Properties of Chromy's sampling procedure

Guillaume Chauvet

Properties of Chromy's sampling procedure

Guillaume Chauvet*

Abstract

Chromy (1979) proposed a unequal probability sampling algorithm, which enables to select a sample in one pass of the sampling frame only. This is the default sequential method used in the `SURVEYSELECT` procedure of the `SAS` software. In this article, we study the properties of Chromy sampling. We prove that the Horvitz-Thompson is asymptotically normally distributed, and give an explicit expression for the second-order inclusion probabilities. This makes it possible to estimate the variance unbiasedly for the randomized version of the method programmed in the `SURVEYSELECT` procedure.

1 Introduction

Chromy (1979) proposed a fixed-size unequal probability sampling design which is strictly sequential, in the sense that a sample is selected in one pass of the sampling frame only. This algorithm benefits from a stratification effect, in the sense that the selected units are well spread over the population like with systematic sampling. The drawback is that many second-order inclusion probabilities are zero, making unbiased variance estimation not possible.

Chromy (1979) therefore proposed to partially randomize the order of the units in the population before applying the sampling algorithm. This randomization is sufficient to guarantee that the second-order inclusion probabilities are positive. The randomized Chromy algorithm is the default sequential method currently available in the `SURVEYSELECT` procedure of the `SAS` software. The method has been extensively used for sample surveys, see

* *Univ. Rennes, Ensai, CNRS, IRMAR - UMR 6625*

for example Mills et al. (2018), Radwin et al. (2018), Schmitt et al. (2018) and Rust et al. (2019) for recent examples.

So far, the properties of this sampling algorithm have not been fully investigated, and this is the purpose of the current paper. We prove that Chromy sampling is equivalent to ordered pivotal sampling (Deville and Tillé, 1998; Chauvet, 2012), in the sense that both algorithms lead to the same sampling design. This leads to the Horvitz-Thompson being consistent and asymptotically normally distributed, under weak assumptions. This also leads to an explicit formula for the second-order inclusion probabilities, making unbiased variance estimation possible for randomized Chromy sampling.

The paper is organized as follows. In Section 2, the notation and the assumptions are given. Chromy sampling and ordered pivotal sampling are introduced in Sections 3 and 4. The equivalence between both sampling designs is proved in Section 5, and the properties of Chromy sampling are studied. The results of a small simulation study are given in Section 6. We conclude in Section 7. The proofs are gathered in the Appendix.

2 Notation and assumptions

We consider a finite population U of N sampling units that may be represented by integers $k = 1, \ldots, N$. Denote by $\pi = (\pi_1, \ldots, \pi_N)^\top$ a vector of probabilities, with $0 < \pi_k < 1$ for any unit $k \in U$, and with $n = \sum_{k \in U} \pi_k$ the expected sample size. A random sample S is selected in U by means of a sampling design $p(\cdot)$ with parameter π, in the sense that the expected number of draws for unit k is π_k. We let I_k denote the number of times that unit k is selected in the sample, and we note $I = (I_1, \ldots, I_N)^\top$.

The set of probabilities π may be defined proportionally on some positive auxiliary variable known for any unit in the population, which leads to unequal probability sampling with probabilities proportional to size (π-ps) (Särndal et al., 1992, Section 3.6.2). The sampling algorithm proposed by Chromy (1979) may handle π_k's greater than 1, in which case a same unit may be selected several times in the sample. In this paper, we focus on the fairly usual situation when all π_k's lie between 0 and 1, which means that the sampling design is without replacement. In this case, we may interpret π_k as

the probability for unit k to be included in the sample, and I_k is the sample membership indicator for unit k. We are interested in the total $t_y = \sum_{k \in U} y_k$ of some variable of interest taking the value y_k for unit $k \in U$. The Horvitz-Thompson estimator is

$$\hat{t}_{y\pi} \quad = \quad \sum_{k \in U} \frac{y_k}{\pi_k} I_k. \tag{2.1}$$

In order to study Chromy sampling, some additional notation is needed. We let $V_k = \sum_{l=1}^{k} \pi_l$ denote the cumulated inclusion probabilities up to unit k, with $V_0 = 0$. The integer part of V_k is the largest integer smaller than V_k, and is denoted as V_k^I. The difference between V_k and its integer part V_k^I is the fractional part, and is denoted as $V_k^F = V_k - V_k^I$. For example, if $V_k = 3.6$ we have $V_k^I = 3$ and $V_k^F = 0.6$, and if $V_k = 4.0$ we have $V_k^I = 4$ and $V_k^F = 0$.

