



HAL
open science

Asymptotic properties of cross-classified sampling designs

Jean Rubin, Guillaume Chauvet

► **To cite this version:**

Jean Rubin, Guillaume Chauvet. Asymptotic properties of cross-classified sampling designs. 2024. hal-04579261

HAL Id: hal-04579261

<https://hal.science/hal-04579261>

Preprint submitted on 17 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asymptotic properties of cross-classified sampling designs

Jean Rubin^{1,2} and Guillaume Chauvet¹

¹Ensaï (Irmair), Campus de Ker Lann, Bruz - France

²Insee

May 17, 2024

Abstract

We investigate the family of cross-classified sampling designs across an arbitrary number of dimensions. We introduce a variance decomposition that enables the derivation of general asymptotic properties for these designs and the development of straightforward and asymptotically unbiased variance estimators. Additionally, we demonstrate the suitability of weighted bootstrap techniques for CCS, given the availability of a weighted bootstrap technique in each dimension. Our conclusions are supported by an extensive simulation study. Finally, we apply the proposed methods to a French longitudinal survey conducted among children.

1 Introduction

Multistage sampling designs are commonly used in household and health surveys. In case of two-stage sampling, the population units are grouped into large blocks (e.g. municipalities or counties), called Primary Sampling Units (PSUs), which are sampled at the first stage. At the second stage, a list of population units is obtained inside the selected PSUs, and a sample of these units is drawn. A detailed treatment of multistage sampling may be found in Cochran (1977), Särndal et al. (1992) and Fuller (2011). In some situations, a population unit \mathbf{k} is more easily represented as a D -uple vector (k_1, \dots, k_D) . For example, in the ELFE maternity survey (Juillard et al., 2017), a sample is obtained by selecting a sample of maternity units (k_1), a sample of days (unit k_2), and by crossing the two samples. In this case, a population unit is therefore given by a day-maternity couple (k_1, k_2) .

In such situations, Ohlsson (1996) introduced the two-dimensional cross-classified sampling (CCS) design, under which independent samples S_1 and S_2 are selected in each dimension. By taking the Cartesian product of S_1 and S_2 , the final sample $S = S_1 \times S_2$ is naturally obtained. Other examples of cross-classified sampling designs include consumer price index surveys (Dalén and Ohlsson, 1995) and business surveys (Skinner, 2015). To produce reliable estimators with associated confidence intervals, some basic statistical properties are needed for cross-classified sampling designs, including the consistency and the asymptotic normality of Horvitz-Thompson estimators. It is also desirable to provide (at least approximately) unbiased variance estimators. Thereupon, it would also be of both theoretical and practical interest to derive appropriate bootstrap methods for cross-classified sampling. With the notable exception of Skinner (2015), who proposed a bootstrap algorithm for with-replacement sampling designs in each dimension, this last topic has not been studied in the literature.

In this work, we extend cross-classified sampling to an arbitrary number of dimensions. Under some mild conditions, we prove the consistency and the asymptotic normality of the Horvitz-Thompson (HT)-estimator. Using the Hoeffding-Sobol decomposition (Hoeffding, 1948), we generalize the variance formula given in Ohlsson (1996) and prove that the variance of the HT-estimator can be decomposed into a sum of multiple terms with different orders of magnitude. By identifying the leading terms in this variance decomposition, we obtain simple, consistent variance estimators. This decomposition is also used to derive bootstrap methods suitable for cross-classified sampling.

The article is structured as follows. In Section 2, we introduce our notations and outline our main assumptions. Section 3 utilizes the Hoeffding-Sobol decomposition to derive a comprehensive variance decomposition and establish the consistency and asymptotic normality of the HT-estimator. We also present straightforward and consistent variance estimators, which are exemplified in Section 4. Section 5 introduces and examines a weighted bootstrap method, while its application is illustrated in Section 6. Section 7 presents the results of a simulation study. In Section 8, we apply the proposed methods to the ELFE survey. The proofs are provided in the Appendix.

2 Multi-dimensional cross-classified sampling

2.1 Notation

Suppose that we are interested in a finite population $\mathcal{U} = \prod_{d=1}^D \mathcal{U}_d$ which can be seen as the cartesian product of D finite populations of respective sizes N_d , $d = 1, \dots, D$. The size of the product population \mathcal{U} is therefore $N = \prod_{d=1}^D N_d$. For each unit $\mathbf{k} = (k_1, \dots, k_D) \in \mathcal{U}$, a variable of interest y takes the value $y_{\mathbf{k}}$. We are interested in estimating the population total

$$Y = \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}}. \quad (2.1)$$

For computations considered in the following sections, it is convenient to introduce some notations for sub-totals. Let $I \subseteq \{1, \dots, D\}$ denote a subset of dimensions, and let $\mathcal{U}_I = \prod_{d \in I} \mathcal{U}_d$ be the product population associated to these dimensions. For any $\mathbf{k}' \in \mathcal{U}_I$, we let

$$Y_{\mathbf{k}'} = \sum_{\substack{\mathbf{l} \in \mathcal{U} \\ \forall d \in I, l_d = k'_d}} y_{\mathbf{l}} \quad (2.2)$$

denote the sub-total of y when the set of coordinates in \mathcal{U}_I remains fixed and equal to \mathbf{k}' .

For $d \in \{1, \dots, D\}$, we let $p_d(\cdot)$ denote a sampling design used in the population \mathcal{U}_d . Under a D -dimensional cross-classified sampling design, D independent samples S_d are selected in the populations \mathcal{U}_d , $d = 1, \dots, D$, and their cartesian product $S := \prod_{d=1}^D S_d$ is the overall sample. Therefore, the resulting sampling design $p(\cdot)$ is such that

$$\forall d \in \{1, \dots, D\}, \forall s_d \in \mathcal{P}(\mathcal{U}_d), p\left(\prod_{d=1}^D s_d\right) = \prod_{d=1}^D p_d(s_d). \quad (2.3)$$

In the particular case when $D = 2$, we obtain the usual two-dimensional cross-classified design introduced by Ohlsson (1996). It is somewhat similar to a two-stage sampling design where each unit k_1 in \mathcal{U}_1 would be a PSU, while \mathcal{U}_2 would be the sub-population of SSUs for any k_1 . The main difference is that, due to the independence in the selection of S_1 and S_2 , the same subsample of SSUs is used inside any PSU: in other words, the cross-classified design does not verify the so-called *independence property* for two-stage sampling (Särndal et al., 1992, pages 134-135). However, the *invariance property* is satisfied, namely the sampling design used in a given dimension does not depend on the outcome of the sampling performed in another dimension, see Särndal et al. (1992, Section 4.3.1) and Juillard et al. (2017).

In the population \mathcal{U}_d , we let δ_k^d denote the sample membership indicator of the unit k in the sample S_d . We let π_k^d denote the probability that k is selected in S_d , then $n_d = \sum_{k \in \mathcal{U}_d} \pi_k^d$ denote the expected size of the sample S_d , and π_{kl}^d denote the probability that units k and l are jointly selected in S_d . We will assume in the following that these probabilities are all strictly positive. Finally we also use the notation

$$\Delta_{kl}^d = Cov(\delta_k^d, \delta_l^d) = \pi_{kl}^d - \pi_k^d \pi_l^d \text{ for any } k, l \in \mathcal{U}_d. \quad (2.4)$$

2.2 Estimation

We consider weighted total estimators as:

$$\hat{Y} = \sum_{\mathbf{k} \in S} y_{\mathbf{k}} \prod_{d=1}^D w_{k_d}^d(S_d), \quad (2.5)$$

where $\{w_k^d(S_d)\}_{k \in \mathcal{U}_d}$ is a set of estimation weights available for each dimension $d \in \{1, \dots, D\}$. We suppose that $\{w_k^d(S_d)\}_{k \in \mathcal{U}_d}$ depends only on the sample S_d , and not on the samples selected in the other dimensions. When there is no risk of confusion, we simplify the notation as $w_k^d(S_d) \equiv w_k^d$. We also suppose that for any $d \in \{1, \dots, D\}$ and $k \in \mathcal{U}_d$, $w_k^d = 0$ if $k \notin S_d$. This implies that each unit $\mathbf{k} \in \mathcal{U}$ has an associated weight $w_{\mathbf{k}} = \prod_{d=1}^D w_{k_d}^d$ that can be decomposed as a product of independent weights, leading to the following compact expression:

$$\hat{Y} = \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} w_{\mathbf{k}}. \quad (2.6)$$

In the important specific case when $w_k^d = \{\pi_k^d\}^{-1} \delta_k^d$ for $k \in \mathcal{U}_d$, we obtain the Horvitz-Thompson (HT) estimator

$$\hat{Y}_{\pi} = \sum_{\mathbf{k} \in S} \frac{y_{\mathbf{k}}}{\prod_{d=1}^D \pi_{k_d}^d} = \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} \prod_{d=1}^D \frac{\delta_{k_d}^d}{\pi_{k_d}^d}. \quad (2.7)$$

We may further introduce for $\mathbf{k} \in \mathcal{U}$ the quantities $\delta_{\mathbf{k}} = \prod_{d=1}^D \delta_{k_d}^d$ and $\pi_{\mathbf{k}} = \prod_{d=1}^D \pi_{k_d}^d$, leading to the compact expression

$$\hat{Y}_{\pi} = \sum_{\mathbf{k} \in S} \frac{y_{\mathbf{k}}}{\pi_{\mathbf{k}}} = \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} \frac{\delta_{\mathbf{k}}}{\pi_{\mathbf{k}}}. \quad (2.8)$$

We will also use the notation $\pi_{\mathbf{k}\mathbf{l}} = \prod_{d=1}^D \pi_{k_d l_d}^d$ for any $\mathbf{k}, \mathbf{l} \in \mathcal{U}$.

2.3 Assumptions

In this paper, we consider the following assumptions:

H1. There exists some constant α such that

$$\frac{1}{N} \sum_{\mathbf{k} \in \mathcal{U}} (y_{\mathbf{k}})^2 \leq \alpha. \quad (2.9)$$

H2. For any $d = 1, \dots, D$, we have $n_d \rightarrow \infty$ and $N_d \rightarrow \infty$, and there exists some constant $f_d \in [0, 1]$ such that

$$\frac{n_d}{N_d} \rightarrow f_d. \quad (2.10)$$

H3. For any $d = 1, \dots, D$, there exists some constant $\lambda_d > 0$ such that

$$\forall k \in \mathcal{U}_d, \pi_k^d \geq \lambda_d \frac{n_d}{N_d}. \quad (2.11)$$

H4. For every $d = 1, \dots, D$, there exists a constant γ_d such that

$$\forall k \neq l \in \mathcal{U}_d, |\Delta_{k,l}^d| \leq \gamma_d \frac{n_d}{N_d^2}. \quad (2.12)$$

The assumption (H1) pertains to the variable of interest by positing that it possesses a finite moment of order 2. Assumptions (H2)-(H4) are associated with the sampling design. The asymptotic framework is delineated in (H2), wherein it is explicitly assumed that the sample size $n_d \rightarrow \infty$ across all dimensions. This assumption is imperative for achieving the consistency of the Horvitz-Thompson estimator, as described in Section 3.2. Furthermore, (H3) presupposes that within each dimension d , the first-order inclusion probabilities maintain a lower bound of order n_d/N_d . On the other hand, assumption (H4) addresses the second-order inclusion probabilities. The parameter $|\Delta_{k,l}^d|$ can be interpreted as a gauge of interdependence in the selection of units $k, l \in \mathcal{U}_d$ within the sample S_d . These quantities assume a value of 0 if the units are independently selected in S_d , a concept akin to Poisson sampling. Notably, this assumption holds true for simple random sampling and rejective sampling (Hájek, 1964). In summary, assumptions (H1)-(H4) adhere to conventional standards. Specifically, in the realm of two-dimensional cross-classified sampling, they align with assumptions (H1)-(H3) outlined in Juillard et al. (2017).

3 Properties of total estimators

3.1 Hoeffding-Sobol variance decomposition of the CCS

From equation (2.5), we can derive the following general variance formula:

$$V_p(\hat{Y}) = \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} y_{\mathbf{k}} y_{\mathbf{l}} \left[\prod_{d=1}^D E_p(w_{k_d}^d w_{l_d}^d) - \prod_{d=1}^D E_p(w_{k_d}^d) E_p(w_{l_d}^d) \right]. \quad (3.1)$$

The proof is provided in Appendix A.1. However, the unbiased variance estimator derived from equation (3.1) may not be practical in many sampling designs, as illustrated by Ohlsson (1996). To explore the statistical properties of estimators and propose simpler variance estimators, we adopt an alternative variance decomposition approach, extending Ohlsson's approach for the two-dimensional case. We express \hat{Y} as a sum of uncorrelated components, as outlined in Proposition 1, with its proof given in Appendix A.2. This decomposition, inspired by the works of Hoeffding (1948), is commonly known as the Hoeffding-Sobol or functional ANOVA decomposition.

