
VB2	-3.98	0.87	3.5e+16	-1.12	0.87	4.3e+15	2.31	0.91	1.2e+15	3.56	0.89	8.1e+14	
VB3	0.28	0.97	3.9e+16	5.95	1.03	5.1e+15	10.74	1.09	1.4e+15	12.6	1.08	9.9e+14	
VB4	-21.24	0.61	2.5e+16	-9.83	0.76	3.8e+15	-4.83	0.83	1e+15	-3.04	0.82	7.5e+14	
CPS	VHT	—	0.99	4.1e+16	—	0.99	5.2e+15	—	0.98	1.6e+15	—	0.99	8.9e+14
	VSYG	—	1	4.1e+16	—	1	5.3e+15	—	1	1.6e+15	—	1	8.9e+14
	VMC1	-0.04	1	4.1e+16	-0.23	0.98	5.2e+15	-0.37	0.99	1.6e+15	0.26	0.99	8.9e+14
	VMC2	-0.14	0.99	4.1e+16	-0.26	1	5.3e+15	-0.59	1.01	1.6e+15	0.08	1	8.9e+14
	VD1	-4.42	0.88	3.6e+16	-5.76	0.86	4.5e+15	-7.34	0.84	1.3e+15	-7.27	0.83	7.5e+14
	VD2	-3.95	0.89	3.7e+16	-4.9	0.88	4.6e+15	-5.8	0.87	1.4e+15	-5.19	0.87	7.8e+14
	VD3	-3.47	0.9	3.7e+16	-3.24	0.91	4.8e+15	-2.05	0.93	1.5e+15	0.26	0.95	8.5e+14
	VFP	-4.78	0.88	3.6e+16	-0.11	0.99	5.2e+15	-0.7	0.98	1.6e+15	0.16	1	8.9e+14
	VBe	0.23	0.99	4.1e+16	-0.73	0.99	5.2e+15	-1.86	0.97	1.5e+15	-1.71	0.96	8.6e+14
	Vt1	0.47	1	4.1e+16	0.23	1	5.3e+15	0.62	1	1.6e+15	2.18	1.01	9e+14
	VMT1	-3.77	0.89	3.6e+16	-6.25	0.84	4.4e+15	-8.31	0.81	1.3e+15	-8.38	0.81	7.2e+14
	VMT2	-3.49	0.9	3.7e+16	-3.31	0.91	4.8e+15	-2.16	0.92	1.5e+15	0.13	0.95	8.5e+14
	VMT3	-3.97	0.89	3.7e+16	-4.97	0.88	4.6e+15	-5.9	0.87	1.4e+15	-5.3	0.87	7.8e+14
	VMT4	-9.24	0.8	3.3e+16	-8.98	0.81	4.2e+15	-10.2	0.79	1.3e+15	-10.05	0.79	7e+14
	VMT5	-6.6	0.84	3.4e+16	-8.99	0.8	4.2e+15	-10.97	0.77	1.2e+15	-11.01	0.76	6.8e+14
	VB1	-3.95	0.89	3.7e+16	-4.12	0.89	4.7e+15	-3.67	0.89	1.4e+15	-1.95	0.91	8.1e+14
VB2	-6.11	0.84	3.5e+16	-7.43	0.82	4.3e+15	-7.5	0.81	1.3e+15	-6.01	0.83	7.4e+14	
VB3	-1.79	0.94	3.9e+16	-0.81	0.97	5.1e+15	0.16	0.98	1.6e+15	2.11	1	8.9e+14	
VB4	-22.73	0.6	2.5e+16	-15.59	0.72	3.8e+15	-13.89	0.74	1.2e+15	-12.15	0.76	6.8e+14	
Rand. Sys.	VHT	—	1.03	4.7e+16	—	0.98	5.2e+15	—	0.99	1.4e+15	—	0.99	9.1e+14
	VSYG	—	1	4.6e+16	—	1	5.3e+15	—	1	1.4e+15	—	1	9.3e+14
	VMC1	-0.08	1.01	4.6e+16	-1.13	0.96	5.1e+15	-0.89	0.98	1.4e+15	-0.49	0.99	9.2e+14
	VMC2	-0.49	0.98	4.5e+16	-0.96	0.97	5.2e+15	-0.87	1	1.4e+15	-0.44	1	9.3e+14
	VD1	-4.89	0.9	4.1e+16	-6.62	0.86	4.6e+15	-8.88	0.83	1.2e+15	-10.57	0.82	7.6e+14
	VD2	-4.39	0.91	4.1e+16	-5.78	0.88	4.7e+15	-7.39	0.86	1.2e+15	-8.61	0.85	7.9e+14
	VD3	-3.88	0.92	4.2e+16	-4.16	0.9	4.8e+15	-3.7	0.92	1.3e+15	-3.46	0.93	8.6e+14
	VFP	-5.26	0.89	4e+16	-1.11	0.98	5.2e+15	-2.46	0.98	1.4e+15	-3.53	0.97	9e+14
	VBe	-0.46	0.99	4.5e+16	-1.47	0.99	5.3e+15	-3.3	0.97	1.4e+15	-4.54	0.96	8.9e+14
	Vt1	-0.22	1	4.5e+16	-0.55	1	5.3e+15	-0.88	1	1.4e+15	-1.06	1	9.3e+14
	VMT1	-4.14	0.9	4.1e+16	-7.12	0.84	4.5e+15	-9.81	0.81	1.1e+15	-11.64	0.79	7.3e+14
	VMT2	-3.93	0.91	4.2e+16	-4.22	0.9	4.8e+15	-3.78	0.92	1.3e+15	-3.54	0.93	8.6e+14
	VMT3	-4.43	0.9	4.1e+16	-5.84	0.87	4.7e+15	-7.46	0.86	1.2e+15	-8.67	0.85	7.9e+14
	VMT4	-9.68	0.81	3.7e+16	-9.81	0.8	4.3e+15	-11.69	0.78	1.1e+15	-13.25	0.77	7.1e+14
	VMT5	-6.96	0.85	3.9e+16	-9.83	0.79	4.2e+15	-12.41	0.76	1.1e+15	-14.18	0.74	6.9e+14
	VB1	-4.39	0.91	4.1e+16	-5.02	0.89	4.7e+15	-5.26	0.89	1.2e+15	-5.54	0.89	8.2e+14
VB2	-6.45	0.86	3.9e+16	-8.3	0.81	4.3e+15	-9	0.81	1.1e+15	-9.46	0.81	7.5e+14	
VB3	-2.32	0.95	4.3e+16	-1.74	0.96	5.1e+15	-1.52	0.97	1.4e+15	-1.62	0.98	9e+14	
VB4	-23.3	0.6	2.7e+16	-16.36	0.71	3.8e+15	-15.39	0.74	1e+15	-15.34	0.74	6.9e+14	

Table B.60: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f . Sukhatme population ($N = 34$), $CV(Y) = 0.5$, $CV(X) = 0.77$