A unit k is a *cross-border* if $V_{k-1} \leq i$ and $V_k > i$ for some integer $i = 1, \ldots, n-1$. The cross-border units are denoted as $k_i,\ i = 1, \ldots, n-1$, and we note $a_i = i - V_{k_i - 1}$ and $b_i = V_{k_i} - i$. The cross-border units define a partition of the population into microstrata $U_i,\ i = 1, \ldots, n$, which are defined as

$$U_i \quad = \quad \{k \in U;\ k_{i-1} \leq k \leq k_i\} \text{ with } k_0 = 0 \text{ and } k_n = N + 1. \tag{2.2}$$

The quantities are presented in Figure 1 for illustration.

We consider the following assumptions, which are the same than in Chauvet and Le Gleut (2019):

H1: There exists some constants $0 < f_0$ and $f_1 < 1$ such that for any $k \in U$:

$$f_0 \frac{n}{N} \quad \leq \pi_k \leq \quad f_1. \tag{2.3}$$

H2: There exists some constant C_1 such that:

$$\sum_{k \in U} \pi_k \left(\frac{y_k}{\pi_k} - \frac{t_y}{n} \right)^4 \quad \leq \quad C_1 N^4 n^{-3}. \tag{2.4}$$

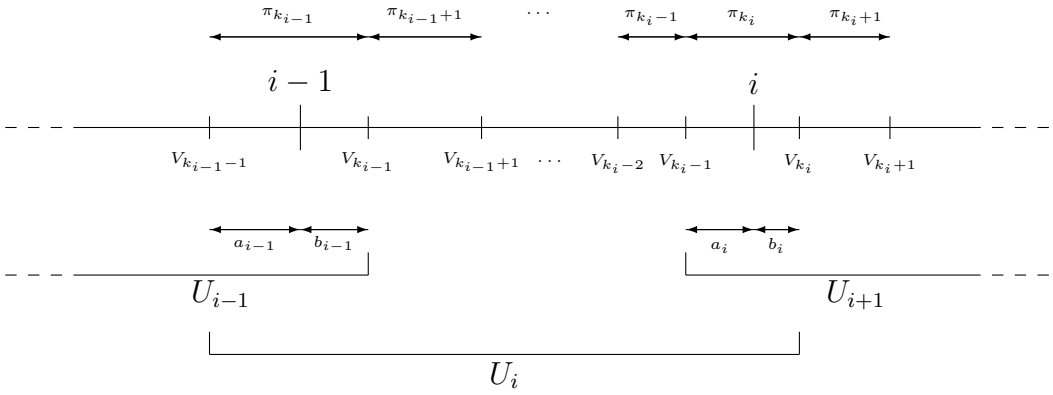

Figure 1: Inclusion probabilities and cross-border units in microstratum U_i, for population U

H3: There exists some constant $C_2 > 0$ such that:

$$\sum_{i=1}^{n} \sum_{k \in U_i} \alpha_{ik} \left(\frac{y_k}{\pi_k} - \sum_{l \in U_i} \alpha_{il} \frac{y_l}{\pi_l} \right)^2 \geq C_2 N^2 n^{-1}, \tag{2.5}$$

where for any unit $k \in U_i$ we take:

$$\alpha_{ik} = \begin{cases} b_{i-1} & \text{if } k = k_{i-1}, \\ \pi_k & \text{if } k_{i-1} < k < k_i, \\ a_i & \text{if } k = k_i, \end{cases}$$

with the convention that $b_0 = a_n = 0$.

It is assumed in (H1) that the first-order inclusion probabilities are bounded away from 1. This is not a severe restriction in practice, since a unit with an inclusion probability close to 1 is usually placed into a take-all stratum, i.e. the probability is rounded to 1 and the unit is not involved in the selection process. It is also assumed in (H1) that the first-order inclusion probabilities have a lower bound of order n/N, which ensures that no design-weight $d_k = 1/\pi_k$ is disproportionately larger than the others.

Under the condition (H1), the condition (H2) holds in particular if the variable y has a finite moment of order 4. This seems a fair assumption in practice, unless the variable of interest is heavily skewed like in wealth sur-

veys, for example.