Proposition 1. *We express \hat{Y} as:*

$$\hat{Y} = \sum_{I \subseteq \{1, \dots, D\}} \hat{Y}^I, \quad (3.2)$$

where for any subset $I \subseteq \{1, \dots, D\}$:

$$\hat{Y}^I = \sum_{I' \in \mathcal{P}(I)} (-1)^{|I|-|I'|} E_p \left\{ \hat{Y} | (S_d)_{d \in I'} \right\} \quad (3.3)$$

$$= \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} \prod_{d \in I} \left\{ w_{k_d}^d - E_p(w_{k_d}^d) \right\} \prod_{d \notin I} E_p(w_{k_d}^d). \quad (3.4)$$

Additionally, the components $\{\hat{Y}^I\}_{I \subseteq \{1, \dots, D\}}$ are uncorrelated, and

$$V_p(\hat{Y}) = \sum_{I \subseteq \{1, \dots, D\}} V_p(\hat{Y}^I), \quad (3.5)$$

where for any $I \neq \emptyset$:

$$V_p(\hat{Y}^I) = \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} y_{\mathbf{k}} y_{\mathbf{l}} \prod_{d \in I} \text{Cov}_p(w_{k_d}^d, w_{l_d}^d) \prod_{d \notin I} E_p(w_{k_d}^d) E_p(w_{l_d}^d). \quad (3.6)$$

From Proposition 1, we derive an unbiased variance estimator for \hat{Y} , as summarized in Corollary 1, with its proof provided in Appendix A.3.

Corollary 1. *Assume that the set of weights is adapted to the sampling design, such that for any $d \in \{1, \dots, D\}$:*

$$\forall k \in \mathcal{U}_d, E_p(w_k^d) = 1. \quad (3.7)$$

Additionally, assume that for any $d \in \{1, \dots, D\}$ and for any pair of units $k, l \in \mathcal{U}_d$, $\widehat{\text{Cov}}^d(w_k^d, w_l^d)$ serves as an unbiased estimator of $\text{Cov}_p(w_k^d, w_l^d)$ constructed from S_d , and that $\widehat{\text{Cov}}^d(w_k^d, w_l^d) = 0$ when either k or l is not in S_d . Consequently, $V_p(\hat{Y})$ can be estimated unbiasedly as follows

$$\hat{V}_p(\hat{Y}) = \sum_{I \subseteq \{1, \dots, D\}} \hat{V}_p(\hat{Y}^I), \quad (3.8)$$

where

$$\hat{V}_p(\hat{Y}^I) = \sum_{\mathbf{k}, \mathbf{l} \in S} y_{\mathbf{k}} y_{\mathbf{l}} \prod_{d \in I} \widehat{Cov}^d(w_{k_d}^d, w_{l_d}^d) \prod_{d \notin I} \frac{1}{\pi_{k_d, l_d}^d}. \quad (3.9)$$

3.2 Properties of the Horvitz-Thompson estimator

We now delve into the HT-estimator, a widely used method in practical scenarios. In this case, the variance formula (3.1) simplifies to

$$V_p(\hat{Y}_\pi) = \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} \frac{y_{\mathbf{k}} y_{\mathbf{l}}}{\pi_{\mathbf{k}} \pi_{\mathbf{l}}} (\pi_{\mathbf{k}\mathbf{l}} - \pi_{\mathbf{k}} \pi_{\mathbf{l}}). \quad (3.10)$$

By using Proposition 1 and Corollary 1, we derive an alternative variance decomposition for \hat{Y}_π alongside an unbiased variance estimator, succinctly presented in Corollary 2. The proof is explained in Appendix B.1.

Corollary 2. *The variance of the HT-estimator can be expressed as:*

$$V_p(\hat{Y}_\pi) = \sum_{I \subseteq \{1, \dots, D\}} V_p(\hat{Y}_\pi^I), \quad (3.11)$$

where for any $I \subseteq \{1, \dots, D\}$

$$\hat{Y}_\pi^I = \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} \prod_{d \in I} \left(\frac{\delta_{k_d}^d}{\pi_{k_d}^d} - 1 \right), \quad (3.12)$$

with the convention that $\hat{Y}_\pi^\emptyset = Y$. Moreover, if $I \neq \emptyset$, we have

$$V_p(\hat{Y}_\pi^I) = \sum_{\mathbf{k}', \mathbf{l}' \in \mathcal{U}_I} \frac{Y_{\mathbf{k}'} Y_{\mathbf{l}'}}{\pi_{\mathbf{k}'} \pi_{\mathbf{l}'}} \Delta_{\mathbf{k}'\mathbf{l}'}, \quad (3.13)$$

where $\Delta_{\mathbf{k}'\mathbf{l}'} = \prod_{d \in I} \Delta_{k'_d l'_d}^d$ and $\pi_{\mathbf{k}'} = \prod_{d \in I} \pi_{k'_d}^d$ for any $\mathbf{k}', \mathbf{l}' \in \mathcal{U}_I$. This variance is estimated unbiasedly by

$$\hat{V}_p(\hat{Y}_\pi^I) = \sum_{\mathbf{k}, \mathbf{l} \in S} \frac{y_{\mathbf{k}} y_{\mathbf{l}}}{\pi_{\mathbf{k}\mathbf{l}}} \prod_{d \in I} \frac{\Delta_{k_d, l_d}^d}{\pi_{k_d}^d \pi_{l_d}^d}. \quad (3.14)$$

The Hoeffding-Sobol decomposition allows us to express the HT-estimator as the sum of 2^D uncorrelated terms, as presented in (3.12), leading to the unbiased variance estimator detailed in (3.14). In the case of $D = 2$, this decomposition simplifies to:

$$\hat{Y}_\pi = \hat{Y}_\pi^\emptyset + \hat{Y}_\pi^{\{1\}} + \hat{Y}_\pi^{\{2\}} + \hat{Y}_\pi^{\{1,2\}}, \quad (3.15)$$

with

$$\begin{aligned}
\hat{Y}_\pi^{\{1\}} &= \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} \left(\frac{\delta_{k_1}^{(1)}}{\pi_{k_1}^{(1)}} - 1 \right), \\
\hat{Y}_\pi^{\{2\}} &= \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} \left(\frac{\delta_{k_2}^{(2)}}{\pi_{k_2}^{(2)}} - 1 \right), \\
\hat{Y}_\pi^{\{1,2\}} &= \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} \left(\frac{\delta_{k_1}^{(1)}}{\pi_{k_1}^{(1)}} - 1 \right) \left(\frac{\delta_{k_2}^{(2)}}{\pi_{k_2}^{(2)}} - 1 \right).
\end{aligned} \tag{3.16}$$

This decomposition mirrors the one proposed in Ohlsson (1996), and Corollary 2 extends his theorem 3.1 to any number of dimensions. This case is further explored in Section 4, focusing particularly on variance estimation for specific sampling designs.

It is important to specify the orders of magnitude of the terms in the decomposition (3.11). Proposition 2 addresses this, with the proof provided in Appendix B.2.

Proposition 2. *Under assumptions (H1), (H3), and (H4), for any non-empty subset I of $1, \dots, D$:*

$$V_p(\hat{Y}_\pi^I) = O\left(\frac{N^2}{\prod_{d \in I} n_d}\right). \tag{3.17}$$

Proposition 2 yields significant implications. Firstly, it indicates that the dominant terms in the Hoeffding-Sobol variance decomposition correspond to subsets I of size 1, while other terms are of negligible order under assumption (H2). Thus, it is feasible to obtain an approximately unbiased variance estimator by focusing solely on the D terms in (3.14) associated with singletons. This leads to the simplified variance estimator:

$$\hat{V}^{SIMPL}(\hat{Y}_\pi) = \sum_{\mathbf{k}, l \in S} \frac{y_{\mathbf{k}} y_l}{\pi_{\mathbf{k}l}} \left(\sum_{d=1}^D \frac{\Delta_{k_d, l_d}^d}{\pi_{k_d}^d \pi_{l_d}^d} \right). \tag{3.18}$$

However, this simplified variance estimator neglects the (non-negative) terms of higher order in the Hoeffding-Sobol decomposition, resulting in negative bias if sample sizes n_d are moderately large. Secondly, assumption (H2) must be retained for the HT-estimator to remain consistent. In other words, if the sample size n_d in some dimension d is bounded, the variance of $N^{-1} \hat{Y}_\pi^d$ does not converge to zero, rendering the Horvitz-Thompson estimator inconsistent. Weak consistency and asymptotic normality of the HT-estimator are established in Proposition 3, with the proof provided in Appendix B.3.

Proposition 3. *Suppose that assumptions (H1)-(H4) hold. Then*

$$V_p(N^{-1}\hat{Y}_\pi) = O(n_m^{-1}), \quad (3.19)$$

where $n_m = \min_{d=1,\dots,D} n_d$. Additionally, suppose that:

H5. *There exists some constant $C > 0$ such that*

$$V_p(\hat{Y}_\pi) \geq CN^2 n_m^{-1}. \quad (3.20)$$

H6. *For any $d = 1, \dots, D$:*

$$\frac{\hat{Y}_\pi^{\{d\}}}{\sqrt{V_p(\hat{Y}_\pi^{\{d\}})}} \rightarrow_{\mathcal{L}} \mathcal{N}(0, 1), \quad (3.21)$$

where $\rightarrow_{\mathcal{L}}$ denotes convergence in distribution.

H7. *For any $d = 1, \dots, D$, there exists some constant $\gamma_d \geq 0$ such that*

$$\frac{V_p(\hat{Y}_\pi^{\{d\}})}{V_p(\hat{Y}_\pi)} \rightarrow (\gamma_d)^2. \quad (3.22)$$

Then

$$\frac{\hat{Y}_\pi - Y}{\sqrt{V_p(\hat{Y}_\pi)}} \rightarrow_{\mathcal{L}} \mathcal{N}(0, 1). \quad (3.23)$$

Assumption (H5) posits that the variance of the HT-estimator is non-vanishing, following the typical order of magnitude $O(N^2 n_m^{-1})$. Proposition 3 establishes that if the HT-estimator is asymptotically normally distributed in each dimension, as stated in (H6), then it is also asymptotically normally distributed under CCS. This generalizes Theorem 1 in Juillard et al. (2017) to an arbitrary number of dimensions.

3.3 Plug-in variance estimation

Another viable variance estimator arises from the decomposition in equation (3.11). For any term $V_p(\hat{Y}_\pi^I)$ of the variance decomposition, a plug-in estimator based on the expression in (3.13) is obtained:

$$\hat{V}^{PLUG}(\hat{Y}_\pi^I) = \sum_{\mathbf{k}', \mathbf{l}' \in S_I} \frac{\hat{Y}_{\mathbf{k}'}}{\pi_{\mathbf{k}'}} \frac{\hat{Y}_{\mathbf{l}'}}{\pi_{\mathbf{l}'}} \frac{\Delta_{\mathbf{k}'\mathbf{l}'}}{\pi_{\mathbf{k}'\mathbf{l}'}} \quad (3.24)$$

where each $\mathbf{k}' \in S_I$ in the partial sum $Y_{\mathbf{k}'}$ is replaced with the unbiased estimator given by:

$$\hat{Y}_{\mathbf{k}'} = \sum_{\substack{\mathbf{l} \in S \\ \forall d \in I, l_d = k'_d}} \frac{y_{\mathbf{l}}}{\prod_{d \notin I} \pi_{l_d}}. \quad (3.25)$$

This estimator employs a Horvitz-Thompson-like variance estimator, leading to the plug-in variance estimator:

$$\hat{V}^{PLUG}(\hat{Y}_{\pi}) = \sum_{I \subseteq \{1, \dots, D\}} \hat{V}^{PLUG}(\hat{Y}_{\pi}^I). \quad (3.26)$$

Following the result obtained in Proposition 2, this estimator may be further simplified by restricting the sum in (3.26) to the D terms associated with singletons. This yields the second simplified variance estimator:

$$\hat{V}^{SIMPL2}(\hat{Y}_{\pi}) = \sum_{d=1}^D \sum_{k_d, l_d \in S_d} \frac{\hat{Y}_{k_d} \hat{Y}_{l_d} \Delta_{k_d, l_d}^d}{\pi_{k_d}^d \pi_{l_d}^d \pi_{k_d, l_d}^d}. \quad (3.27)$$

By construction, this approximation considers only first-order variance terms, simplifying higher-order interactions within these variance terms. Proposition 4 reveals that the bias of the plug-in estimators can be expressed in terms of the $V_p(\hat{Y}_{\pi}^I)$, with the proof provided in Appendix B.4.

Proposition 4. *For every non-empty subset I of $\{1, \dots, D\}$, we have*

$$E_p \left[\hat{V}^{PLUG}(\hat{Y}_{\pi}^I) \right] = V_p(\hat{Y}_{\pi}^I) + \sum_{I' \supseteq I} V_p(\hat{Y}_{\pi}^{I'}). \quad (3.28)$$

It follows from Proposition 2 that the second term on the right-hand side of (3.28) is asymptotically negligible under assumptions (H1)-(H4). Thus, $\hat{V}^{PLUG}(\hat{Y}_{\pi}^I)$ is asymptotically unbiased for $V_p(\hat{Y}_{\pi}^I)$. Likewise, $\hat{V}^{PLUG}(\hat{Y}_{\pi})$ is asymptotically unbiased for $V_p(\hat{Y}_{\pi})$. Equation (3.28) indicates that the plug-in variance estimator overestimates in particular the variance contribution associated with singletons, which constitutes the major part of the variance. If the sample sizes n_d are moderately large, the plug-in variance estimator is therefore expected to be conservative.