sampling	estimator	f=5%			f=10%			f=15%			f=20%		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	VHT	—	0.99	3.7e+16	—	0.98	6e+15	—	1.01	1.6e+15	—	0.99	9.7e+14
	VS _{YG}	—	1	3.7e+16	—	1	6.1e+15	—	1	1.6e+15	—	1	9.8e+14
	VM _{C1}	1.05	0.98	3.6e+16	1.52	1.01	6.2e+15	0.9	1.03	1.7e+15	-0.25	0.99	9.7e+14
	VM _{C2}	1.26	1	3.7e+16	1.56	1.02	6.2e+15	0.42	1.02	1.6e+15	0.35	1.01	9.9e+14
	VD ₁	-2.46	0.88	3.3e+16	-3.63	0.88	5.4e+15	-4.65	0.88	1.4e+15	-5.35	0.87	8.5e+14
	VD ₂	-2.09	0.89	3.3e+16	-3.08	0.89	5.4e+15	-3.73	0.9	1.5e+15	-4.22	0.89	8.7e+14
	VD ₃	-1.71	0.89	3.3e+16	-2.03	0.91	5.5e+15	-1.45	0.93	1.5e+15	-1.12	0.93	9.1e+14
	VFP	-2.74	0.87	3.2e+16	0.53	0.98	6e+15	0.11	1	1.6e+15	-0.36	0.99	9.7e+14
	VBe	0.25	0.95	3.5e+16	-0.01	0.97	5.9e+15	-0.79	0.98	1.6e+15	-0.67	0.99	9.7e+14
	VT ₁	0.32	0.96	3.5e+16	0.48	0.98	6e+15	0.51	1	1.6e+15	1.09	1	9.8e+14
	VM _{T1}	-1.92	0.88	3.3e+16	-4.38	0.85	5.2e+15	-5.82	0.85	1.4e+15	-6.67	0.83	8.2e+14
	VM _{T2}	-1.79	0.89	3.3e+16	-2.1	0.9	5.5e+15	-1.55	0.93	1.5e+15	-1.19	0.93	9.1e+14
	VM _{T3}	-2.17	0.88	3.3e+16	-3.15	0.89	5.4e+15	-3.83	0.9	1.5e+15	-4.29	0.89	8.7e+14
	VM _{T4}	-5.64	0.82	3.1e+16	-5.75	0.84	5.1e+15	-6.54	0.85	1.4e+15	-7.18	0.83	8.2e+14
	VM _{T5}	-3.76	0.85	3.1e+16	-6.16	0.82	5e+15	-7.56	0.82	1.3e+15	-8.39	0.8	7.9e+14
	VB ₁	-2.09	0.89	3.3e+16	-2.59	0.89	5.5e+15	-2.4	0.92	1.5e+15	-2.29	0.91	8.9e+14
VB ₂	-3.35	0.85	3.2e+16	-5.05	0.83	5.1e+15	-5.19	0.85	1.4e+15	-5.2	0.84	8.2e+14	
VB ₃	-0.82	0.92	3.4e+16	-0.12	0.96	5.9e+15	0.39	0.99	1.6e+15	0.61	0.98	9.6e+14	
VB ₄	-19	0.63	2.3e+16	-12.55	0.75	4.6e+15	-11.47	0.79	1.3e+15	-11.24	0.79	7.7e+14	
Tillé	VHT	—	1.07	4.4e+16	—	0.99	5e+15	—	1.07	1.6e+15	—	1.06	9e+14
	VS _{YG}	—	1	4.1e+16	—	1	5e+15	—	1	1.5e+15	—	1	8.5e+14
	VM _{C1}	0.26	1.07	4.4e+16	-0.14	0.99	5e+15	0.91	1.07	1.6e+15	0.9	1.06	9e+14
	VM _{C2}	-2.84	0.98	4e+16	-1.48	0.97	4.9e+15	-2.45	1	1.5e+15	-2.14	1	8.5e+14
	VD ₁	-2	0.95	3.9e+16	-1.1	0.9	4.5e+15	-0.31	0.95	1.4e+15	0.02	0.93	7.9e+14
	VD ₂	-1.65	0.96	4e+16	-0.55	0.91	4.6e+15	0.65	0.96	1.4e+15	1.24	0.96	8.1e+14
	VD ₃	-1.3	0.97	4e+16	0.53	0.93	4.7e+15	3.09	1	1.5e+15	4.65	1	8.5e+14
	VFP	-2.26	0.95	3.9e+16	2.94	1	5e+15	4.73	1.08	1.6e+15	5.42	1.07	9.1e+14
	VBe	1.06	1.04	4.3e+16	2.63	1	5e+15	3.35	1.04	1.5e+15	3.94	1.04	8.8e+14
	VT ₁	1.16	1.04	4.3e+16	3.11	1.01	5.1e+15	4.83	1.06	1.5e+15	6.16	1.06	9e+14
	VM _{T1}	-1.62	0.95	3.9e+16	-1.78	0.87	4.4e+15	-1.49	0.91	1.3e+15	-1.33	0.9	7.6e+14
	VM _{T2}	-1.35	0.97	4e+16	0.48	0.92	4.6e+15	2.99	1	1.5e+15	4.55	1	8.5e+14
	VM _{T3}	-1.7	0.96	3.9e+16	-0.61	0.91	4.6e+15	0.56	0.96	1.4e+15	1.14	0.95	8.1e+14
	VM _{T4}	-5.19	0.89	3.7e+16	-3.27	0.86	4.3e+15	-2.28	0.91	1.3e+15	-1.91	0.9	7.6e+14
	VM _{T5}	-3.46	0.91	3.8e+16	-3.61	0.84	4.2e+15	-3.31	0.88	1.3e+15	-3.14	0.86	7.3e+14
	VB ₁	-1.65	0.96	4e+16	-0.02	0.91	4.6e+15	2.11	0.98	1.4e+15	3.39	0.98	8.3e+14
VB ₂	-3.08	0.92	3.8e+16	-2.45	0.85	4.3e+15	-0.78	0.91	1.3e+15	0.37	0.9	7.6e+14	
VB ₃	-0.22	1	4.1e+16	2.4	0.98	4.9e+15	4.99	1.06	1.5e+15	6.41	1.06	9e+14	
VB ₄	-18.25	0.69	2.8e+16	-10.45	0.77	3.9e+15	-7.46	0.85	1.2e+15	-6.21	0.85	7.2e+14	
CPS	VHT	—	0.98	4.9e+16	—	1.01	4.7e+15	—	0.99	1.5e+15	—	0.99	1.1e+15
	VS _{YG}	—	1	5.1e+16	—	1	4.6e+15	—	1	1.6e+15	—	1	1.1e+15
	VM _{C1}	-0.53	0.99	5e+16	1.09	1.02	4.7e+15	-0.96	0.95	1.5e+15	0	0.99	1.1e+15
	VM _{C2}	-0.76	0.99	5e+16	0.27	1	4.7e+15	-0.85	0.96	1.5e+15	-0.11	1	1.1e+15
	VD ₁	-3.97	0.9	4.5e+16	-3.23	0.9	4.2e+15	-4.45	0.87	1.4e+15	-5.47	0.86	9.3e+14
	VD ₂	-3.67	0.9	4.6e+16	-2.67	0.91	4.2e+15	-3.56	0.89	1.4e+15	-4.33	0.89	9.5e+14
	VD ₃	-3.38	0.91	4.6e+16	-1.57	0.93	4.3e+15	-1.3	0.92	1.4e+15	-1.24	0.93	9.9e+14
	VFP	-4.18	0.89	4.5e+16	0.8	1.01	4.7e+15	0.17	0.99	1.5e+15	-0.32	0.99	1.1e+15
	VBe	-0.27	1	5e+16	0.06	0.99	4.6e+15	-0.21	0.98	1.5e+15	-0.97	0.98	1e+15
	VT ₁	-0.17	1	5.1e+16	0.57	1	4.6e+15	1.01	1	1.6e+15	0.82	1	1.1e+15
	VM _{T1}	-3.92	0.88	4.5e+16	-3.83	0.87	4.1e+15	-5.6	0.84	1.3e+15	-6.82	0.83	8.9e+14
	VM _{T2}	-3.38	0.91	4.6e+16	-1.64	0.93	4.3e+15	-1.37	0.92	1.4e+15	-1.33	0.93	9.9e+14
	VM _{T3}	-3.67	0.9	4.6e+16	-2.74	0.91	4.2e+15	-3.63	0.89	1.4e+15	-4.41	0.88	9.5e+14
	VM _{T4}	-7.09	0.84	4.2e+16	-5.35	0.86	4e+15	-6.35	0.84	1.3e+15	-7.29	0.83	8.9e+14
	VM _{T5}	-5.72	0.85	4.3e+16	-5.62	0.84	3.9e+15	-7.35	0.81	1.3e+15	-8.53	0.8	8.6e+14
	VB ₁	-3.67	0.9	4.6e+16	-2.13	0.92	4.3e+15	-2.22	0.9	1.4e+15	-2.42	0.91	9.7e+14
VB ₂	-5.39	0.86	4.3e+16	-4.44	0.86	4e+15	-4.99	0.84	1.3e+15	-5.38	0.83	8.9e+14	
VB ₃	-1.95	0.95	4.8e+16	0.18	0.98	4.6e+15	0.55	0.98	1.5e+15	0.54	0.98	1.1e+15	
VB ₄	-19.17	0.66	3.4e+16	-12.46	0.77	3.6e+15	-11.36	0.78	1.2e+15	-11.23	0.79	8.4e+14	
Rand. Sys.	VHT	—	1.03	3.4e+16	—	1.01	5.3e+15	—	0.98	1.6e+15	—	1	1e+15
	VS _{YG}	—	1	3.3e+16	—	1	5.3e+15	—	1	1.7e+15	—	1	1e+15
	VM _{C1}	2.63	1.08	3.6e+16	0.07	0.99	5.2e+15	-1.27	0.97	1.6e+15	0.51	1	1e+15
	VM _{C2}	0.54	1.01	3.4e+16	-0.27	0.98	5.2e+15	-0.83	0.99	1.7e+15	0.31	1	1e+15
	VD ₁	-2.57	0.91	3e+16	-3.78	0.9	4.7e+15	-6.11	0.87	1.5e+15	-6.22	0.87	9e+14
	VD ₂	-2.2	0.92	3.1e+16	-3.22	0.91	4.8e+15	-5.26	0.88	1.5e+15	-5.09	0.89	9.2e+14
	VD ₃	-1.83	0.93	3.1e+16	-2.16	0.93	4.9e+15	-3.12	0.91	1.5e+15	-2.03	0.93	9.6e+14
	VFP	-2.84	0.91	3e+16	0.27	1.01	5.3e+15	-1.63	0.98	1.6e+15	-1.16	0.99	1e+15
	VBe	0.2	1	3.3e+16	-0.28	0.99	5.2e+15	-1.43	0.99	1.7e+15	-1.73	0.98	1e+15
	VT ₁	0.26	1	3.3e+16	0.21	1	5.3e+15	-0.36	1	1.7e+15	0.04	1	1e+15
	VM _{T1}	-2.06	0.91	3e+16	-4.46	0.87	4.6e+15	-7.28	0.83	1.4e+15	-7.54	0.83	8.6e+14
	VM _{T2}	-1.91	0.93	3.1e+16	-2.23	0.92	4.9e+15	-3.17	0.91	1.5e+15	-2.11	0.93	9.6e+14
	VM _{T3}	-2.27	0.92	3.1e+16	-3.29	0.91	4.8e+15	-5.3	0.88	1.5e+15	-5.16	0.89	9.2e+14
	VM _{T4}	-5.74	0.86	2.8e+16	-5.89	0.86	4.5e+15	-7.97	0.83	1.4e+15	-8.02	0.84	8.6e+14
	VM _{T5}	-3.89	0.88	2.9e+16	-6.23	0.84	4.4e+15	-9	0.8	1.3e+15	-9.24	0.8	8.3e+14
	VB ₁	-2.2	0.92	3.1e+16	-2.72	0.92	4.8e+15	-4	0.9	1.5e+15	-3.2	0.91	9.4e+14
VB ₂	-3.49	0.88	2.9e+16	-5.11	0.85	4.5e+15	-6.77	0.83	1.4e+15	-6.11	0.84	8.7e+14	
VB ₃	-0.92	0.96	3.2e+16	-0.34	0.98	5.2e+15	-1.22	0.97	1.6e+15	-0.28	0.98	1e+15	
VB ₄	-19.04	0.66	2.2e+16	-12.82	0.77	4e+15	-12.86	0.78	1.3e+15	-11.98	0.79	8.2e+14	