4 Illustrations of the simplified estimations

In this section, we demonstrate the proposed variance estimators in the two-dimensional scenario. The general formulas for the case $D = 2$ are initially presented in Section 4.1. We delve into the scenario of simple random sampling in each dimension in Section 4.2, and the scenario of Poisson sampling in each dimension in Section 4.3.

4.1 General case

Expanding upon the specific Hoeffding-Sobol decomposition in the case $D = 2$, as derived in equation (3.15), we derive the variance decomposition:

$$V_p\left(\hat{Y}_\pi\right) = V_p\left(\hat{Y}_\pi^{\{1\}}\right) + V_p\left(\hat{Y}_\pi^{\{2\}}\right) + V_p\left(\hat{Y}_\pi^{\{1,2\}}\right), \quad (4.1)$$

where

$$\begin{aligned} V_p\left(\hat{Y}_\pi^{\{1\}}\right) &= \sum_{k_1, l_1 \in \mathcal{U}_1} \frac{Y_{k_1 \bullet} Y_{l_1 \bullet}}{\pi_{k_1}^{(1)} \pi_{l_1}^{(1)}} \Delta_{k_1 l_1}^{(1)} \text{ with } Y_{k_1 \bullet} = \sum_{k_2 \in \mathcal{U}_2} y_{k_1 k_2}, \\ V_p\left(\hat{Y}_\pi^{\{2\}}\right) &= \sum_{k_2, l_2 \in \mathcal{U}_2} \frac{Y_{\bullet k_2} Y_{\bullet l_2}}{\pi_{k_2}^{(2)} \pi_{l_2}^{(2)}} \Delta_{k_2 l_2}^{(2)} \text{ with } Y_{\bullet k_2} = \sum_{k_1 \in \mathcal{U}_1} y_{k_1 k_2}, \\ V_p\left(\hat{Y}_\pi^{\{1,2\}}\right) &= \sum_{k_1, l_1 \in \mathcal{U}_1} \sum_{k_2, l_2 \in \mathcal{U}_2} \frac{y_{k_1 k_2} y_{l_1 l_2}}{\pi_{k_1}^{(1)} \pi_{k_2}^{(2)} \pi_{l_1}^{(1)} \pi_{l_2}^{(2)}} \Delta_{k_1 l_1}^{(1)} \Delta_{k_2 l_2}^{(2)}, \end{aligned} \quad (4.2)$$

as described in Ohlsson (1996, Theorem 3.1). Utilizing equation (3.14), the corresponding unbiased variance estimator is:

$$\hat{V}_p\left(\hat{Y}_\pi\right) = \hat{V}_p\left(\hat{Y}_\pi^{\{1\}}\right) + \hat{V}_p\left(\hat{Y}_\pi^{\{2\}}\right) + \hat{V}_p\left(\hat{Y}_\pi^{\{1,2\}}\right), \quad (4.3)$$

where

$$\begin{aligned} \hat{V}_p\left(\hat{Y}_\pi^{\{1\}}\right) &= \sum_{k_1, l_1 \in \mathcal{S}_1} \frac{\Delta_{k_1 l_1}^{(1)}}{\pi_{k_1 l_1}^{(1)} \pi_{k_1}^{(1)} \pi_{l_1}^{(1)}} \sum_{k_2, l_2 \in \mathcal{S}_2} \frac{y_{k_1 k_2} y_{l_1 l_2}}{\pi_{k_2 l_2}^{(2)}}, \\ \hat{V}_p\left(\hat{Y}_\pi^{\{2\}}\right) &= \sum_{k_2, l_2 \in \mathcal{S}_2} \frac{\Delta_{k_2 l_2}^{(2)}}{\pi_{k_2 l_2}^{(2)} \pi_{k_2}^{(2)} \pi_{l_2}^{(2)}} \sum_{k_1, l_1 \in \mathcal{S}_1} \frac{y_{k_1 k_2} y_{l_1 l_2}}{\pi_{k_1 l_1}^{(1)}}, \\ \hat{V}_p\left(\hat{Y}_\pi^{\{1,2\}}\right) &= \sum_{k_1, l_1 \in \mathcal{S}_1} \sum_{k_2, l_2 \in \mathcal{S}_2} \frac{y_{k_1 k_2} y_{l_1 l_2}}{\pi_{k_1 l_1}^{(1)} \pi_{k_2 l_2}^{(2)}} \frac{\Delta_{k_1 l_1}^{(1)}}{\pi_{k_1}^{(1)} \pi_{l_1}^{(1)}} \frac{\Delta_{k_2 l_2}^{(2)}}{\pi_{k_2}^{(2)} \pi_{l_2}^{(2)}}. \end{aligned} \quad (4.4)$$

The first simplified variance estimator provided in (3.18) can be reformulated as:

$$\hat{V}^{SIMPL}\left(\hat{Y}_\pi\right) = \hat{V}_p\left(\hat{Y}_\pi^{\{1\}}\right) + \hat{V}_p\left(\hat{Y}_\pi^{\{2\}}\right). \quad (4.5)$$

The second simplified variance estimator provided in (3.27) can be reformulated as:

$$\hat{V}^{SIMPL2}\left(\hat{Y}_\pi\right) = \hat{V}^{PLUG}\left(\hat{Y}_\pi^{\{1\}}\right) + \hat{V}^{PLUG}\left(\hat{Y}_\pi^{\{2\}}\right), \quad (4.6)$$

where

$$\begin{aligned} \hat{V}^{PLUG}\left(\hat{Y}_\pi^{\{1\}}\right) &= \sum_{k_1, l_1 \in \mathcal{S}_1} \frac{\hat{Y}_{k_1 \bullet} \hat{Y}_{l_1 \bullet}}{\pi_{k_1}^{(1)} \pi_{l_1}^{(1)} \pi_{k_1 l_1}^{(1)}} \text{ with } \hat{Y}_{k_1 \bullet} = \sum_{k_2 \in \mathcal{S}_2} \frac{y_{k_1 k_2}}{\pi_{k_2}^{(2)}}, \\ \hat{V}^{PLUG}\left(\hat{Y}_\pi^{\{2\}}\right) &= \sum_{k_2, l_2 \in \mathcal{S}_2} \frac{\hat{Y}_{\bullet k_2} \hat{Y}_{\bullet l_2}}{\pi_{k_2}^{(2)} \pi_{l_2}^{(2)} \pi_{k_2 l_2}^{(2)}} \text{ with } \hat{Y}_{\bullet k_2} = \sum_{k_1 \in \mathcal{S}_1} \frac{y_{k_1 k_2}}{\pi_{k_1}^{(1)}}. \end{aligned} \quad (4.7)$$

4.2 Simple random sampling

We examine the scenario where simple random sampling of size n_d is employed in each dimension $d = 1, 2$. The sampling fraction in dimension d is denoted as $f_d = n_d/N_d$. The Horvitz-Thompson estimator can be expressed as:

$$\hat{Y}_\pi = \frac{N_1 N_2}{n_1 n_2} \sum_{k_1 \in S_1} \sum_{k_2 \in S_2} y_{k_1 k_2} \quad (4.8)$$

By employing equation (4.1), the Hoeffding-Sobol variance decomposition becomes

$$V_p(\hat{Y}_\pi) = N_1^2 N_2^2 \left[(1 - f_1) \frac{S_1^2}{n_1} + (1 - f_2) \frac{S_2^2}{n_2} + (1 - f_1)(1 - f_2) \frac{S_{12}^2}{n_1 n_2} \right] \quad (4.9)$$

where

$$\begin{aligned} S_1^2 &= \frac{1}{N_1 - 1} \sum_{k_1 \in \mathcal{U}_1} (\bar{Y}_{k_1 \bullet} - \bar{Y}_{\bullet \bullet})^2 \text{ with } \bar{Y}_{k_1 \bullet} = \frac{1}{N_2} \sum_{k_2 \in \mathcal{U}_2} y_{k_1 k_2}, \\ S_2^2 &= \frac{1}{N_2 - 1} \sum_{k_2 \in \mathcal{U}_2} (\bar{Y}_{\bullet k_2} - \bar{Y}_{\bullet \bullet})^2 \text{ with } \bar{Y}_{\bullet k_2} = \frac{1}{N_1} \sum_{k_1 \in \mathcal{U}_1} y_{k_1 k_2}, \\ S_{12}^2 &= \frac{1}{N_1 - 1} \frac{1}{N_2 - 1} \sum_{k_1 \in \mathcal{U}_1} \sum_{k_2 \in \mathcal{U}_2} (y_{k_1 k_2} - \bar{Y}_{k_1 \bullet} - \bar{Y}_{\bullet k_2} + \bar{Y}_{\bullet \bullet})^2, \end{aligned} \quad (4.10)$$

and where $\bar{Y}_{\bullet \bullet} = N^{-1}Y$ is the population mean. This expression matches that found in Ohlsson (1996, Example 1.1) and Skinner (2015, Theorem 2.1).

Applying equation (4.3), we derive a term-by-term unbiased estimator. As indicated by Skinner (2015), this process involves substituting each term in (4.10) with the following unbiased estimators:

$$\begin{aligned} \hat{S}_1^2 &= \frac{1}{n_1 - 1} \sum_{k_1 \in S_1} (\bar{y}_{k_1 \bullet} - \bar{y}_{\bullet \bullet})^2 - (1 - f_2) \frac{\hat{S}_{12}^2}{n_2}, \\ \hat{S}_2^2 &= \frac{1}{n_2 - 1} \sum_{k_2 \in S_2} (\bar{y}_{\bullet k_2} - \bar{y}_{\bullet \bullet})^2 - (1 - f_1) \frac{\hat{S}_{12}^2}{n_1}, \\ \hat{S}_{12}^2 &= \frac{1}{n_1 - 1} \frac{1}{n_2 - 1} \sum_{k_1 \in S_1} \sum_{k_2 \in S_2} (y_{k_1 k_2} - \bar{y}_{k_1 \bullet} - \bar{y}_{\bullet k_2} + \bar{y}_{\bullet \bullet})^2, \end{aligned} \quad (4.11)$$

where

$$\bar{y}_{k_1 \bullet} = \frac{1}{n_2} \sum_{k_2 \in S_2} y_{k_1 k_2} \text{ and } \bar{y}_{\bullet k_2} = \frac{1}{n_1} \sum_{k_1 \in S_1} y_{k_1 k_2} \quad (4.12)$$

represent the unbiased estimators of the partial means $\bar{Y}_{k_1 \bullet}$ and $\bar{Y}_{\bullet k_2}$, respectively, with $\bar{y}_{\bullet \bullet}$ denoting the sample mean. It is worth noting that both \hat{S}_1^2 and \hat{S}_2^2 incorporate correction terms for unbiasedness, which become negligible with sufficiently large sample sizes.

The first proposed variance estimator $\hat{V}^{SIMPL}(\hat{Y}_\pi)$, provided in (3.18), is obtained by eliminating the estimator of the interaction term in (4.9), resulting in:

$$\hat{V}^{SIMPL}(\hat{Y}_\pi) = N_1^2 N_2^2 \left[(1 - f_1) \frac{\hat{S}_1^2}{n_1} + (1 - f_2) \frac{\hat{S}_2^2}{n_2} \right]. \quad (4.13)$$

It is important to note that this estimator still incorporates correction terms within \hat{S}_1^2 and \hat{S}_2^2 , necessitating the computation of \hat{S}_{12}^2 . On the other hand, the second proposed variance estimator $\hat{V}^{SIMPL2}(\hat{Y}_\pi)$ not only eliminates the interaction term \hat{S}_{12}^2 but also these correction terms, yielding:

$$\hat{V}^{SIMPL2}(\hat{Y}_\pi) = N_1^2 N_2^2 \left[(1 - f_1) \frac{\hat{S}_1^{2,PLUG}}{n_1} + (1 - f_2) \frac{\hat{S}_2^{2,PLUG}}{n_2} \right] \quad (4.14)$$

where

$$\begin{aligned} \hat{S}_1^{2,PLUG} &= \frac{1}{n_1 - 1} \sum_{k_1 \in S_1} (\bar{y}_{k_1 \bullet} - \bar{y}_{\bullet \bullet})^2, \\ \hat{S}_2^{2,PLUG} &= \frac{1}{n_2 - 1} \sum_{k_2 \in S_2} (\bar{y}_{\bullet k_2} - \bar{y}_{\bullet \bullet})^2, \end{aligned} \quad (4.15)$$

represent the plug-in estimators of S_1^2 and S_2^2 , respectively.