Table B.61: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f . Sukhatme population ($N = 34$), $CV(Y) = 0.5$, $CV(X) = 0.67$

sampling	estimator	f=5%			f=10%			f=15%			f=20%		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	V _{HT}	—	1.49	5e+13	—	1.72	5.6e+12	—	2.13	1.8e+12	—	2.32	1.2e+12
	V _{SYG}	—	1	3.4e+13	—	1	3.2e+12	—	1	8.4e+11	—	1	5.2e+11
	V _{MC1}	0.51	1.34	4.5e+13	0.2	1.57	5.1e+12	-0.07	2.1	1.8e+12	-0.53	2.34	1.2e+12
	V _{MC2}	0.62	1.01	3.4e+13	0.01	0.97	3.1e+12	0.66	1	8.4e+11	-0.16	0.97	5e+11
	V _{D1}	0.5	0.96	3.2e+13	0.12	0.94	3e+12	-0.3	0.91	7.6e+11	-1.11	0.89	4.6e+11
	V _{D2}	0.77	0.97	3.2e+13	0.53	0.94	3.1e+12	0.36	0.92	7.7e+11	-0.29	0.91	4.7e+11
	V _{D3}	1.04	0.97	3.3e+13	1.53	0.97	3.1e+12	2.82	0.96	8.1e+11	3.34	0.96	5e+11
	V _{FP}	0.3	0.95	3.2e+13	0.37	1	3.2e+12	0.26	0.98	8.2e+11	-0.28	0.97	5e+11
	V _{Be}	0.53	0.99	3.3e+13	0.52	0.99	3.2e+12	0.59	0.98	8.2e+11	0.03	0.97	5e+11
	V _{T1}	0.42	0.99	3.3e+13	0.65	1	3.2e+12	1.59	1.01	8.5e+11	1.86	1.01	5.2e+11
	V _{MT1}	1.85	0.97	3.3e+13	1.07	0.93	3e+12	0.11	0.89	7.5e+11	-1	0.87	4.5e+11
	V _{MT2}	1	0.97	3.3e+13	1.49	0.96	3.1e+12	2.79	0.96	8.1e+11	3.31	0.96	5e+11
	V _{MT3}	0.73	0.96	3.2e+13	0.49	0.94	3.1e+12	0.32	0.92	7.7e+11	-0.32	0.9	4.7e+11
	V _{MT4}	-1.99	0.91	3.1e+13	-1.55	0.9	2.9e+12	-1.79	0.88	7.4e+11	-2.55	0.86	4.5e+11
	V _{MT5}	0.41	0.94	3.2e+13	-0.35	0.9	2.9e+12	-1.29	0.86	7.3e+11	-2.38	0.84	4.3e+11
	V _{B1}	0.77	0.97	3.2e+13	1.11	0.96	3.1e+12	2.15	0.95	8e+11	2.49	0.95	4.9e+11
V _{B2}	0.81	0.95	3.2e+13	1.24	0.92	3e+12	2.31	0.91	7.6e+11	2.65	0.9	4.6e+11	
V _{B3}	0.72	0.98	3.3e+13	0.99	0.99	3.2e+12	1.99	0.99	8.4e+11	2.33	1	5.1e+11	
V _{B4}	-17.87	0.68	2.3e+13	-12.77	0.78	2.5e+12	-11.99	0.8	6.7e+11	-12.12	0.8	4.1e+11	
Tillé	V _{HT}	—	2.76	8.9e+13	—	3.1	9.1e+12	—	4.51	3.4e+12	—	5.09	2.2e+12
	V _{SYG}	—	1	3.2e+13	—	1	2.9e+12	—	1	7.6e+11	—	1	4.4e+11
	V _{MC1}	-0.85	2.91	9.4e+13	1.86	3.03	8.9e+12	-2.33	4.55	3.4e+12	-0.28	5.1	2.3e+12
	V _{MC2}	-0.05	0.98	3.2e+13	0.25	0.97	2.8e+12	-0.15	1	7.6e+11	-0.14	0.96	4.3e+11
	V _{D1}	2.37	1.02	3.3e+13	5.19	1.06	3.1e+12	5.86	1.03	7.8e+11	7.53	1.04	4.6e+11
	V _{D2}	2.65	1.03	3.3e+13	5.63	1.07	3.1e+12	6.56	1.05	7.9e+11	8.42	1.06	4.7e+11
	V _{D3}	2.93	1.04	3.4e+13	6.7	1.09	3.2e+12	9.21	1.1	8.3e+11	12.4	1.13	5e+11
	V _{FP}	2.17	1.02	3.3e+13	5.51	1.13	3.3e+12	6.49	1.12	8.4e+11	8.5	1.14	5e+11
	V _{Be}	2.4	1.06	3.4e+13	5.53	1.11	3.3e+12	6.73	1.1	8.3e+11	8.64	1.12	5e+11
	V _{T1}	2.29	1.06	3.4e+13	5.68	1.12	3.3e+12	7.86	1.14	8.6e+11	10.72	1.17	5.2e+11
	V _{MT1}	3.76	1.04	3.4e+13	6.19	1.05	3.1e+12	6.28	1.01	7.6e+11	7.63	1.01	4.5e+11
	V _{MT2}	2.89	1.04	3.3e+13	6.65	1.09	3.2e+12	9.16	1.1	8.3e+11	12.35	1.13	5e+11
	V _{MT3}	2.61	1.03	3.3e+13	5.59	1.07	3.1e+12	6.52	1.05	7.9e+11	8.38	1.06	4.7e+11
	V _{MT4}	-0.16	0.97	3.1e+13	3.44	1.02	3e+12	4.27	1	7.6e+11	5.96	1.01	4.5e+11
	V _{MT5}	2.29	1.01	3.3e+13	4.7	1.02	3e+12	4.79	0.98	7.4e+11	6.14	0.98	4.4e+11
	V _{B1}	2.65	1.03	3.3e+13	6.26	1.08	3.2e+12	8.49	1.08	8.2e+11	11.47	1.11	4.9e+11
V _{B2}	2.7	1.02	3.3e+13	6.39	1.05	3.1e+12	8.64	1.03	7.8e+11	11.6	1.05	4.7e+11	
V _{B3}	2.6	1.05	3.4e+13	6.13	1.12	3.3e+12	8.33	1.13	8.6e+11	11.33	1.17	5.2e+11	
V _{B4}	-16.34	0.72	2.3e+13	-8.33	0.88	2.6e+12	-6.49	0.9	6.8e+11	-4.35	0.93	4.1e+11	
CPS	V _{HT}	—	1.27	4.3e+13	—	1.56	5.3e+12	—	2.15	1.8e+12	—	2.44	1.2e+12