4.3 Poisson sampling

We now consider the scenario where Poisson sampling is applied independently in each dimension $d = 1, 2$. The HT-estimator takes the form:

$$\hat{Y}_\pi = \sum_{k_1 \in S_1} \sum_{k_2 \in S_2} \frac{y_{k_1 k_2}}{\pi_{k_1}^{(1)} \pi_{k_2}^{(2)}}. \quad (4.16)$$

By employing equation (4.1), the Hoeffding-Sobol variance decomposition becomes:

$$\begin{aligned} V_p(\hat{Y}_\pi) &= \sum_{k_1 \in \mathcal{U}_1} \left(\frac{1 - \pi_{k_1}^{(1)}}{\pi_{k_1}^{(1)}} \right) Y_{k_1 \bullet}^2 + \sum_{k_2 \in \mathcal{U}_2} \left(\frac{1 - \pi_{k_2}^{(2)}}{\pi_{k_2}^{(2)}} \right) Y_{\bullet k_2}^2 \\ &+ \sum_{k_1 \in \mathcal{U}_1} \sum_{k_2 \in \mathcal{U}_2} \left(\frac{1 - \pi_{k_1}^{(1)}}{\pi_{k_1}^{(1)}} \right) \left(\frac{1 - \pi_{k_2}^{(2)}}{\pi_{k_2}^{(2)}} \right) y_{k_1 k_2}^2, \end{aligned} \quad (4.17)$$

where $Y_{k_1 \bullet}$ and $Y_{\bullet k_2}$ are defined in equation (4.2). A similar variance decomposition is presented in Example 3.1 of Ohlsson (1996).

By utilizing equation (4.3), the variance is estimated without bias by substituting each term in (4.17) with an unbiased counterpart, resulting in the variance estimator:

$$\begin{aligned}
\hat{V}_p(\hat{Y}_\pi) &= \sum_{k_1 \in S_1} \frac{1}{\pi_{k_1}^{(1)}} \left(\frac{1 - \pi_{k_1}^{(1)}}{\pi_{k_1}^{(1)}} \right) \left(\hat{Y}_{k_1 \bullet}^2 - \sum_{k_2 \in S_2} \frac{1}{\pi_{k_2}^{(2)}} \left(\frac{1 - \pi_{k_2}^{(2)}}{\pi_{k_2}^{(2)}} \right) y_{k_1 k_2}^2 \right) \\
&+ \sum_{k_2 \in S_2} \frac{1}{\pi_{k_2}^{(2)}} \left(\frac{1 - \pi_{k_2}^{(2)}}{\pi_{k_2}^{(2)}} \right) \left(\hat{Y}_{\bullet k_2}^2 - \sum_{k_1 \in S_1} \frac{1}{\pi_{k_1}^{(1)}} \left(\frac{1 - \pi_{k_1}^{(1)}}{\pi_{k_1}^{(1)}} \right) y_{k_1 k_2}^2 \right) \\
&+ \sum_{k_1 \in S_1} \sum_{k_2 \in S_2} \left(\frac{1 - \pi_{k_1}^{(1)}}{\pi_{k_1}^{(1)}} \right) \left(\frac{1 - \pi_{k_2}^{(2)}}{\pi_{k_2}^{(2)}} \right) \frac{y_{k_1 k_2}^2}{\pi_{k_1}^{(1)} \pi_{k_2}^{(2)}},
\end{aligned} \tag{4.18}$$

where $\hat{Y}_{k_1 \bullet}$ and $\hat{Y}_{\bullet k_2}$ are defined in equation (4.7). It is worth noting that the two first terms in the right-hand side of (4.18) incorporate correction terms for unbiasedness, which become negligible with sufficiently large sample sizes.

The first proposed variance estimator $\hat{V}^{SIMPL}(\hat{Y}_\pi)$, presented in (3.18), is derived by eliminating the third interaction term in (4.18), yielding:

$$\begin{aligned}
\hat{V}^{SIMPL}(\hat{Y}_\pi) &= \sum_{k_1 \in S_1} \frac{1}{\pi_{k_1}^{(1)}} \left(\frac{1 - \pi_{k_1}^{(1)}}{\pi_{k_1}^{(1)}} \right) \left(\hat{Y}_{k_1 \bullet}^2 - \sum_{k_2 \in S_2} \frac{1}{\pi_{k_2}^{(2)}} \left(\frac{1 - \pi_{k_2}^{(2)}}{\pi_{k_2}^{(2)}} \right) y_{k_1 k_2}^2 \right) \\
&+ \sum_{k_2 \in S_2} \frac{1}{\pi_{k_2}^{(2)}} \left(\frac{1 - \pi_{k_2}^{(2)}}{\pi_{k_2}^{(2)}} \right) \left(\hat{Y}_{\bullet k_2}^2 - \sum_{k_1 \in S_1} \frac{1}{\pi_{k_1}^{(1)}} \left(\frac{1 - \pi_{k_1}^{(1)}}{\pi_{k_1}^{(1)}} \right) y_{k_1 k_2}^2 \right).
\end{aligned} \tag{4.19}$$

The second proposed variance estimator $\hat{V}^{SIMPL2}(\hat{Y}_\pi)$ further removes the correction terms, resulting in:

$$\hat{V}^{SIMPL2}(\hat{Y}_\pi) = \sum_{k_1 \in S_1} \left(\frac{1 - \pi_{k_1}^{(1)}}{\pi_{k_1}^{(1)}} \right) \frac{\hat{Y}_{k_1 \bullet}^2}{\pi_{k_1}^{(1)}} + \sum_{k_2 \in S_2} \left(\frac{1 - \pi_{k_2}^{(2)}}{\pi_{k_2}^{(2)}} \right) \frac{\hat{Y}_{\bullet k_2}^2}{\pi_{k_2}^{(2)}}. \tag{4.20}$$

5 Weighted bootstrap method for CCS

We now introduce a weighted bootstrap method for CSS, assuming that for each dimension $d \in 1, \dots, D$, a weighted bootstrap method tailored to the weights $(w_{k_d}^d)_{k_d \in \mathcal{U}d}$ is available. Specifically, we consider a set of weights $(w_{k_d}^{d*})_{k_d \in S_d}$ for each $d \in 1, \dots, D$ that satisfies the first and second-order moment constraints outlined by Beaumont and Patak (2012):

$$\forall k_d \in S_d, E_*(w_{k_d}^{d*}) = w_{k_d}^d, \tag{5.1}$$

$$\forall k_d, l_d \in S_d, Cov_*(w_{k_d}^{d*}, w_{l_d}^{d*}) = \widehat{Cov}^d(w_{k_d}^d, w_{l_d}^d), \tag{5.2}$$

where $E_*(\cdot)$, $V_*(\cdot)$ and $Cov_*(\cdot, \cdot)$ denote the (bootstrap) expectation, variance and covariance, conditioned on the original sample S .

A natural bootstrap version for CSS of the weighted estimator in (2.5) is:

$$\hat{Y}^* = \sum_{\mathbf{k} \in S} y_{\mathbf{k}} \prod_{d=1}^D w_{k_d}^{d*} \quad (5.3)$$

where the weights are simulated independently in each dimension $d \in 1, \dots, D$. Applying Proposition 1 conditionally on the original sample S , we obtain the Hoeffding-Sobol decomposition:

$$\begin{aligned} \hat{Y}^* &= \sum_{I \subseteq \{1, \dots, D\}} \hat{Y}^{I*}, \\ \text{with } \hat{Y}^{I*} &= \sum_{\mathbf{k} \in S} y_{\mathbf{k}} \prod_{d \in I} (w_{k_d}^{d*} - w_{k_d}^d) \prod_{d \notin I} w_{k_d}^d \text{ for any nonempty } I \subseteq \{1, \dots, D\}. \end{aligned} \quad (5.4)$$

By applying equation (3.6), we obtain:

$$\begin{aligned} V_*(\hat{Y}^{I*}) &= \sum_{\mathbf{k}, \mathbf{l} \in S} y_{\mathbf{k}} y_{\mathbf{l}} \prod_{d \in I} Cov_*(w_{k_d}^{d*}, w_{l_d}^{d*}) \prod_{d \notin I} E_*(w_{k_d}^{d*}) E_*(w_{l_d}^{d*}) \\ &= \sum_{\mathbf{k}, \mathbf{l} \in S} y_{\mathbf{k}} y_{\mathbf{l}} \prod_{d \in I} \widehat{Cov}_p(w_{k_d}^d, w_{l_d}^d) \prod_{d \notin I} w_{k_d}^d w_{l_d}^d. \end{aligned} \quad (5.5)$$

The last line in (5.5) follows from equations (5.1) and (5.3).

In the case of Horvitz-Thompson estimation, for any $k_d, l_d \in S_d$, we have:

$$w_{k_d}^d = \frac{1}{\pi_{k_d}^d} \quad \text{and} \quad \widehat{Cov}^d(w_{k_d}^d, w_{l_d}^d) = \frac{\Delta_{k_d l_d}^d}{\pi_{k_d, l_d}^d} \frac{1}{\pi_{k_d}^d \pi_{l_d}^d}.$$

For the bootstrap version \hat{Y}_π^* of the HT-estimator, equations (5.4) and (5.5) lead to:

$$\hat{Y}_\pi^* = \sum_{I \subseteq \{1, \dots, D\}} \hat{Y}_\pi^{I*} \quad (5.6)$$

and

$$V_*(\hat{Y}_\pi^{I*}) = \sum_{\mathbf{k}, \mathbf{l} \in S} y_{\mathbf{k}} y_{\mathbf{l}} \prod_{d \in I} \frac{\Delta_{k_d l_d}^d}{\pi_{k_d, l_d}^d} \frac{1}{\pi_{k_d}^d \pi_{l_d}^d} \prod_{d \notin I} \frac{1}{\pi_{k_d}^d \pi_{l_d}^d} \quad (5.7)$$

$$= \sum_{\mathbf{k}', \mathbf{l}' \in S_I} \frac{\hat{Y}_{\mathbf{k}'} \hat{Y}_{\mathbf{l}'}}{\pi_{\mathbf{k}'} \pi_{\mathbf{l}'}} \frac{\Delta_{\mathbf{k}' \mathbf{l}'}}{\pi_{\mathbf{k}' \mathbf{l}'}}. \quad (5.8)$$

This corresponds to the plug-in variance estimator for \hat{Y}_π^I given in (3.24). Therefore, $V_*(\hat{Y}_\pi^*)$ equals the plug-in variance estimator $\hat{V}^{PLUG}(\hat{Y}_\pi)$ as defined in (3.26), and is also asymptotically

unbiased by Proposition 4. This is a very important result. It indicates that if we have bootstrap procedures that satisfy conditions (5.1) and (5.3) in each dimension, then under our conditions (H1)-(H4), it suffices to independently apply these bootstrap procedures in each dimension to obtain an asymptotically unbiased bootstrap variance estimator for CCS. In the following section, we illustrate this result for two bootstrap procedures in the case of simple random sampling in each dimension.

6 Case study: simple random sampling in each dimension

In this section, we focus on the particularly significant case of CCS in dimension $D = 2$, where simple random sampling is employed in each dimension. We investigate the application of the pseudo-population bootstrap in Section 6.1, and the utilization of the rescaled bootstrap is examined in Section 6.2.

6.1 Pseudo-Population Bootstrap method

We introduce a pseudo-population bootstrap approach tailored for simple random sampling without replacement in each dimension, drawing inspiration from the method proposed by Gross (1980). For ease of exposition, let us assume that both N_2/n_2 and N_1/n_1 are integers. Once S_1 and S_2 are sampled, the methodology involves constructing a pseudo-population \mathcal{U}_d^* by replicating each unit in S_d for N_d/n_d times for each $d = 1, 2$. Subsequently, simple random sampling without replacement is applied in each pseudo-population to generate the bootstrap sample S_d^* . Consequently, the bootstrap estimator is given by:

$$\hat{Y}_\pi^* = \frac{N_1 N_2}{n_1 n_2} \sum_{k_1 \in S_1^*} \sum_{k_2 \in S_2^*} y_{k_1 k_2}. \quad (6.1)$$

We can easily check that $E_*(\hat{Y}_\pi^*) = \hat{Y}_\pi$. Furthermore, the bootstrap variance is

$$V_*(\hat{Y}_\pi^*) = N_1^2 N_2^2 \left[(1 - f_1) \frac{S_1^{2*}}{n_1} + (1 - f_2) \frac{S_2^{2*}}{n_2} + (1 - f_1)(1 - f_2) \frac{S_{12}^{2*}}{n_1 n_2} \right] \quad (6.2)$$

with

$$\begin{aligned} S_1^{2*} &= \frac{N_1(n_1 - 1)}{n_1(N_1 - 1)} \left[\hat{S}_1^2 + (1 - f_2) \frac{\hat{S}_{12}^2}{n_2} \right], \\ S_2^{2*} &= \frac{N_2(n_2 - 1)}{n_2(N_2 - 1)} \left[\hat{S}_2^2 + (1 - f_1) \frac{\hat{S}_{12}^2}{n_1} \right], \\ S_{12}^{2*} &= \frac{N_1(n_1 - 1)}{n_1(N_1 - 1)} \frac{N_2(n_2 - 1)}{n_2(N_2 - 1)} \hat{S}_{12}^2. \end{aligned} \quad (6.3)$$

Under our asymptotic framework where $n_1, n_2 \rightarrow \infty$, the bootstrap variance asymptotically converges to the unbiased variance estimator of $V_p(\hat{Y}_\pi)$, as illustrated in equation (4.11). However, when $n_1 \rightarrow \infty$ and n_2 is bounded, the component S_1^2 tends to be overestimated, while the components S_2^2 and S_{12}^2 tend to be underestimated. When $n_2 \rightarrow \infty$ and n_1 is bounded, the component S_2^2 tends to be overestimated whereas the components S_1^2 and S_{12}^2 tend to be underestimated.