	V _{SYG}	—	1	3.4e+13	—	1	3.4e+12	—	1	8.4e+11	—	1	4.9e+11
	V _{MC1}	-0.43	1.31	4.4e+13	-0.47	1.59	5.4e+12	1.09	2.16	1.8e+12	0.28	2.46	1.2e+12
	V _{MC2}	-0.12	1.01	3.4e+13	-0.67	1.02	3.4e+12	0.31	1.01	8.4e+11	0.35	1.02	5e+11
	V _{D1}	-0.19	0.96	3.2e+13	-0.9	0.94	3.2e+12	-0.25	0.93	7.8e+11	-0.76	0.91	4.5e+11
	V _{D2}	0.08	0.97	3.3e+13	-0.5	0.95	3.2e+12	0.4	0.94	7.9e+11	0.06	0.92	4.6e+11
	V _{D3}	0.35	0.98	3.3e+13	0.49	0.97	3.3e+12	2.88	0.98	8.3e+11	3.72	0.98	4.9e+11
	V _{FP}	-0.38	0.96	3.2e+13	-0.62	1.01	3.4e+12	0.34	1	8.4e+11	0.05	0.99	4.9e+11
	V _{Be}	-0.1	1	3.4e+13	-0.45	1	3.4e+12	0.54	0.99	8.3e+11	0.36	0.99	4.9e+11
	V _{T1}	-0.21	1	3.4e+13	-0.33	1.01	3.4e+12	1.54	1.02	8.6e+11	2.19	1.03	5.1e+11
	V _{MT1}	1.14	0.98	3.3e+13	0.01	0.93	3.2e+12	0.15	0.91	7.6e+11	-0.64	0.88	4.4e+11
	V _{MT2}	0.31	0.98	3.3e+13	0.45	0.97	3.3e+12	2.84	0.98	8.2e+11	3.68	0.98	4.9e+11
	V _{MT3}	0.04	0.97	3.2e+13	-0.54	0.95	3.2e+12	0.37	0.94	7.9e+11	0.03	0.92	4.6e+11
	V _{MT4}	-2.66	0.92	3.1e+13	-2.55	0.91	3.1e+12	-1.75	0.9	7.5e+11	-2.21	0.88	4.4e+11
	V _{MT5}	-0.29	0.95	3.2e+13	-1.4	0.91	3.1e+12	-1.25	0.88	7.4e+11	-2.02	0.86	4.2e+11
	V _{B1}	0.08	0.97	3.3e+13	0.08	0.96	3.3e+12	2.2	0.97	8.1e+11	2.86	0.97	4.8e+11
V _{B2}	0.1	0.96	3.2e+13	0.17	0.93	3.1e+12	2.36	0.93	7.8e+11	3.04	0.92	4.5e+11	
V _{B3}	0.06	0.99	3.3e+13	-0.02	1	3.4e+12	2.04	1.02	8.5e+11	2.68	1.02	5e+11	
V _{B4}	-18.37	0.69	2.3e+13	-13.61	0.79	2.7e+12	-11.93	0.82	6.8e+11	-11.86	0.82	4e+11	
Rand. Sys.	V _{HT}	—	1.26	4e+13	—	1.5	5.1e+12	—	1.89	1.6e+12	—	2.06	1.1e+12
	V _{SYG}	—	1	3.2e+13	—	1	3.4e+12	—	1	8.4e+11	—	1	5.3e+11
	V _{MC1}	-0.27	1.31	4.2e+13	-0.45	1.55	5.2e+12	0.39	1.93	1.6e+12	1.16	2.14	1.1e+12
	V _{MC2}	-0.44	1	3.2e+13	0.14	1.01	3.4e+12	-0.47	1	8.3e+11	1	1	5.3e+11
	V _{D1}	-0.54	0.96	3.1e+13	-0.77	0.93	3.1e+12	-2.56	0.9	7.5e+11	-2.04	0.87	4.6e+11
	V _{D2}	-0.27	0.97	3.1e+13	-0.37	0.94	3.2e+12	-1.92	0.91	7.6e+11	-1.24	0.89	4.7e+11
	V _{D3}	-0.01	0.98	3.1e+13	0.62	0.96	3.2e+12	0.48	0.95	8e+11	2.35	0.95	5e+11
	V _{FP}	-0.73	0.96	3e+13	-0.45	0.99	3.3e+12	-2.02	0.97	8.1e+11	-1.21	0.95	5.1e+11
	V _{Be}	-0.52	1	3.2e+13	-0.32	0.98	3.3e+12	-1.69	0.97	8.1e+11	-0.89	0.95	5.1e+11
	V _{T1}	-0.63	1	3.2e+13	-0.19	0.99	3.3e+12	-0.75	0.99	8.3e+11	0.88	0.99	5.3e+11
	V _{MT1}	0.81	0.97	3.1e+13	0.14	0.92	3.1e+12	-2.16	0.88	7.3e+11	-1.94	0.85	4.5e+11
	V _{MT2}	-0.04	0.97	3.1e+13	0.58	0.96	3.2e+12	0.45	0.95	7.9e+11	2.31	0.94	5e+11
	V _{MT3}	-0.31	0.97	3.1e+13	-0.4	0.94	3.2e+12	-1.95	0.91	7.6e+11	-1.26	0.89	4.7e+11
	V _{MT4}	-3	0.92	2.9e+13	-2.42	0.9	3e+12	-4.02	0.87	7.3e+11	-3.47	0.85	4.5e+11
	V _{MT5}	-0.61	0.95	3e+13	-1.27	0.9	3e+12	-3.53	0.85	7.1e+11	-3.3	0.83	4.4e+11
	V _{B1}	-0.27	0.97	3.1e+13	0.21	0.95	3.2e+12	-0.18	0.94	7.8e+11	1.51	0.93	4.9e+11
V _{B2}	-0.22	0.95	3e+13	0.29	0.92	3.1e+12	-0.01	0.89	7.5e+11	1.66	0.88	4.7e+11	
V _{B3}	-0.33	0.98	3.1e+13	0.13	0.99	3.3e+12	-0.35	0.98	8.2e+11	1.35	0.98	5.2e+11	
V _{B4}	-18.74	0.68	2.2e+13	-13.46	0.78	2.6e+12	-14.01	0.79	6.6e+11	-12.96	0.79	4.2e+11	

Table B.62: Monte Carlo Relative Bias, Relative Stability and MSE of different variance estimators under Brewer, Tillé, CPS and Randomised Systematic designs and different sampling fractions f . Sukhatme population ($N = 34$), $CV(Y) = 0.5$, $CV(X) = 0.5$

sampling	estimator	f=5%			f=10%			f=15%			f=20%		
		RB	RS	MSE	RB	RS	MSE	RB	RS	MSE	RB	RS	MSE
Brewer	VHT	—	5.64	1.2e+13	—	13.38	1.8e+12	—	35.24	7.3e+11	—	58.43	5.2e+11
	VSYG	—	1	2.1e+12	—	1	1.4e+11	—	1	2.1e+10	—	1	9e+09
	VMC1	1.06	3.36	7.1e+12	-0.73	10.2	1.4e+12	-3.05	30.29	6.3e+11	-2	52.23	4.7e+11
	VMC2	0.83	1.09	2.3e+12	-0.64	1.07	1.4e+11	-1.36	1.06	2.2e+10	0.27	1.07	9.6e+09
	VD1	3.22	1.18	2.5e+12	1.57	1.17	1.6e+11	0.99	1.19	2.5e+10	2.76	1.21	1.1e+10
	VD2	3.34	1.19	2.5e+12	1.8	1.18	1.6e+11	1.34	1.21	2.5e+10	3.2	1.23	1.1e+10
	VD3	3.46	1.19	2.5e+12	2.16	1.19	1.6e+11	2	1.23	2.5e+10	4.05	1.26	1.1e+10
	VFP	3.12	1.18	2.5e+12	-0.92	1.08	1.5e+11	-1.49	1.1	2.3e+10	0.25	1.11	1e+10
	VB _e	1.04	1.1	2.3e+12	-0.27	1.08	1.5e+11	-0.52	1.1	2.3e+10	1.39	1.12	1e+10
	VT _i	0.73	1.09	2.3e+12	-0.93	1.06	1.4e+11	-1.64	1.05	2.2e+10	-0.05	1.06	9.5e+09
	VMT1	5	1.24	2.6e+12	3.47	1.24	1.7e+11	2.5	1.25	2.6e+10	4.04	1.26	1.1e+10
	VMT2	3.47	1.19	2.5e+12	2.14	1.19	1.6e+11	1.97	1.22	2.5e+10	4.02	1.25	1.1e+10
	VMT3	3.35	1.19	2.5e+12	1.78	1.18	1.6e+11	1.31	1.2	2.5e+10	3.17	1.22	1.1e+10
	VMT4	1.76	1.15	2.4e+12	0.62	1.15	1.6e+11	0.13	1.17	2.4e+10	1.91	1.19	1.1e+10
	VMT5	4.17	1.22	2.6e+12	2.65	1.22	1.7e+11	1.69	1.23	2.6e+10	3.22	1.24	1.1e+10
	Tillé	VB1	3.34	1.19	2.5e+12	1.93	1.18	1.6e+11	1.64	1.21	2.5e+10	3.61	1.24
VB2		4.49	1.23	2.6e+12	3.94	1.26	1.7e+11	4.26	1.33	2.8e+10	6.62	1.38	1.2e+10
VB3		2.19	1.14	2.4e+12	-0.08	1.11	1.5e+11	-0.98	1.11	2.3e+10	0.59	1.1	1e+10
VB4		-15.63	0.77	1.6e+12	-13.3	0.83	1.1e+11	-14.69	0.84	1.7e+10	-14.21	0.83	7.5e+09
VHT		—	3.64	8.6e+12	—	11.17	1.7e+12	—	30.94	7.1e+11	—	50.18	5e+11
VSYG		—	1	2.4e+12	—	1	1.5e+11	—	1	2.3e+10	—	1	1e+10
VMC1		1.18	3.95	9.3e+12	0.14	11.24	1.7e+12	-2.47	31.01	7.1e+11	0.05	50.35	5.1e+11
VMC2		1.75	1	2.4e+12	-0.25	1	1.5e+11	-0.22	1	2.3e+10	-0.76	1	1e+10
VD1		3.64	1.07	2.5e+12	1.57	1.08	1.6e+11	1.47	1.08	2.5e+10	0.62	1.08	1.1e+10
VD2		3.76	1.07	2.5e+12	1.8	1.08	1.6e+11	1.83	1.09	2.5e+10	1.05	1.09	1.1e+10
VD3		3.89	1.07	2.5e+12	2.17	1.09	1.6e+11	2.5	1.11	2.5e+10	1.89	1.12	1.1e+10
VFP		3.54	1.06	2.5e+12	-0.92	0.99	1.5e+11	-1.02	0.99	2.3e+10	-1.84	1	1e+10
VB _e		1.46	0.99	2.3e+12	-0.24	1	1.5e+11	0.04	1	2.3e+10	-0.66	1	1e+10
VT _i		1.15	0.98	2.3e+12	-0.88	0.97	1.5e+11	-1.05	0.96	2.2e+10	-2.02	0.95	9.5e+09
VMT1		5.43	1.12	2.6e+12	3.47	1.14	1.7e+11	2.99	1.13	2.6e+10	1.87	1.12	1.1e+10
VMT2		3.89	1.07	2.5e+12	2.14	1.09	1.6e+11	2.47	1.11	2.5e+10	1.86	1.11	1.1e+10
VMT3	3.77	1.07	2.5e+12	1.78	1.08	1.6e+11	1.8	1.09	2.5e+10	1.02	1.09	1.1e+10	
VMT4	2.18	1.04	2.5e+12	0.62	1.05	1.6e+11	0.61	1.06	2.4e+10	-0.21	1.06	1.1e+10	
VMT5	4.59	1.1	2.6e+12	2.65	1.12	1.7e+11	2.18	1.11	2.5e+10	1.07	1.11	1.1e+10	
VB1	3.76	1.07	2.5e+12	1.94	1.09	1.6e+11	2.14	1.1	2.5e+10	1.45	1.1	1.1e+10	
VB2	4.91	1.11	2.6e+12	3.95	1.16	1.7e+11	4.77	1.2	2.8e+10	4.41	1.22	1.2e+10	
VB3	2.61	1.03	2.4e+12	-0.07	1.01	1.5e+11	-0.49	1	2.3e+10	-1.5	1	1e+10	
VB4	-15.27	0.69	1.6e+12	-13.29	0.76	1.1e+11	-14.26	0.76	1.7e+10	-15.99	0.76	7.7e+09	
CPS	VHT	—	2.84	6.4e+12	—	9.55	1.4e+12	—	28.23	6.2e+11	—	49.47	4.7e+11
	VSYG	—	1	2.3e+12	—	1	1.4e+11	—	1	2.2e+10	—	1	9.6e+09
	VMC1	-0.17	3.04	6.9e+12	-0.51	9.63	1.4e+12	5.1	28.31	6.3e+11	1.44	49.33	4.7e+11
	VMC2	-0.48	1.01	2.3e+12	-0.95	1	1.4e+11	-0.26	1	2.2e+10	-0.13	1	9.6e+09
	VD1	1.61	1.09	2.5e+12	1.57	1.11	1.6e+11	2.54	1.13	2.5e+10	2.47	1.14	1.1e+10
	VD2	1.73	1.09	2.5e+12	1.79	1.11	1.6e+11	2.9	1.14	2.5e+10	2.91	1.15	1.1e+10
	VD3	1.86	1.09	2.5e+12	2.15	1.12	1.6e+11	3.56	1.16	2.6e+10	3.76	1.17	1.1e+10
	VFP	1.52	1.08	2.5e+12	-0.9	1.02	1.5e+11	0.01	1.04	2.3e+10	-0.03	1.04	1e+10
	VB _e	-0.5	1.01	2.3e+12	-0.29	1.02	1.5e+11	0.91	1.04	2.3e+10	1.06	1.04	1e+10
	VT _i	-0.79	1	2.3e+12	-0.95	1	1.4e+11	-0.26	0.99	2.2e+10	-0.39	0.99	9.5e+09
	VMT1	3.35	1.14	2.6e+12	3.46	1.17	1.7e+11	4.08	1.18	2.6e+10	3.75	1.18	1.1e+10
	VMT2	1.86	1.09	2.5e+12	2.13	1.12	1.6e+11	3.54	1.15	2.6e+10	3.73	1.17	1.1e+10
	VMT3	1.74	1.09	2.5e+12	1.77	1.11	1.6e+11	2.87	1.14	2.5e+10	2.88	1.15	1.1e+10
	VMT4	0.18	1.06	2.4e+12	0.61	1.09	1.6e+11	1.67	1.11	2.4e+10	1.62	1.12	1.1e+10
	VMT5	2.53	1.12	2.5e+12	2.64	1.15	1.7e+11	3.27	1.16	2.6e+10	2.93	1.16	1.1e+10
	VB1	1.73	1.09	2.5e+12	1.93	1.12	1.6e+11	3.2	1.15	2.5e+10	3.31	1.16	1.1e+10
VB2	2.85	1.13	2.6e+12	3.93	1.19	1.7e+11	5.89	1.26	2.8e+10	6.33	1.29	1.2e+10	
VB3	0.62	1.05	2.4e+12	-0.07	1.04	1.5e+11	0.51	1.05	2.3e+10	0.3	1.04	1e+10	
VB4	-16.89	0.71	1.6e+12	-13.29	0.78	1.1e+11	-13.43	0.78	1.7e+10	-14.47	0.78	7.5e+09	
Rand. Sys.	VHT	—	2.75	6.3e+12	—	9.11	1.3e+12	—	26.79	6e+11	—	45.85	4.5e+11
	VSYG	—	1	2.3e+12	—	1	1.4e+11	—	1	2.2e+10	—	1	9.8e+09
	VMC1	1.3	3.11	7.2e+12	-2.5	9.56	1.4e+12	-0.88	28.02	6.3e+11	-0.05	47.79	4.7e+11
	VMC2	0.9	0.99	2.3e+12	-0.01	0.99	1.4e+11	-0.4	0.98	2.2e+10	1.37	0.97	9.5e+09
	VD1	3.3	1.08	2.5e+12	2.07	1.1	1.6e+11	2.26	1.12	2.5e+10	4.33	1.13	1.1e+10
	VD2	3.43	1.09	2.5e+12	2.3	1.1	1.6e+11	2.61	1.13	2.5e+10	4.78	1.14	1.1e+10
	VD3	3.55	1.09	2.5e+12	2.66	1.11	1.6e+11	3.28	1.14	2.6e+10	5.64	1.17	1.1e+10
	VFP	3.21	1.08	2.5e+12	-0.4	1.02	1.5e+11	-0.26	1.03	2.3e+10	1.78	1.03	1e+10
	VB _e	1.14	1.01	2.3e+12	0.25	1.02	1.5e+11	0.69	1.03	2.3e+10	2.9	1.03	1e+10
	VT _i	0.84	1	2.3e+12	-0.41	0.99	1.4e+11	-0.46	0.98	2.2e+10	1.41	0.98	9.5e+09
	VMT1	5.09	1.14	2.6e+12	3.97	1.16	1.7e+11	3.79	1.17	2.6e+10	5.63	1.17	1.1e+10
	VMT2	3.56	1.09	2.5e+12	2.64	1.11	1.6e+11	3.25	1.14	2.5e+10	5.61	1.16	1.1e+10
	VMT3	3.43	1.09	2.5e+12	2.28	1.1	1.6e+11	2.59	1.12	2.5e+10	4.75	1.14	1.1e+10
	VMT4	1.85	1.05	2.4e+12	1.11	1.08	1.6e+11	1.39	1.1	2.4e+10	3.47	1.11	1.1e+10
	VMT5	4.25	1.12	2.6e+12	3.14	1.14	1.7e+11	2.98	1.15	2.6e+10	4.8	1.15	1.1e+10
	VB1	3.43	1.09	2.5e+12	2.43	1.11	1.6e+11	2.92	1.13	2.5e+10	5.19	1.15	1.1e+10
VB2	4.57	1.13	2.6e+12	4.43	1.18	1.7e+11	5.59	1.24	2.8e+10	8.26	1.28	1.3e+10	
VB3	2.28	1.05	2.4e+12	0.44	1.04	1.5e+11	0.25	1.03	2.3e+10	2.11	1.03	1e+10	
VB4	-15.53	0.7	1.6e+12	-12.83	0.77	1.1e+11	-13.65	0.77	1.7e+10	-12.92	0.76	7.4e+09	