6.2 Rescaled bootstrap method

We propose a rescaled bootstrap method tailored for simple random sampling without replacement in each dimension, drawing inspiration from the approach proposed in Rao et al. (1992). The Horvitz-Thompson estimator employs weights of the form $w_{k_1 k_2} = w_{k_1}^{(1)} w_{k_2}^{(2)}$ for each unit $(k_1, k_2) \in \mathcal{U}$, where $w_{k_1}^{(1)} = 1/f_1$ and $w_{k_2}^{(2)} = 1/f_2$. The method proposed in Rao et al. (1992) involves constructing bootstrap weights $w_{k_d}^{d*}$ for $d = 1, 2$ by multiplying the original weights by an adjustment factor $a_{k_d}^{d*}$, resulting in $w_{k_d}^{d*} = a_{k_d}^{d*} w_{k_d}^d$. These adjustment factors are determined by sampling multiplicities $(m_{k_d}^*)_{k_d \in S_d}$ from a multinomial distribution $\mathcal{M}\left(n_d^*, \frac{1}{n_d}, \dots, \frac{1}{n_d}\right)$, leading to:

$$a_{k_d}^{d*} = 1 + \sqrt{\frac{n_d^*(1-f_d)}{n_d-1}} \left(\frac{n_d m_{k_d}^{d*}}{n_d^*} - 1 \right) \text{ for any } k_d \in S_d, \quad (6.4)$$

where n_d^* represents the resampling size, commonly chosen as $n_d^* = n_d - 1$.

The rescaled bootstrap for CCS involves constructing bootstrap weights $w_{k_1 k_2}^* = w_{k_1}^{(1*)} w_{k_2}^{(2*)}$ by employing the rescaled bootstrap in each dimension. Performing straightforward algebra yields:

$$V_*(\hat{Y}_\pi^*) = N_1^2 N_2^2 \left[(1-f_1) \frac{S_1^{2*}}{n_1} + (1-f_2) \frac{S_2^{2*}}{n_2} + (1-f_1)(1-f_2) \frac{S_{12}^{2*}}{n_1 n_2} \right], \quad (6.5)$$

with

$$\begin{aligned} S_1^{2*} &= \hat{S}_1^2 + (1-f_2) \frac{\hat{S}_{12}^2}{n_2}, \\ S_2^{2*} &= \hat{S}_2^2 + (1-f_1) \frac{\hat{S}_{12}^2}{n_1}, \\ S_{12}^{2*} &= \hat{S}_{12}^2. \end{aligned} \quad (6.6)$$

While the second-order moment condition is met in each dimension, the method by Rao and Wu does not ensure that the bootstrap variance estimator is unbiased for CCS. However, it does ensure that under our asymptotic framework where $n_1, n_2 \rightarrow \infty$, the bootstrap variance estimator converges asymptotically to the unbiased variance estimator of $V_p(\hat{Y}_\pi)$. Additionally, if n_1 (respectively,

n_2) is bounded, then the component S_2 (respectively, S_1) tends to be overestimated, rendering our proposed adaptation of the Rao-Wu-Yue bootstrap a conservative approach.

7 Simulation study

In this section, we conduct a simulation study in the $D = 2$ dimensional case, inspired by the setup outlined in Juillard et al. (2017). Our code to realize these simulations is accessible on github in the following link: <https://github.com/Jean-Rubin/CrossClassifiedSampling>. Our objective is to assess the performance of the proposed variance estimators for estimating both a total and a ratio. The underlying generation model we consider is:

$$y_{k_1 k_2} = \mu + \sigma_1 U_{k_1} + \sigma_2 V_{k_2} + \sigma_{12} W_{k_1 k_2}, \quad (7.1)$$

where the $U_{k_1}, V_{k_2}, W_{k_1 k_2}$ are independent standard normal variables.

7.1 Variance estimation for a total

In this simulation study, we generate data according to model (7.1) with the following parameters: $N_1 = N_2 = 1000$, $\mu = 200$, and $\sigma_1 = \sigma_2 = 5$. The parameter σ_{12} varies in $\{5, 10, 50\}$ to examine the impact of interaction variance.

We perform $T = 10,000$ iterations of sample selection using CCS with simple random sampling in each dimension, with varying sample sizes n_1 and n_2 chosen from $\{5, 10, 100, 500\}$. For each iteration, we estimate the total Y , and the estimator for the t^{th} sample is denoted as $\hat{Y}^{(t)}$. We assess the performance of the *Unbiased* variance estimator (see equation 4.11), the simplified variance estimator $\hat{V}^{SIMPL}(\hat{Y}_\pi)$ (*Simpl*) given in equation (4.13), and the simplified variance estimator $\hat{V}^{SIMPL2}(\hat{Y}_\pi)$ (*Simpl2*) given in equation (4.14). We also compute the bootstrap variance estimator $\hat{V}_{\text{Gross}}^{(t)}$ associated to the pseudo-population bootstrap (*Gross*), $\hat{V}_{\text{RaoWu}}^{(t)}$ associated to the rescaled bootstrap (*RaoWu*), and $\hat{V}_{\text{Skinner}}^{(t)}$ associated to a bootstrap procedure for with-replacement sampling in each dimension proposed by Skinner (2015) (*Skinner*). We use $B = 1,000$ bootstrap resamples for each estimator. The bootstrap variance estimator $\hat{V}^{(t)}$ is computed using the formula:

$$\hat{V}^{(t)} = \frac{1}{B-1} \sum_{b=1}^B \left(\hat{Y}^{(b^*)(t)} - \frac{1}{B} \sum_{b=1}^B \hat{Y}^{(b^*)(t)} \right)^2. \quad (7.2)$$

To evaluate these estimators, we compute the Monte-Carlo relative bias (RB) and relative stability (RS) as follows:

$$\begin{aligned}\text{RB}(\hat{V}) &= 100 \times \frac{\frac{1}{T} \sum_{t=1}^T \hat{V}^{(t)} - V}{V}, \\ \text{RS}(\hat{V}) &= 100 \times \frac{\sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{V}^{(t)} - V)^2}}{V},\end{aligned}$$

where V is an approximation of the true variance of \hat{Y} , computed from an independent run of 100,000 simulations. Additionally, confidence intervals for the bootstrap methods are computed using the reverse percentile method. More precisely, assume that the bootstrap estimates are reordered as $\hat{Y}^{(1^*)(t)} \leq \hat{Y}^{(2^*)(t)} \leq \dots \leq \hat{Y}^{(B^*)(t)}$. Then the reverse percentile confidence interval is obtained by considering $(\hat{Y}^* - \hat{Y})$ as an approximation of the distribution of $\hat{Y} - Y$, which leads to:

$$\text{CI}_\alpha^{\text{REV},(t)}(Y) = \left[2\hat{Y}^{(t)} - \hat{Y}^{(U^*)(t)}, 2\hat{Y}^{(t)} - \hat{Y}^{(L^*)(t)} \right]. \quad (7.3)$$

For $\text{CI}_\alpha^{\text{REV},(t)}(Y)$ to have the coverage level α , the values of L and U are given by $L = \lfloor \frac{1-\alpha}{2} B \rfloor$ and $U = \lfloor \frac{1+\alpha}{2} B \rfloor$. This comprehensive evaluation allows us to assess the accuracy, stability, and coverage of the proposed variance estimators under various conditions.

The simulation results are presented in Table 1. As anticipated, the relative bias of the two proposed simplified variance estimators approaches zero as both sample sizes increase. However, the first simplified variance estimator exhibits negative bias when both sample sizes are small or when σ_{12} is large, as the dropped variance component in these scenarios is non-negligible. Similarly, in these situations, the second simplified variance estimator displays positive bias, consistent with the heuristic reasoning outlined in Section 3.3. All three analytic variance estimators demonstrate similar levels of stability.

Turning to the bootstrap methods, Skinner's approach is conservative across all cases, as expected given its design for with-replacement sampling. As anticipated, the relative bias of both the Gross bootstrap and the Rao-Wu-Yue bootstrap variance estimators diminishes as both sample sizes increase. The bias of these two bootstrap variance estimators increases with σ_{12} , aligning with the formulation of the bootstrap variance outlined in equations (6.2) and (6.5). Overall, the bootstrap variance estimator exhibits smaller bias and greater stability under Gross's method, particularly for moderate sample sizes. While the coverage rates of the bootstrap confidence intervals are well

maintained when the variance estimators are approximately unbiased, they tend to be conservative when the variance estimators are conservative, as one would expect.

7.2 Variance estimation for a ratio

In the second population, the values for three variables of interest are generated according to the model

$$z_{k_1 k_2} = \sigma_1 U_{k_1} + \sigma_2 V_{k_2} + \sigma_{12} W_{k_1 k_2}, \quad (7.4)$$

$$x_{k_1 k_2} = \mu_x + \alpha z_{k_1, k_2} + (1 - \alpha)(\sigma'_1 U'_{k_1} + \sigma'_2 V'_{k_2} + \sigma'_{12} W'_{k_1 k_2}) \quad (7.5)$$

$$y_{k_1 k_2} = \mu_y + \beta z_{k_1, k_2} + (1 - \beta)(\sigma''_1 U''_{k_1} + \sigma''_2 V''_{k_2} + \sigma''_{12} W''_{k_1 k_2}). \quad (7.6)$$

The variables $U_{k_1}, U'_{k_1}, U''_{k_1}, V_{k_2}, V'_{k_2}, V''_{k_2}$, and $W_{k_1 k_2}, W'_{k_1 k_2}, W''_{k_1 k_2}$, are independent standard normal variables. The parameter α in equation (7.5) is used to control the correlation between variables x and z . Similarly, the parameter β in equation (7.6) is used to control the correlation between variables y and z . Thus, high values of α and β lead to high correlations between x and y . We used $\mu_x = 100$, $\mu_y = 300$, and $\sigma_1 = \sigma_2 = \sigma_{12} = \sigma'_1 = \sigma''_1 = 5, \sigma'_2 = \sigma''_2 = 10, \sigma'_{12} = \sigma''_{12} = 15$. Finally we considered $\alpha \in \{0.1, 0.5, 0.8\}$ and $\beta = 0.5$.

A Monte-carlo procedure similar to the one in the previous section is then applied to the substitution estimator $\hat{R} = \hat{Y}/\hat{X}$ of the ratio $R = Y/X$. For each sample $t \in \{1, \dots, T\}$, we produce bootstrap estimations $\hat{R}^{(b^*)}(t) = \hat{Y}^{(b^*)}(t)/\hat{X}^{(b^*)}(t)$ from the bootstrap versions $(\hat{Y}^{(b^*)}(t), \hat{X}^{(b^*)}(t))$ of $(\hat{Y}^{(t)}, \hat{X}^{(t)})$. We obtain a bootstrap variance estimator, and a confidence interval by means of the reverse percentile bootstrap. For the analytic variance estimators, linearized variance estimators are obtained by replacing the variable $y_{k_1 k_2}$ by the linearized variable $\nu_{k_1 k_2} = (y_{k_1 k_2} - \hat{R}x_{k_1 k_2})/\hat{X}$ in the three possible variance estimators.

The simulation results are presented in Table 2. The results are very similar to those obtained for total estimation. We have compiled in Table 3 the Monte-carlo mean values for the dispersions $S_{1,\eta}^2, S_{2,\eta}^2$ and $S_{12,\eta}^2$ associated with the variable $\eta_{k_1 k_2} = y_{k_1 k_2} - R x_{k_1 k_2}$. Results show that $S_{12,\eta}^2$ is comparatively larger as α decreases, and as for the estimation of a total, the bias of the variance estimators is larger in this situation.