R PACKAGES

In this appendix, the R packages we developed for performing the analyses presented in this thesis will be briefly described.

C.1 JIPAPPROX

This package provides functions to approximate joint-inclusion probabilities in Unequal Probability Sampling, or to find Monte Carlo approximations of first and second-order inclusion probabilities of a general sampling design. The package is currently available on CRAN¹ (the Comprehensive R Archive Network). The main functions are:

- `jip_approx()`: returns a matrix of approximated joint-inclusion probabilities for Unequal Probability Sampling designs with high entropy;
- `jip_MonteCarlo()`: produces a matrix of first and second-order inclusion probabilities for a given sampling design, approximated through Monte Carlo simulation.
- `HTvar()`: returns the Horvitz-Thompson or Sen-Yates-Grundy variance or their estimates, computed using true inclusion probabilities or an approximation obtained by functions `jip_approx()` or `jip_MonteCarlo()`.

A brief example of the usage of package `jipApprox` is proposed below, for more details, the reader is referred to the package manual².

```
library(UPSvarApprox)

### Generate population data
N <- 20; n <- 5
x <- rgamma(N, scale=10, shape=5)
y <- abs( 2*x + 3.7*sqrt(x) * rnorm(N) )
pik <- n * x/sum(x)

### Approximate joint-inclusion probabilities
### for high entropy designs
pikl <- jip_approx(pik, method='Hajek')
pikl <- jip_approx(pik, method='HartleyRao')
```

¹ <https://CRAN.R-project.org/package=jipApprox>

² <https://cran.r-project.org/web/packages/jipApprox/jipApprox.pdf>

```
### Approximate inclusion probabilities
### through Monte Carlo simulation
pikl <- jip_MonteCarlo(x=pik, n=n,
                      replications=100, design="brewer")
```

C.2 UPSVARAPPROX

The purpose of package `UPSvarApprox` is to provide functions for the approximation of the variance of the Horvitz–Thompson estimator under Unequal Probability Sampling designs using only first-order inclusion probabilities; hence the name, which is an abbreviation for *Unequal Probability Sampling variance approximation*. The package is currently available on CRAN³.

The main function is `approx_var_est()`, which estimates the variance by means of any of the 18 approximate variance estimators defined in section 2.6. The package also implements function `Var_approx()`, which computes an approximation of the population variance of the Horvitz–Thompson estimator. This is clearly not useful in a real context, but could be of interest in simulation studies.

A brief example of the usage of package `UPSvarApprox` is proposed below; for more details, the reader is referred to the package manual⁴.

```
library(UPSvarApprox)

### Generate population data
N <- 500; n <- 50
x <- rgamma(N, scale=10, shape=5)
y <- abs( 2*x + 3.7*sqrt(x) * rnorm(N) )
pik <- n * x/sum(x)
s <- sample(N, n)
ys <- y[s]
piks <- pik[s]

### Variance approximations
VarApprox(y, pik, n, method="Hajek1")
VarApprox(y, pik, n, method="FixedPoint")

### Approximate variance estimators
## Estimators of class 2
approx_var_est(ys, piks, method="Deville1")
approx_var_est(ys, piks, method="FixedPoint")

## Estimators of class 3
approx_var_est(ys, pik, method="Tille", sample=s)
approx_var_est(ys, pik, method="Berger", sample=s)
```

³ <https://CRAN.R-project.org/package=UPSvarApprox>

⁴ <https://cran.r-project.org/web/packages/UPSvarApprox/UPSvarApprox.pdf>

C.3 BOOTSTRAPFP

Package `bootstrapFP` implements the main bootstrap algorithms for finite population inference available in literature, for estimating the variance of the Horvitz–Thompson total estimator.

The package is available on CRAN⁵ and includes pseudo–population and direct bootstrap methods for SRS and UPS designs, as well as the generalised bootstrap by Beaumont and Patak (2012).

An example of its usage follows, more details are available in the help pages of the package⁶.

```
library(bootstrapFP)

### Generate population data
N <- 500; n <- 50
x <- rgamma(N, scale=10, shape=5)
y <- abs( 2*x + 3.7*sqrt(x) * rnorm(N) )
pik <- n * x/sum(x)

### Draw a dummy sample
s <- sample(N, n)

### Estimate bootstrap variance
bootstrapFP(y=y[s], pik=n/N, B=1000, method="ppSitter")
bootstrapFP(y=y[s], pik=pik[s], B=1000, method="ppHolmberg",
            design='brewer')
bootstrapFP(y=y[s], pik=pik[s], B=1000, D=10, method="ppChauvet")
bootstrapFP(y=y[s], pik=n/N, B=1000, method="dRaoWu")
bootstrapFP(y=y[s], pik=n/N, B=1000, method="dSitter")
bootstrapFP(y=y[s], pik=pik[s], B=1000, method="dAntalTille_UPS",
            design='brewer')
bootstrapFP(y=y[s], pik=n/N, B=1000, method="wRaoWuYue")
bootstrapFP(y=y[s], pik=n/N, B=1000, method="wChipperfieldPreston")
bootstrapFP(y=y[s], pik=pik[s], B=1000, method="wGeneralised",
            distribution='normal')
```

⁵ <https://CRAN.R-project.org/package=bootstrapFP>

⁶ <https://cran.r-project.org/web/packages/bootstrapFP/bootstrapFP.pdf>

C.4 OTHER PACKAGES

The following R packages, available on GitHub, were also developed in order to perform some of the analyses presented in the thesis:

- `fpdust`⁷: implements the FPDUST and PPS FPDUST spatial sampling designs used for the application to the roaming–dog survey described in chapter 5;
- `robustHT`⁸: provides functions for the estimation of the conditional bias and for the computation of the Robust Horvitz–Thompson estimator (see chapter 4).

⁷ <https://github.com/rhobis/fpdust>

⁸ <https://github.com/rhobis/robustHT>

REFERENCES

- Aires, Nibia (1999). 'Algorithms to Find Exact Inclusion Probabilities for Conditional Poisson Sampling and Pareto π ps Sampling Designs'. In: *Methodology And Computing In Applied Probability* 1.4, pp. 457–469. DOI: 10.1023/A:1010091628740.
- Amado, Conceição and Ana M. Pires (2004). 'Robust Bootstrap with Non Random Weights Based on the Influence Function'. In: *Communications in Statistics - Simulation and Computation* 33.2, pp. 377–396. DOI: 10.1081/SAC-120037242.
- Antal, Erika and Yves Tillé (2011). 'A Direct Bootstrap Method for Complex Sampling Designs From a Finite Population'. In: *Journal of the American Statistical Association* 106.494, pp. 534–543. DOI: 10.1198/jasa.2011.tm09767.
- Antal, Erika and Yves Tillé (2014). 'A New Resampling Method for Sampling Designs Without Replacement: the Doubled Half Bootstrap'. In: *Computational Statistics* 29.5, pp. 1345–1363. DOI: 10.1007/s00180-014-0495-0.
- Arcos, Antonio et al. (2015). *Frames2: Estimation in Dual Frame Surveys*. R package version 0.2.1. URL: <http://CRAN.R-project.org/package=Frames2>.
- Asok, C. and B. V. Sukhatme (1976). 'On Sampford's Procedure of Unequal Probability Sampling without Replacement'. In: *Journal of the American Statistical Association* 71.356, pp. 912–918. DOI: 10.1080/01621459.1976.10480968.
- Barabesi, L., L. Fattorini, and G. Ridolfi (1997). 'Two-phase surveys of elusive populations'. In: *Proceedings of the Statistic Canada Symposium*. Vol. 97, pp. 285–287.
- Barranco-Chamorro, I. et al. (2007). 'Local influence on the ratio and Horvitz-Thompson estimators through the second order approach'. In: *Applied Mathematics and Computation* 190.1, pp. 851–865. DOI: <https://doi.org/10.1016/j.amc.2007.01.093>.
- Beaumont, Jean-François, David Haziza, and A. Ruiz-Gazen (2013). 'A unified approach to robust estimation in finite population sampling'. In: *Biometrika* 100.3, pp. 555–569. DOI: doi.org/10.1093/biomet/ast010.
- Beaumont, Jean-François and Zdenek Patak (2012). 'On the Generalized Bootstrap for Sample Surveys with Special Attention to Poisson Sampling'. In: *International Statistical Review* 80.1, pp. 127–148. DOI: 10.1111/j.1751-5823.2011.00166.x.
- Berger, Yves G. (1998a). 'Rate of convergence for asymptotic variance of the Horvitz-Thompson estimator'. In: *Journal of Statistical Plan-*

- ning and Inference* 74.1, pp. 149–168. DOI: [https://doi.org/10.1016/S0378-3758\(98\)00107-4](https://doi.org/10.1016/S0378-3758(98)00107-4).
- Berger, Yves G. (1998b). 'Variance estimation using list sequential scheme for unequal probability sampling'. In: *Journal of Official Statistics* 14.3, pp. 315–323. URL: <https://eprints.soton.ac.uk/34115/>.
- Berger, Yves G. (2004). 'A Simple Variance Estimator for Unequal Probability Sampling without Replacement'. In: *Journal of Applied Statistics* 31.3, pp. 305–315. DOI: 10.1080/0266476042000184046.
- Berger, Yves G. (2007). 'A Jackknife Variance Estimator for Unistage Stratified Samples with Unequal Probabilities'. In: *Biometrika* 94.4, pp. 953–964. URL: <http://www.jstor.org/stable/20441428>.
- Berger, Yves G. and J. N. K. Rao (2006). 'Adjusted jackknife for imputation under unequal probability sampling without replacement'. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68.3, pp. 531–547. DOI: 10.1111/j.1467-9868.2006.00555.x.
- Berger, Yves G. and Chris J. Skinner (2005). 'A Jackknife Variance Estimator for Unequal Probability Sampling'. In: *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 67.1, pp. 79–89. URL: <http://www.jstor.org/stable/3647601>.
- Bertail, Patrice and Pierre Combris (1997). 'Bootstrap Généralisé d'un Sondage'. In: *Annales d'Économie et de Statistique* 46, pp. 49–83. URL: <http://www.jstor.org/stable/20076068>.
- Bethlehem, Jelke G. and Maarten H. Schuerhoff (1984). 'Second-Order Inclusion Probabilities in Sequential Sampling Without Replacement with Unequal Probabilities'. In: *Biometrika* Vol. 71, No. 3, pp. 642–644.
- Bickel, P. J. and D. A. Freedman (1984). 'Asymptotic Normality and the Bootstrap in Stratified Sampling'. In: *The Annals of Statistics* 12.2, pp. 470–482. DOI: 10.1214/aos/1176346500.
- Booth, James G., Ronald W. Butler, and Peter Hall (1994). 'Bootstrap Methods for Finite Populations'. In: *Journal of the American Statistical Association* 89.428, pp. 1282–1289. DOI: 10.1080/01621459.1994.10476868.
- Brewer, Kenneth R. W. (1975). 'A Simple Procedure For Sampling π pswor'. In: *Australian Journal of Statistics* 17 (3), pp. 166–172.
- Brewer, Kenneth R. W. (2002). 'Combined Survey Sampling Inference: Weighing of Basu's Elephant'. In:
- Brewer, Kenneth R. W. and Martin E. Donadio (2003). 'The high entropy variance of the Horvitz–Thompson estimator'. In: *Survey Methodology* 29.2, pp. 189–196.
- Brewer, Kenneth R. W. and Muhammad Hanif (1982). *Sampling with unequal probabilities*. Springer-Verlag.
- Chao, Min-Te (1982). 'A general purpose unequal probability sampling plan'. In: *Biometrika* Vol. 69.No. 3, pp. 653–656.

- Chao, Min-Te and Shaw-Hwa Lo (1985). 'A Bootstrap Method for Finite Population'. In: *Sankhyā: The Indian Journal of Statistics, Series A*, pp. 399–405.
- Chao, Min-Te and Shaw-Hwa Lo (1994). 'Maximum Likelihood Summary and the Bootstrap Method in Structured Finite Populations'. In: *Statistica Sinica*, pp. 389–406.
- Chauvet, Guillaume (2007). 'Méthodes de Bootstrap en Population Finie'. PhD thesis. ENSAI.
- Chen, Xiang-Hui, Arthur P. Dempster, and Jun S. Liu (1994). 'Weighted finite population sampling to maximize entropy'. In: *Biometrika* 81.3, pp. 457–469. DOI: 10.1093/biomet/81.3.457.
- Chipperfield, James and John Preston (2007). 'Efficient Bootstrap for Business Surveys'. In: *Survey Methodology* 33.2, pp. 167–172.
- Cochran, W. G. (1977). *Sampling Techniques*. Wiley.
- Cook, R Dennis (1986). 'Assessment of local influence'. In: *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 133–169.
- Deville, Jean-Claude (1993). 'Estimation de la Variance Pour les Enquêtes en Deux Phases'. In: *Manuscript. INSEE, Paris*.
- Deville, Jean-Claude (1999). 'Variance Estimation for Complex Statistics and Estimators: Linearization and Residual Techniques'. In: *Survey methodology* 25.2, pp. 193–204.
- Deville, Jean-Claude (2000). 'Note sur l'algorithme de Chen, Dempster et Liu'. In: *Hand-written note. CREST-ENSAI, Rennes*.
- Deville, Jean-Claude and Yves Tillé (2005). 'Variance approximation under balanced sampling'. In: *Journal of Statistical Planning and Inference* 128.2, pp. 569–591. ISSN: 0378-3758. DOI: <https://doi.org/10.1016/j.jspi.2003.11.011>.
- Efron, B. (1979). 'Bootstrap Methods: Another Look at the Jackknife'. In: *Annals of Statistics* 7, pp. 1–26.
- Estevao, V., M.A. Hidiroglou, and Carl-Erik Särndal (1995). 'Methodological principles for a generalized estimation system at Statistics Canada'. In: *Journal of Official Statistics* 11.2, pp. 181–204.
- Fattorini, L. (2006). 'Applying the Horvitz: Thompson Criterion in Complex Designs: A Computer-Intensive Perspective for Estimating Inclusion Probabilities'. In: *Biometrika* 93.2, pp. 269–278.
- Fattorini, L. and G. Ridolfi (1997). 'A sampling design for areal units based on spatial variability'. In: *Metron* 55.1-2, pp. 59–72.
- Grafström, Anton (2010). 'Entropy of unequal probability sampling designs'. In: *Statistical Methodology* 7.2, pp. 84–97. ISSN: 1572-3127. DOI: <https://doi.org/10.1016/j.stamet.2009.10.005>.
- Grafström, Anton and Niklas L. P. Lundström (2013). 'Why Well Spread Probability Samples Are Balanced'. In: *Open Journal of Statistics* 3.1, pp. 36–41. DOI: <https://doi.org/10.4236/ojs.2013.31005>.
- Grafström, Anton and Yves Tillé (2012). 'Doubly balanced spatial sampling with spreading and restitution of auxiliary totals'. In: *Environmetrics* 24.2, pp. 120–131. DOI: 10.1002/env.2194.