Table 2: Relative biases, relative stabilities and coverage probabilities of multiple estimators of the variance of \hat{R} in percentage

α	n_1	n_2	RB										RS										CI _{95%}											
			Unbiased					Simpl					Gross					Unbiased					Simpl2					Gross					RaoWu	Skinner
			Unbiased	Simpl	Simpl2	Gross	RaoWu	Skinner	Unbiased	Simpl	Simpl2	Gross	RaoWu	Skinner	Unbiased	Simpl	Simpl2	Gross	RaoWu	Skinner	Unbiased	Simpl	Simpl2	Gross	RaoWu	Skinner								
0.1	5	5	-0.1	-25.6	26.1	22.4	60.6	61.1	70.7	76.6	77.5	67.7	60.6	61.1	93.8	96.3	60.6	61.1	93.8	96.8	60.6	61.1	93.8	96.8	96.3	96.3	96.3							
	10	10	0.2	-14.0	14.2	16.6	30.6	31.6	45.6	46.8	48.7	47.1	57.9	58.7	94.9	95.8	57.9	58.7	94.9	95.8	57.9	58.7	94.9	95.8	95.8	95.8	95.8							
	10	100	0.4	-6.1	5.9	3.4	12.1	16.4	35.4	35.5	35.8	33.8	39.0	40.8	94.1	95.2	39.0	40.8	94.1	94.8	39.0	40.8	94.1	94.8	95.2	95.2	95.2							
	100	100	0.3	-1.1	2.2	2.3	3.3	15.0	11.6	11.7	11.7	12.5	13.1	20.5	95.1	96.2	13.1	20.5	95.1	95.1	13.1	20.5	95.1	95.1	96.2	96.2	96.2							
	500	500	-0.6	-0.7	-0.4	0.4	0.4	101.5	3.9	4.0	3.9	6.0	6.0	102.2	94.9	99.4	6.0	102.2	94.9	95.2	6.0	102.2	94.9	95.2	99.4	99.4	99.4							
0.5	5	5	0.4	-24.2	24.5	17.3	50.8	51.8	66.0	70.8	70.2	58.4	85.2	87.6	93.9	96.4	85.2	87.6	93.9	96.3	85.2	87.6	93.9	96.3	96.4	96.4	96.4							
	10	10	-0.1	-14.6	13.7	15.5	29.1	31.2	42.1	44.2	43.9	42.3	52.5	54.1	95.3	96.1	52.5	54.1	95.3	96.0	52.5	54.1	95.3	96.0	96.1	96.1	96.1							
	10	100	-0.9	-5.1	3.9	-0.1	8.2	12.5	38.7	38.9	39.0	35.5	39.4	41.4	93.5	94.8	39.4	41.4	93.5	94.6	39.4	41.4	93.5	94.6	94.8	94.8	94.8							
	100	100	0.1	-1.2	1.6	2.1	3.3	14.8	11.4	11.5	11.6	12.5	12.9	20.3	96.3	96.3	12.9	20.3	94.8	94.7	12.9	20.3	94.8	94.7	96.3	96.3	96.3							
	500	500	0.1	0.1	0.3	1.3	1.5	103.3	3.9	4.0	4.0	6.1	6.2	104.0	95.4	99.4	6.2	104.0	95.4	95.2	6.2	104.0	95.4	95.2	99.4	99.4	99.4							
0.8	5	5	-0.7	-14.6	14.1	1.5	30.5	31.8	60.6	62.4	62.5	49.2	69.4	70.0	92.1	94.6	69.4	70.0	92.1	94.9	69.4	70.0	92.1	94.9	94.6	94.6	94.6							
	10	10	1.3	-7.6	7.8	4.0	16.1	17.3	37.2	37.2	37.0	34.1	40.8	41.5	93.9	95.1	40.8	41.5	93.9	95.3	40.8	41.5	93.9	95.3	95.1	95.1	95.1							
	10	100	-0.5	-1.7	1.4	-6.1	3.4	6.5	39.2	39.6	39.2	36.1	39.7	40.6	92.0	93.5	39.7	40.6	92.0	93.4	39.7	40.6	92.0	93.4	93.5	93.5	93.5							
	100	100	-0.4	-0.1	0.4	1.0	1.9	13.2	10.5	10.5	10.5	11.3	11.6	18.4	96.1	96.1	11.6	18.4	95.1	94.7	11.6	18.4	95.1	94.7	96.1	96.1	96.1							
	500	500	0.6	0.6	0.8	0.2	0.3	100.8	3.2	3.3	3.3	5.5	5.5	101.4	94.7	99.3	5.5	101.4	94.7	94.9	5.5	101.4	94.7	94.9	99.3	99.3	99.3							

Table 3: Monte Carlo mean value for the population dispersions $S_{1,\eta}^2$, $S_{2,\eta}^2$ and $S_{12,\eta}^2$

α	$S_{1,\eta}^2$	$S_{2,\eta}^2$	$S_{12,\eta}^2$
0.1	192	760	1,696
0.5	88	276	586
0.8	105	151	227

8 Application to the ELFE survey

ELFE is a French longitudinal survey that tracks children from birth to adulthood (Pirus et al., 2010; Charles et al., 2020). It commenced in France in 2011, with a survey of babies born in maternity wards, excluding the most premature infants. The sample was selected using two-dimensional cross-classified sampling, achieved by combining the population of 544 maternity wards with the population of 365 days. The population of maternity wards was divided into 5 strata of equal size based on the number of births in 2008. Similarly, the population of days was divided into 4 strata, each representing a season. The sampling design can be viewed as stratified simple random sampling in each dimension (Juillard et al., 2017). All babies born in the selected maternity wards on the designated days were surveyed, resulting in approximately 18,000 children after adjusting for unit non-response due to parental refusal to participate.

In this application, we are interested in the sub-population of maternity wards that hosted between 1,010 and 1,418 births in 2018. It represents our population \mathcal{U}_1 in the first dimension, of size $N_1 = 109$. The responding maternity wards in \mathcal{U}_1 are seen as coming from a simple random sample, corresponding to a sample S_1 of size $n_1 = 62$. The population \mathcal{U}_2 in the second dimension is that of all days, of size $N_2 = 365$. The sample of days selected for ELFE is seen as coming from a simple random sample, corresponding to a sample S_2 of size $n_2 = 25$. Thus, we obtain a sample $S = S_1 \times S_2$ of 1,550 maternity wards \times days.

We are interested in estimating the total number of births, the total number and percentage of twin births, the total number and percentage of births according to the marital status of the mother (married or remarried; divorced, single, or widowed; civil solidarity pact), and the average birth weight of the babies. We are also interested in variables related to the mother: we estimate the

average weight gain between the beginning and the end of pregnancy, and the average BMI at the end of pregnancy. For a total Y , we use the Horvitz-Thompson estimator \hat{Y}_π given in equation (2.8). For a ratio $R = Y^1/Y^2$, we use the substitution estimator $\hat{R}_\pi = \hat{Y}_\pi^1/\hat{Y}_\pi^2$.

We first consider analytic variance estimators. For the estimator of the total \hat{Y}_π , we consider the *Unbiased* variance estimator (see equation 4.11), the simplified variance estimator $\hat{V}^{SIMPL}(\hat{Y}_\pi)$ (*Simpl*) given in equation (4.13), and the simplified variance estimator $\hat{V}^{SIMPL2}(\hat{Y}_\pi)$ (*Simpl2*) given in equation (4.14). For the substitution estimator $\hat{R}_\pi = \hat{Y}_\pi^1/\hat{Y}_\pi^2$, we use linearization variance estimators obtained by replacing in the three previous estimators the variable of interest $y_{k_1k_2}$ with the linearized variable $e_{k_1k_2} = \{\hat{Y}_\pi^2\}^{-1}\{y_{k_1k_2}^2 - \hat{R}_\pi y_{k_1k_2}^1\}$. We also consider two bootstrap variance estimators for both \hat{Y}_π and \hat{R}_π . The first procedure is the pseudo-population bootstrap (see Section 6.1), with the variant proposed by Booth et al. (1994) to take into account the fact that N_1/n_1 and N_2/n_2 are not integers. The second procedure is the rescaled bootstrap (see Section 6.2). The results are presented in terms of coefficients of variation. The analytic and the bootstrap variance estimators are also plugged into normality-based confidence intervals, with a nominal one-tailed error rate of 2.5 %. For both bootstrap procedures, we also use a confidence interval based on the reverse percentile method (see equation 7.3).

The findings are presented in Table 4 concerning the estimation of counts for variables. It is observed that, in general, the various methods for estimating variance yield similar results, with the exception being the estimation of twin counts. Once again, this scenario demonstrates a notable disparity in the interaction term, with $\hat{S}_{12}^2 = 4.8 \cdot 10^{-2}$, $\hat{S}_1^2 = 4.4 \cdot 10^{-4}$, and $\hat{S}_2^2 = 8.2 \cdot 10^{-4}$. Notably, the variance estimator *Simpl* exhibits a negative bias in this case. Conversely, the variance estimator *Simpl2* displays a positive bias, albeit less pronounced than the bootstrap variance estimators. These observations extend to the estimation of ratios, as evidenced in Table 5.

9 Conclusion

In this paper, we investigated CSS designs across an arbitrary number of dimensions. We established the weak consistency and asymptotic normality of the HT-estimator under mild conditions. Specifically, we highlighted that these asymptotic properties necessitate the sample size to approach infinity in each dimension, presenting a significant limitation of CSS designs compared to multi-

Table 4: Estimations of the numbers for count variables, with estimated coefficient of variations and confidence intervals

	Birth	Twins	Marital status			
			Married/remarried	Divorced/single /widowed	Civil solidarity pact	
	\hat{Y}_π	49 487	975	21 766	513	27 20
<i>Unbiased</i>	\widehat{CV}	10.5%	27.7%	11.6%	10.6%	26.0%
	IC_{norm}	[39 291, 59 683]	[446, 1 504]	[16 807, 26 725]	[21 559, 32 857]	[252, 774]
<i>Simpl</i>	\widehat{CV}	10.4%	23.7%	11.4%	10.4%	21.8%
	IC_{norm}	[39 366, 59 608]	[523, 1 427]	[16 888, 26 644]	[21 641, 32 775]	[293, 733]
<i>Simpl2</i>	\widehat{CV}	10.6%	31.2%	11.8%	10.7%	29.5%
	IC_{norm}	[39 216, 59 758]	[379, 1 571]	[16 727, 26 805]	[21 477, 32 939]	[216, 810]
<i>Gross</i>	\widehat{CV}	10.6%	35.2%	11.9%	10.9%	32.8%
	IC_{norm}	[39 203, 59 771]	[302, 1 648]	[16 694, 26 838]	[21 382, 33 034]	[183, 843]
	IC_{rev}	[39 580, 60 166]	[205, 1 540]	[16 710, 26 874]	[21 510, 33 086]	[154, 821]
<i>Rao-Wu</i>	\widehat{CV}	10.6%	34.1%	11.9%	10.8%	33.2%
	IC_{norm}	[39 248, 59 726]	[324, 1 626]	[16 701, 26 831]	[21 447, 32 969]	[179, 847]
	IC_{rev}	[39 635, 59 822]	[249, 1 525]	[16 816, 26 841]	[21 594, 33 009]	[140, 801]

Table 5: Estimations of ratios, with estimated coefficient of variations and confidence intervals

	Twins	Marital status		Baby Birth weight	Mother	
		Married/remarried	Divorced/single /widowed		Weight gain	IMC
\hat{R}_π	1.97%	43.98%	54.98%	3 234	13.39	28.70
\widehat{CV}	24.3%	4.0%	3.2%	0.3%	0.9%	0.5%
IC_{norm}	[1.03%, 2.91%]	[40.56%, 47.40%]	[51.55%, 58.41%]	[3 217, 3 252]	[13.16, 13.62]	[28.43, 28.96]
\widehat{CV}	19.7%	3.6%	2.9%	0.2%	0.6%	0.4%
IC_{norm}	[1.21%, 2.73%]	[40.86%, 47.10%]	[51.85%, 58.11%]	[3 223, 3 246]	[13.24, 13.54]	[28.48, 28.91]
\widehat{CV}	28.1%	4.3%	3.4%	0.3%	1.1%	0.6%
IC_{norm}	[0.89%, 3.05%]	[40.29%, 47.67%]	[51.28%, 58.68%]	[3 196, 3 239]	[12.87, 13.44]	[28.11, 28.74]
\widehat{CV}	32.6%	4.7%	3.7%	0.4%	1.3%	0.6%
IC_{norm}	[0.71%, 3.23%]	[39.96%, 48.00%]	[50.96%, 59.00%]	[3 209, 3 260]	[13.05, 13.73]	[28.34, 29.05]
IC_{rev}	[0.60%, 3.07%]	[39.89%, 47.99%]	[50.93%, 59.11%]	[3 209, 3 260]	[13.04, 13.73]	[28.33, 29.06]
\widehat{CV}	31.1%	4.6%	3.7%	0.4%	1.3%	0.6%
IC_{norm}	[0.77%, 3.17%]	[40.04%, 47.92%]	[51.04%, 58.92%]	[3 209, 3 259]	[13.06, 13.72]	[28.35, 29.04]
IC_{rev}	[0.63%, 3.01%]	[39.92%, 47.92%]	[51.07%, 58.96%]	[3 209, 3 259]	[13.04, 13.70]	[28.34, 29.04]

stage sampling designs, where only the number of PSUs needs to approach infinity to ensure these properties (Chauvet and Vallée, 2020).

Additionally, we delved into the intricate matter of variance estimation for CSS (Juillard et al., 2017). We derived the Hoeffding-Sobol decomposition of the HT-estimator and proposed two straightforward variance estimators by focusing on the leading terms in this decomposition. Theoretical and empirical analyses demonstrate the consistency of these estimators as sample sizes increase. Our empirical findings indicate that the plug-in variance estimator tends to be conservative with smaller sample sizes, exhibiting a moderate positive bias as long as interaction dispersion remains moderate. Therefore, if an analytic variance estimator is preferred, we recommend its use.

Furthermore, we explored bootstrap variance estimation, an area which is little addressed in the literature concerning CSS. By employing a conditional Hoeffding-Sobol decomposition, we established that weighted bootstrap techniques are applicable for CSS, provided a weighted bootstrap technique is available for the HT-estimator in each dimension. This greatly facilitates the implementation of simple bootstrap procedures for CSS. Particularly, we proposed adaptations of the population bootstrap and the rescaled bootstrap for simple random sampling in each dimension. Empirical results validate the asymptotic unbiasedness of both bootstrap techniques as sample sizes increase. Additionally, they suggest that the population bootstrap yields variance estimators that are less biased and more stable with smaller sample sizes compared to the rescaled bootstrap.