- Gross, S. (1980). 'Median Estimation in Sample Surveys'. In: *Proceedings of the Section on Survey Research Methods*. Vol. 1814184. American Statistical Association.
- Hájek, Jaroslav (1964). 'Asymptotic Theory of Rejective Sampling with Varying Probabilities from a Finite Population'. In: *The Annals of Mathematical Statistics* 35.4, pp. 1491–1523. DOI: 10.1214/aoms/1177700375.
- Hájek, Jaroslav (1981). *Sampling from a finite population*. Dekker.
- Hanurav, T. V. (1967). 'Optimum Utilization of Auxiliary Information: π ps Sampling of Two Units from a Stratum'. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 29.2, pp. 374–391. URL: <http://www.jstor.org/stable/2984597>.
- Hartley, H. O. and J. N. K. Rao (1962). 'Sampling with Unequal Probabilities and without Replacement'. In: *Ann. Math. Statist.* 33.2, pp. 350–374. DOI: 10.1214/aoms/1177704564.
- Haziza, David, Fulvia Mecatti, and J. N. K. Rao (2008). 'Evaluation of some approximate variance estimators under the Rao-Sampford unequal probability sampling design'. In: *Metron - International Journal of Statistics* LXVI.1, pp. 91–108. URL: <https://EconPapers.repec.org/RePEc:mtn:ancoec:080105>.
- Henderson, Tamie (2006). 'Estimating the variance of the Horvitz–Thompson estimator'. In: Bachelor's thesis.
- Holmberg, A (1998). 'A Bootstrap Approach to Probability Proportional-To-Size Sampling'. In: *Proceedings of the Survey Research Methods Section of the American Statistical Association*, pp. 378–383.
- Horvitz, D. G. and D. J. Thompson (1952). 'A generalization of sampling without replacement from a finite universe'. In: *Journal of the American statistical Association* 47.260, pp. 663–685.
- Huber, Peter J. (1964). 'Robust Estimation of a Location Parameter'. In: *The Annals of Mathematical Statistics* 35.1, pp. 73–101.
- Hulliger, Beat (1995). 'Outlier Robust Horvitz-Thompson Estimators'. In: *Survey methodology* 21, pp. 79–87.
- Italian Ministry of Health (2006). *Rendicontazione annuale sul randagismo*.
- Jackman, Jennifer and Andrew N. Rowan (2007). 'Free-roaming dogs in developing countries: The benefits of capture, neuter, and return programs'. In: Washington, DC: Humane Society Press., pp. 55–78.
- Knuth, Donald E (1969). *The Art of Computer Programming. Vol. 2: Seminumerical Algorithms*. AddisonWesley, pp. 229–279.
- Mashreghi, Zeinab, David Haziza, and Christian Léger (2016). 'A Survey of Bootstrap Methods in Finite Population Sampling'. In: *Statistics Surveys* 10, pp. 1–52. DOI: 10.1214/16-SS113.
- Matei, Alina and Yves Tillé (2005). 'Evaluation of variance approximations and estimators in maximum entropy sampling with unequal probability and fixed sample size'. In: *Journal of Official Statistics* 21.4, pp. 543–570.

- Moreno-Rebollo, J. L., A. M. Muñoz-Reyes, and J. M. Muñoz-Pichardo (1999). 'Influence diagnostic in survey sampling: conditional bias'. In: *Biometrika* 86.4, pp. 923–928.
- Quenouille, M. H. (1949). 'Approximate Tests of Correlation in Time-Series'. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 11.1, pp. 68–84. DOI: 10.2307/2983696.
- Rao, J. N. K. (1965). 'On Two Simple Schemes of Unequal Probability Sampling Without Replacement'. In: *Journal of the Indian Statistical Association* 3.4, pp. 173–80.
- Rao, J. N. K. and C. F. J. Wu (1988). 'Resampling Inference with Complex Survey Data'. In: *Journal of the American Statistical Association* 83.401, pp. 231–241. DOI: 10.1080/01621459.1988.10478591.
- Rao, J. N. K., C. F. J. Wu, and K. Yue (1992). 'Some Recent Work on Resampling Methods for Complex Surveys'. In: *Survey methodology* 18.2, pp. 209–217.
- Rosén, B. (1991). 'Variance Estimation for Systematic pps-sampling'. In: *Statistics Sweden*.
- Sampford, M. R. (1967). 'On Sampling Without Replacement With Unequal Probabilities of Selection'. In: *Biometrika* 54.3-4, pp. 499–513.
- Särndal, Carl-Erik (1996). 'Efficient Estimators with Simple Variance in Unequal Probability Sampling'. In: *Journal of the American Statistical Association* 91.435, pp. 1289–1300. DOI: 10.1080/01621459.1996.10476998.
- Särndal, Carl-Erik, Bengt Swensson, and Jan Wretman (1992). *Model assisted survey sampling*. Springer Science & Business Media.
- Sen, A. R. (1953). 'On the estimates in variance in sampling with varying probabilities'. In: *Journal of the Indian Society of Agruculture Statistics* 5, pp. 119–127.
- Shannon, Claude E. (1948). 'A mathematical theory of communication'. In: *The Bell System Technical Journal* 27.3, pp. 379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.
- Sichera, Roberto (2018a). *jipApprox: Approximate Inclusion Probabilities for Survey Sampling*. R package version 0.1.1. URL: <https://CRAN.R-project.org/package=jipApprox>.
- Sichera, Roberto (2018b). *UPSvarApprox: Approximate the variance of the Horvitz-Thompson total estimator*. R package version 0.1.0.
- Sitter, R. R. (1992). 'Comparing Three Bootstrap Methods for Survey Data'. In: *Canadian Journal of Statistics* 20.2, pp. 135–154. DOI: 10.2307/3315464.
- Stehman, Stephen V. and W. Scott Overton (1994). 'Comparison of Variance Estimators of the Horvitz-Thompson Estimator for Randomized Variable Probability Systematic Sampling'. In: *Journal of the American Statistical Association* 89.425, pp. 30–43. URL: <http://www.jstor.org/stable/2291198>.

- Sukhatme, Pandurang V. (1954). *Sampling theory of surveys with applications*. The Indian Society Of Agricultural Statistics; New Delhi.
- Sunter, Alan (1977). 'List Sequential Sampling with Equal or Unequal Probabilities without Replacement'. In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 26.3, pp. 261–268. URL: <http://www.jstor.org/stable/2346966>.
- Sunter, Alan (1986). 'Solutions to the problem of unequal probability sampling without replacement'. In: *International Statistical Review/Revue Internationale de Statistique*, pp. 33–50.
- Thompson, M. E. and Changbao Wu (2008). 'Simulation-based randomized systematic PPS sampling under substitution of units'. In: *Survey Methodology* 34.1, pp. 3–10.
- Tillé, Yves (1996). 'An elimination procedure for unequal probability sampling without replacement'. In: *Biometrika* 83.1, pp. 238–241. DOI: 10.1093/biomet/83.1.238.
- Tillé, Yves (2006). *Sampling algorithms*. Springer.
- Tillé, Yves and Alina Matei (2016). *sampling: Survey Sampling*. R package version 2.8. URL: <https://CRAN.R-project.org/package=sampling>.
- Tukey, J. W. (1958). 'Bias and confidence in not quite large samples'. In: *Annals of Mathematical Statistics* 29, p. 614.
- Wolter, Kirk M. (1985). *Introduction to Variance Estimation*. Springer-Verlag.
- World Organization for Animal Health (2010). *Stray dog population control*.
- World Society for the Protection of Animals (2010). *Surveying roaming dog populations: guidelines on methodology*.
- Wright, R. L. (1983). 'Finite Population Sampling With Equal or Unequal Probabilities Without Replacement'. In: *Journal of the American Statistical Association* 78, pp. 879–884.
- Wright, R. L. (1989). 'Methods and Tools of Load Research'. In: *The MBSS System, Version III*.
- Wu, Xizhi and Zhen Luo (1993). 'Residual sum of squares and multiple potential, diagnostics by a second order local approach'. In: *Statistics & Probability Letters* 16.4, pp. 289–296. DOI: [https://doi.org/10.1016/0167-7152\(93\)90132-3](https://doi.org/10.1016/0167-7152(93)90132-3).
- Yates, F. and P. M. Grundy (1953). 'Selection without replacement from within strata with probability proportional to size'. In: *Journal of the Royal Statistical Society B* 15, pp. 253–261.