Acknowledgement

The ELFE survey is a joint project between the French Institute for Demographic Studies (INED) and the National Institute of Health and Medical Research (INSERM), in partnership with the French blood transfusion service (Etablissement français du sang, EFS), Santé publique France, the National Institute for Statistics and Economic Studies (INSEE), the Direction générale de la santé (DGS, part of the Ministry of Health and Social Affairs), the Direction générale de la prévention des risques (DGPR, Ministry for the Environment), the Direction de la recherche, des études, de l'évaluation et des statistiques (DREES, Ministry of Health and Social Affairs), the Département des études, de la prospective et des statistiques (DEPS, Ministry of Culture), and the Caisse nationale des allocations familiales (CNAF), with the support of the Ministry of Higher Education

and Research and the Institut national de la jeunesse et de l'éducation populaire (INJEP). Via the RECONAI platform, it receives a government grant managed by the National Research Agency under the "Investissements d'avenir" programme (ANR-11-EQPX-0038 and ANR-19-COHO-0001).

A Proofs for Section 3.1

A.1 Proof of equation (3.1)

From equation (2.6), a straightforward computation leads to the result:

$$V_p(\hat{Y}) = \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} Cov_p \left(y_{\mathbf{k}} \prod_{d=1}^D w_{k_d}^d, y_{\mathbf{l}} \prod_{d=1}^D w_{l_d}^d \right) \quad (\text{A.1})$$

$$= \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} y_{\mathbf{k}} y_{\mathbf{l}} \left[E_p \left(\prod_{d=1}^D w_{k_d}^d w_{l_d}^d \right) - E_p \left(\prod_{d=1}^D w_{k_d}^d \right) E_p \left(\prod_{d=1}^D w_{l_d}^d \right) \right] \quad (\text{A.2})$$

$$= \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} y_{\mathbf{k}} y_{\mathbf{l}} \left[\prod_{d=1}^D E_p(w_{k_d}^d w_{l_d}^d) - \prod_{d=1}^D E_p(w_{k_d}^d) E_p(w_{l_d}^d) \right] \quad (\text{A.3})$$

where the last line follows from the independence of the weights in each dimension.

A.2 Proof of Proposition 1

Equations (3.2) as well as (3.5) in the first part of the proposition are general properties of the Hoeffding-Sobol decomposition applied to \hat{Y} seen as a function of the independent variables $(S_d)_{d=1}^D$, the general formula corresponding to (3.3). It remains to check if the equation (3.4) is true in our case by computing for each $I \subseteq \{1, \dots, D\}$ the associated \hat{Y}^I starting from the equation (3.3)

$$\hat{Y}^I = \sum_{I' \in \mathcal{P}(I)} (-1)^{|I|-|I'|} E_p(\hat{Y} | (S_d)_{d \in I'}) \quad (\text{A.4})$$

$$= \sum_{I' \in \mathcal{P}(I)} (-1)^{|I|-|I'|} \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} \prod_{d=1}^D \left[(w_{k_d}^d - E_p(w_{k_d}^d)) \mathbf{1}_{(d \in I')} + E_p(w_{k_d}^d) \right] \quad (\text{A.5})$$

$$= \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} \sum_{I' \in \mathcal{P}(I)} (-1)^{|I|-|I'|} \sum_{I'' \subseteq \{1, \dots, D\}} \prod_{d \in I''} (w_{k_d}^d - E_p(w_{k_d}^d)) \mathbf{1}_{(d \in I')} \prod_{d \notin I''} E_p(w_{k_d}^d). \quad (\text{A.6})$$

By observing that for any $I'' \subseteq \{1, \dots, D\}$, $\prod_{d \in I''} \mathbf{1}_{(d \in I')} = \mathbf{1}_{(I'' \subseteq I')}$ and that $\sum_{I' \in \mathcal{P}(I)} (-1)^{|I|-|I'|} \mathbf{1}_{(I'' \subseteq I')} = \mathbf{1}_{(I''=I)}$, one can conclude by rearranging the sum that

$$\hat{Y}^I = \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} \sum_{I'' \subseteq \{1, \dots, D\}} \prod_{d \in I''} (w_{k_d}^d - E_p(w_{k_d}^d)) \prod_{d \notin I''} E_p(w_{k_d}^d) \sum_{I' \in \mathcal{P}(I)} (-1)^{|I|-|I'|} \mathbf{1}_{(I'' \subseteq I')} \quad (\text{A.7})$$

$$= \sum_{\mathbf{k} \in \mathcal{U}} y_{\mathbf{k}} \prod_{d \in I} (w_{k_d}^d - E_p(w_{k_d}^d)) \prod_{d \notin I} E_p(w_{k_d}^d), \quad (\text{A.8})$$

from which we can deduce the value of $V_p(\hat{Y}^I)$ for $I \neq \emptyset$:

$$V_p(\hat{Y}^I) = \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} y_{\mathbf{k}} y_{\mathbf{l}} \text{Cov}_p \left(\prod_{d \in I} (w_{k_d}^d - E_p(w_{k_d}^d)), \prod_{d \in I} (w_{l_d}^d - E_p(w_{l_d}^d)) \right) \prod_{d \notin I} E_p(w_{k_d}^d) E_p(w_{l_d}^d) \quad (\text{A.9})$$

$$= \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} y_{\mathbf{k}} y_{\mathbf{l}} E_p \left(\prod_{d \in I} (w_{k_d}^d - E_p(w_{k_d}^d))(w_{l_d}^d - E_p(w_{l_d}^d)) \right) \prod_{d \notin I} E_p(w_{k_d}^d) E_p(w_{l_d}^d) \quad (\text{A.10})$$

$$= \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} y_{\mathbf{k}} y_{\mathbf{l}} \prod_{d \in I} \text{Cov}_p(w_{k_d}^d, w_{l_d}^d) \prod_{d \notin I} E_p(w_{k_d}^d) E_p(w_{l_d}^d). \quad (\text{A.11})$$

where the last line follows from the independence of the weights in each dimension.

A.3 Proof of Corollary 1

It suffices to show that for every non-empty subset $I \subseteq \{1, \dots, D\}$, $V_p(\hat{Y}^I)$ is unbiasedly estimated by $\hat{V}_p(\hat{Y}^I)$. It is possible to rewrite the sum over \mathcal{U} by introducing the sample membership indicators. We then obtain

$$E_p(\hat{V}_p(\hat{Y}^I)) = \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} y_{\mathbf{k}} y_{\mathbf{l}} E_p \left[\prod_{d \in I} \widehat{\text{Cov}}^d(w_{k_d}^d, w_{l_d}^d) \prod_{d \notin I} \frac{\delta_{k_d}^d \delta_{l_d}^d}{\pi_{k_d, l_d}^d} \right] \quad (\text{A.12})$$

$$= \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} y_{\mathbf{k}} y_{\mathbf{l}} \prod_{d \in I} E_p \left[\widehat{\text{Cov}}^d(w_{k_d}^d, w_{l_d}^d) \right] \prod_{d \notin I} \frac{E_p(\delta_{k_d}^d \delta_{l_d}^d)}{\pi_{k_d, l_d}^d} \quad (\text{A.13})$$

$$= \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} y_{\mathbf{k}} y_{\mathbf{l}} \prod_{d \in I} \text{Cov}_p(w_{k_d}^d, w_{l_d}^d) \quad (\text{A.14})$$

$$= V_p(\hat{Y}^I). \quad (\text{A.15})$$

where the second line follows from the fact that for a fixed \mathbf{k} and $\mathbf{l} \in \mathcal{U}$, each $\widehat{\text{Cov}}^d(w_{k_d}^d, w_{l_d}^d)$ and $\delta_{k_d}^d \delta_{l_d}^d$ are built from the sample S_d only and are thus all independent from each other as long as they correspond to different dimensions.

B Proofs for Section 3.2

B.1 Proof of Corollary 2

Equation (3.12) is a direct consequence of equation (3.4), applied for $w_{k_d}^d = \delta_{k_d}^d / \pi_{k_d}^d$. Similarly, equation (3.13) follows from equation (3.6). Finally, equation (3.14) follows by applying equation (3.9) with

$$\widehat{\text{Cov}}^d(w_k^d, w_l^d) = \frac{\Delta_{k_d l_d}^d \delta_{k_d}^d \delta_{l_d}^d}{\pi_{k_d, l_d}^d \pi_{k_d}^d \pi_{l_d}^d}. \quad (\text{B.1})$$

B.2 Proof of Proposition 2

Let us introduce the partition $\mathcal{U}_I \times \mathcal{U}_I = \bigcup_{I' \in \mathcal{P}(I)} P_{I'}$ for a given non-empty $I \subseteq \{1, \dots, D\}$, where we defined for each $I' \in \mathcal{P}(I)$, the set $P_{I'}$ by

$$P_{I'} = \{\mathbf{k}', \mathbf{l}' \in \mathcal{U}_I, \forall d \in I, (k'_d \neq l'_d \Leftrightarrow d \in I')\} \quad (\text{B.2})$$

furthermore, note that we can identify every element $\mathbf{k} \in \mathcal{U}$ with a couple $(\mathbf{k}_1, \mathbf{k}_2) \in \mathcal{U}_I \times \mathcal{U}_{I^c}$ where we introduced $\mathcal{U}_{I^c} = \prod_{d \notin I} \mathcal{U}_d$. More precisely, the coordinates of \mathbf{k}_1 are given by the coordinates of \mathbf{k} associated to the dimensions in I and the coordinates of \mathbf{k}_2 are given by the coordinates of \mathbf{k} associated to the dimensions that are not in I . We will therefore denote by $y_{\mathbf{k}_1 \mathbf{k}_2}$ the quantity $y_{\mathbf{k}}$. This splitting allows us for example to rewrite the subtotal $Y_{\mathbf{k}'}$ in the form $Y_{\mathbf{k}'} = \sum_{I' \in \mathcal{U}_{I^c}} y_{\mathbf{k}' I'}$. Now let us fix some subset $I' \in \mathcal{P}(I)$. We can verify using the Cauchy-Schwarz inequality that

$$\sum_{\mathbf{k}' \in \mathcal{U}_I} Y_{\mathbf{k}'}^2 = \sum_{\mathbf{k}' \in \mathcal{U}_I} \left(\sum_{I' \in \mathcal{U}_{I^c}} y_{\mathbf{k}' I'} \right)^2 \leq \left(\prod_{d \notin I} N_d \right) \sum_{\mathbf{k}' \in \mathcal{U}_I} \sum_{I' \in \mathcal{U}_{I^c}} y_{\mathbf{k}' I'}^2 \leq \alpha \frac{N^2}{\prod_{d \in I} N_d} \quad (\text{B.3})$$

by using the assumption (H1).

Now we can similarly identify an element $\mathbf{k}' \in \mathcal{U}_I$ with a couple $(\mathbf{k}'_1, \mathbf{k}'_2) \in \mathcal{U}_{I'} \times \mathcal{U}_{I \setminus I'}$. We will therefore denote by $Y_{\mathbf{k}'_1 \mathbf{k}'_2}$ the subtotal $Y_{\mathbf{k}'}$. Thus, using again the Cauchy-Schwarz inequality, it is now easy to verify that

$$\sum_{(\mathbf{k}', I') \in P_{I'}} Y_{\mathbf{k}'} Y_{I'} \leq \sum_{\substack{\mathbf{k}'', I'' \in \mathcal{U}_{I'} \\ \mathbf{m}'' \in \mathcal{U}_{I \setminus I'}}} Y_{\mathbf{k}'' \mathbf{m}''} Y_{I'' \mathbf{m}''} \quad (\text{B.4})$$

$$= \sum_{\mathbf{m}'' \in \mathcal{U}_{I \setminus I'}} \left(\sum_{\mathbf{k}'' \in \mathcal{U}_{I'}} Y_{\mathbf{k}'' \mathbf{m}''} \right)^2 \quad (\text{B.5})$$

$$\leq \left(\prod_{d \in I'} N_d \right) \sum_{\mathbf{m}'' \in \mathcal{U}_{I \setminus I'}} \sum_{\mathbf{k}'' \in \mathcal{U}_{I'}} Y_{\mathbf{k}'' \mathbf{m}''}^2 \quad (\text{B.6})$$

$$\leq \alpha \frac{N^2}{\prod_{d \in I \setminus I'} N_d}. \quad (\text{B.7})$$

Furthermore we know by construction of $P_{I'}$ that for any $(\mathbf{k}', I') \in P_{I'}$, $k'_d \neq l'_d$ if and only if $d \in I'$ and thus using assumption (H3)-(H4) we get that

$$\forall (\mathbf{k}', I') \in P_{I'}, \forall d \in I', \left| \frac{\Delta_{k'_d, l'_d}^d}{\pi_{k'_d}^d \pi_{l'_d}^d} \right| \leq \frac{\gamma_d}{\lambda_d^2} \frac{1}{n_d}. \quad (\text{B.8})$$

On the other hand, when $d \notin I'$, we have that $k'_d = l'_d$ and thus

$$\forall (\mathbf{k}', I') \in P_{I'}, \forall d \notin I', \left| \frac{\Delta_{k'_d, l'_d}^d}{\pi_{k'_d}^d \pi_{l'_d}^d} \right| = \frac{1 - \pi_{k'_d}^d}{\pi_{k'_d}^d} \leq \frac{1}{\lambda_d} \frac{N_d}{n_d}. \quad (\text{B.9})$$

From there, we can build an upper bound of the sum

$$\sum_{(\mathbf{k}', I') \in P_{I'}} Y_{\mathbf{k}'} Y_{I'} \prod_{d \in I} \frac{\Delta_{k'_d l'_d}^d}{\pi_{k'_d}^d \pi_{l'_d}^d} = \sum_{(\mathbf{k}', I') \in P_{I'}} Y_{\mathbf{k}'} Y_{I'} \prod_{d \in I'} \frac{\Delta_{k'_d l'_d}^d}{\pi_{k'_d}^d \pi_{l'_d}^d} \prod_{d \in I \setminus I'} \frac{\Delta_{k'_d l'_d}^d}{\pi_{k'_d}^d \pi_{l'_d}^d} \quad (\text{B.10})$$

$$= O \left(\frac{N^2}{\prod_{d \in I \setminus I'} N_d} \prod_{d \in I'} \frac{1}{n_d} \prod_{d \in I \setminus I'} \frac{N_d}{n_d} \right) \quad (\text{B.11})$$

$$= O \left(\frac{N^2}{\prod_{d \in I} n_d} \right). \quad (\text{B.12})$$

Now, we can finally conclude that

$$V_p(\hat{Y}^I) = \sum_{I' \in \mathcal{P}(I)} \left(\sum_{(\mathbf{k}', I') \in P_{I'}} Y_{\mathbf{k}'} Y_{I'} \prod_{d \in I} \frac{\Delta_{k'_d l'_d}^d}{\pi_{k'_d}^d \pi_{l'_d}^d} \right) = O \left(\frac{N^2}{\prod_{d \in I} n_d} \right). \quad (\text{B.13})$$

B.3 Proof of Proposition 3

We can write

$$\hat{Y}_\pi - Y = \sum_{d=1}^D \hat{Y}_\pi^{\{d\}} + \Delta, \text{ where } \Delta = \sum_{\substack{I \subset \{1, \dots, D\} \\ \text{Card}(I) \geq 2}} \hat{Y}_\pi^I.$$

We obtain

$$V_p(\hat{Y}_\pi) = \sum_{d=1}^D V_p(\hat{Y}_\pi^{\{d\}}) + V_p(\Delta). \quad (\text{B.14})$$

It follows from Proposition 2 that

$$V_p(N^{-1} \hat{Y}_\pi^{\{d\}}) = O(n_m^{-1}) \quad \text{and} \quad V_p(N^{-1} \Delta) = o(n_m^{-1}). \quad (\text{B.15})$$

Therefore, we obtain (3.19). Note that from equations (B.14) and (B.15), the constants in assumption (H7) are such that $\sum_{d=1}^D (\gamma_d)^2 = 1$.

From equation (B.14), we also obtain

$$\frac{\hat{Y}_\pi - Y}{\sqrt{V_p(\hat{Y}_\pi)}} = \underbrace{\sum_{d=1}^D \gamma_d \frac{\hat{Y}_\pi^{\{d\}}}{\sqrt{V_p(\hat{Y}_\pi^{\{d\}})}}}_{\Delta_1} + \underbrace{\sum_{d=1}^D \left(\sqrt{\frac{V_p(\hat{Y}_\pi^{\{d\}})}{V_p(\hat{Y}_\pi)}} - \gamma_d \right) \frac{\hat{Y}_\pi^{\{d\}}}{\sqrt{V_p(\hat{Y}_\pi^{\{d\}})}}}_{\Delta_2} + \underbrace{\frac{\Delta}{\sqrt{V_p(\hat{Y}_\pi)}}}_{\Delta_3}. \quad (\text{B.16})$$

By using assumption (H5) and the right-hand side of (B.14), we have $\Delta_3 \xrightarrow{P_r} 0$, where $\xrightarrow{P_r}$ stands for the convergence in probability. Also, by using Assumptions (H6) and (H7) and the Slutsky theorem, we obtain that $\Delta_2 \xrightarrow{P_r} 0$. Finally, since the variables $\hat{Y}_\pi^{\{d\}}$ are independent, we obtain from assumption (H6) by standard arguments that $\Delta_1 \xrightarrow{P_r} \mathcal{N}(0, 1)$, which completes the proof.

B.4 Proof of Proposition 4

Let us fix a non-empty subset $I \subseteq \{1, \dots, D\}$. We can write using the definitions of $\hat{V}^{PLUG}(\hat{Y}_\pi^I)$

$$E_p \left[\hat{V}^{PLUG}(\hat{Y}_\pi^I) \right] = \sum_{\mathbf{k}', \mathbf{l}' \in \mathcal{U}_I} \frac{E_p(\hat{Y}_{\mathbf{k}'} \hat{Y}_{\mathbf{l}'} \delta_{\mathbf{k}'} \delta_{\mathbf{l}'})}{\pi_{\mathbf{k}'} \pi_{\mathbf{l}'}} \frac{\Delta_{\mathbf{k}'\mathbf{l}'}}{\pi_{\mathbf{k}'\mathbf{l}'}} \quad (\text{B.17})$$

$$= \sum_{\mathbf{k}', \mathbf{l}' \in \mathcal{U}_I} \frac{E_p(\hat{Y}_{\mathbf{k}'} \hat{Y}_{\mathbf{l}'}) E_p(\delta_{\mathbf{k}'} \delta_{\mathbf{l}'})}{\pi_{\mathbf{k}'} \pi_{\mathbf{l}'}} \frac{\Delta_{\mathbf{k}'\mathbf{l}'}}{\pi_{\mathbf{k}'\mathbf{l}'}} \quad (\text{B.18})$$

$$= \sum_{\mathbf{k}', \mathbf{l}' \in \mathcal{U}_I} \frac{E_p(\hat{Y}_{\mathbf{k}'} \hat{Y}_{\mathbf{l}'})}{\pi_{\mathbf{k}'} \pi_{\mathbf{l}'}} \Delta_{\mathbf{k}'\mathbf{l}'} \quad (\text{B.19})$$

Where the line (B.18) follows from the fact that $\delta_{\mathbf{k}'}$ and $\delta_{\mathbf{l}'}$ are functions of $(S_d)_{d \in I}$ whereas $\hat{Y}_{\mathbf{k}'}$ and $\hat{Y}_{\mathbf{l}'}$ are functions of $(S_d)_{d \notin I}$, and are therefore independent. We can then similarly expand the sum by replacing $\hat{Y}_{\mathbf{k}'}$ and $\hat{Y}_{\mathbf{l}'}$ by their definitions in (3.25):

$$E_p \left[\hat{V}^{PLUG}(\hat{Y}_\pi^I) \right] = \sum_{\mathbf{k}', \mathbf{l}' \in \mathcal{U}_I} \sum_{\substack{\mathbf{k}, \mathbf{l} \in \mathcal{U} \\ \forall d \in I, k_d = k'_d, l_d = l'_d}} \frac{y_{\mathbf{k}} y_{\mathbf{l}}}{\pi_{\mathbf{k}} \pi_{\mathbf{l}}} \prod_{d \in I} \Delta_{k_d l_d}^d E_p \left(\prod_{d \notin I} \delta_{k_d}^d \delta_{l_d}^d \right) \quad (\text{B.20})$$

$$= \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} \frac{y_{\mathbf{k}} y_{\mathbf{l}}}{\pi_{\mathbf{k}} \pi_{\mathbf{l}}} \prod_{d \in I} \Delta_{k_d l_d}^d \prod_{d \notin I} \pi_{k_d}^d \pi_{l_d}^d. \quad (\text{B.21})$$

Now we can furthermore observe that for any $\mathbf{k}, \mathbf{l} \in \mathcal{U}$, we have the following identity

$$\prod_{d \notin I} \pi_{k_d l_d}^d = \prod_{d \notin I} (\Delta_{k_d l_d}^d + \pi_{k_d}^d \pi_{l_d}^d) = \sum_{I' \subseteq I^c} \prod_{d \in I'} \Delta_{k_d l_d}^d \prod_{d \in I^c \setminus I'} \pi_{k_d}^d \pi_{l_d}^d \quad (\text{B.22})$$

which can then be substituted in the previous result to give

$$E_p \left[\hat{V}^{PLUG}(\hat{Y}_\pi^I) \right] = \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} \frac{y_{\mathbf{k}} y_{\mathbf{l}}}{\pi_{\mathbf{k}} \pi_{\mathbf{l}}} \prod_{d \in I} \Delta_{k_d l_d}^d \sum_{I' \subseteq I^c} \prod_{d \in I'} \Delta_{k_d l_d}^d \prod_{d \in I^c \setminus I'} \pi_{k_d}^d \pi_{l_d}^d \quad (\text{B.23})$$

$$= \sum_{I' \subseteq I^c} \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} \frac{y_{\mathbf{k}}}{\prod_{d \in I \cup I'} \pi_{k_d}^d} \frac{y_{\mathbf{l}}}{\prod_{d \in I \cup I'} \pi_{l_d}^d} \prod_{d \in I \cup I'} \Delta_{k_d l_d}^d \quad (\text{B.24})$$

Finally we can note that $\{I \cup I', I' \subseteq I^c\} = \{I', I \subseteq I' \subseteq \{1, \dots, D\}\}$. From this observation, we can make a change of variable and conclude

$$E_p \left[\hat{V}^{PLUG}(\hat{Y}_\pi^I) \right] = \sum_{I' \supseteq I} \sum_{\mathbf{k}, \mathbf{l} \in \mathcal{U}} \frac{y_{\mathbf{k}}}{\prod_{d \in I'} \pi_{k_d}^d} \frac{y_{\mathbf{l}}}{\prod_{d \in I'} \pi_{l_d}^d} \prod_{d \in I'} \Delta_{k_d l_d}^d \quad (\text{B.25})$$

$$= \sum_{I' \supseteq I} \sum_{\substack{\mathbf{k}'', \mathbf{l}'' \in \mathcal{U}_{I'} \\ \mathbf{k}, \mathbf{l} \in \mathcal{U} \\ \forall d \in I', k_d = k''_d, l_d = l''_d}} \frac{y_{\mathbf{k}''}}{\prod_{d \in I'} \pi_{k_d}^d} \frac{y_{\mathbf{l}''}}{\prod_{d \in I'} \pi_{l_d}^d} \prod_{d \in I'} \Delta_{k_d l_d}^d \quad (\text{B.26})$$

$$= \sum_{I' \supseteq I} \sum_{\mathbf{k}'', \mathbf{l}'' \in \mathcal{U}_{I'}} \frac{Y_{\mathbf{k}''}}{\pi_{\mathbf{k}''}} \frac{Y_{\mathbf{l}''}}{\pi_{\mathbf{l}''}} \Delta_{\mathbf{k}''\mathbf{l}''} \quad (\text{B.27})$$

$$= \sum_{I' \supseteq I} V_p(\hat{Y}_\pi^{I'}). \quad (\text{B.28})$$

References

- Beaumont, J.-F. and Patak, Z. (2012). On the generalized bootstrap for sample surveys with special attention to poisson sampling. *International Statistical Review*, 80:127–148.
- Booth, J. G., Butler, R. W., and Hall, P. (1994). Bootstrap methods for finite populations. *Journal of the American Statistical Association*, 89(428):1282–1289.
- Charles, M. A., Thierry, X., Lanoe, J.-L., Bois, C., Dufourg, M.-N., Popa, R., Cheminat, M., Zaros, C., and Geay, B. (2020). Cohort profile: the french national cohort of children (elfe): birth to 5 years. *International journal of epidemiology*, 49(2):368–369j.
- Chauvet, G. and Vallée, A.-A. (2020). Inference for two-stage sampling designs. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82(3):797–815.
- Cochran, W. G. (1977). *Sampling techniques*. John Wiley & Sons, New York-London-Sydney, third edition. Wiley Series in Probability and Mathematical Statistics.
- Dalén, J. and Ohlsson, E. (1995). Variance estimation in the swedish consumer price index. *Journal of Business & Economic Statistics*, 13(3):347–356.
- Fuller, W. A. (2011). *Sampling statistics*, volume 560. John Wiley & Sons.
- Gross, S. (1980). Median estimation in sample surveys. *Proceedings of the Section on Survey Research Methods*, pages 181–184.
- Hájek, J. (1964). Asymptotic theory of rejective sampling with varying probabilities from a finite population. *The Annals of Mathematical Statistics*, pages 1491–1523.
- Hoeffding, W. (1948). A Class of Statistics with Asymptotically Normal Distribution. *The Annals of Mathematical Statistics*, 19(3):293 – 325.
- Juillard, H., Chauvet, G., and Ruiz-Gazen, A. (2017). Estimation under cross-classified sampling with application to a childhood survey. *Journal of the American Statistical Association*, 112(518):850–858.
- Ohlsson, E. (1996). Cross-classified sampling. *Journal of Official Statistics*, 12(3):241.

- Pirus, C., Bois, C., Dufourg, M.-N., Lanoë, J.-L., Vandentorren, S., Leridon, H., Depledge, R., and the Elfe team (2010). Constructing a cohort: Experience with the french elfe project. *Population*, 65:637–670.
- Rao, J., Wu, C., and Yue, K. (1992). Some recent work on resampling methods for complex surveys. *Survey methodology*, 18(2):209–217.
- Särndal, C.-E., Swensson, B., and Wretman, J. (1992). *Model Assisted Survey Sampling*. Springer New York.
- Skinner, C. (2015). Cross-classified sampling: some estimation theory. *Statistics & Probability Letters*, 104:163–168.