Assumption (H3) is somewhat technical, and is used to ensure that the Horvitz-Thompson estimator is not close to being degenerate. This assumption is only needed to prove a central-limit theorem for the Horvitz-Thompson estimator. It requires that the dispersion within the microstrata does not vanish. For example, it does not hold if y_k is proportional to π_k.

3 Chromy sampling

Chromy (1979) proposed a sampling algorithm which is strictly sequential, in the sense that the units in the population are successively considered for possible selection, and the decision for the unit is made at once. The method is presented in Algorithm 1. Let us denote

$$S_c \sim Chr(\pi; U) \tag{3.1}$$

for a sample selected by means of Chromy sampling with parameter π in the population U. The method was originally proposed for π-ps sampling where the inclusion probabilities are defined proportionally on some auxiliary variable, but it is applicable to any set of inclusion probabilities.

The method proceeds by considering at each step $k = 1, \ldots, N$ the unit k for possible selection, and by computing its probability of selection conditionally on the number of units already selected. This algorithm defines a fixed-size sampling design, and Chromy (1979) proves that the parameter π defining the inclusion probabilities is exactly matched. Note that the case $V_k^F > V_{k-1}^F$ (Step 2.a) corresponds to the treatment of a non cross-border unit, while the case $V_{k-1}^F \geq V_k^F$ (Step 2.b) corresponds to the treatment of a cross-border unit. For illustration, the complete probability tree for Chromy sampling on a small population is given in Appendix A.

This algorithm allocates the sample regularly in the population, as stated in Proposition 1. The proof is given in Appendix B. At any step k of the procedure, the number of units selected is equal to the sum of inclusion probabilities up to rounding, a property which is sometimes coined as spatial balancing (Grafström et al., 2012).

Algorithm 1 Chromy sampling with parameter π in the population U

1. At step $k = 1$, take $I_1 = 1$ with probability π_1.

2. At step $k = 2, \ldots, N$:

   (a) If $V_k^F > V_{k-1}^F$, then we take the transition probabilities

$$\begin{aligned} Pr\left(I_k = 1 \left| \sum_{l=1}^{k-1} I_l = V_{k-1}^I \right. \right) &= \frac{V_k^F - V_{k-1}^F}{1 - V_{k-1}^F}, \\ Pr\left(I_k = 1 \left| \sum_{l=1}^{k-1} I_l = V_{k-1}^I + 1 \right. \right) &= 0. \end{aligned}$$

   (b) If $V_{k-1}^F \geq V_k^F$, then we take the transition probabilities

$$\begin{aligned} Pr\left(I_k = 1 \left| \sum_{l=1}^{k-1} I_l = V_{k-1}^I \right. \right) &= 1, \\ Pr\left(I_k = 1 \left| \sum_{l=1}^{k-1} I_l = V_{k-1}^I + 1 \right. \right) &= \frac{V_k^F}{V_{k-1}^F}. \end{aligned}$$

Proposition 1. *For any $k = 1, \ldots, N$, we have*

$$V_k^I \leq \sum_{l=1}^{k} I_l \leq V_k^I + 1. \tag{3.2}$$

A drawback of the method is that, by construction, two non cross-border units inside the same microstratum U_i may not be selected jointly in the sample. Therefore, many second-order inclusion probabilities are equal to 0 and the variance of the Horvitz-Thompson estimator may not be unbiasedly estimated.

For this reason, Chromy (1979) proposed to use a randomized procedure, which is as follows. The population U is viewed as a closed loop, and we consider the set Σ_c of the N possible circular permutations, each of which using a different unit as the first one. The first permutation is σ_1, with the natural order $1, \ldots, N$. For $k = 2, \ldots, N$ the k-th permutation is σ_k where the units are in the order $k, \ldots, N, 1, \ldots, k-1$. The randomized Chromy sample S_{rc} is selected as

$$S_{rc} \sim Chr(\pi^{\sigma_k}; U^{\sigma_k}), \tag{3.3}$$

with σ_k a random permutation selected in Σ_c with probability π_k/n, with U^{σ_k} the population ordered with respect to σ_k, and π^{σ_k} the vector of probabilities ordered accordingly. This is the algorithm currently implemented in the `SURVEYSELECT` procedure of the `SAS` software.

4 Ordered Pivotal sampling

Ordered pivotal sampling (Fuller, 1970; Deville and Tillé, 1998; Chauvet, 2012) is presented in Algorithm 2. This is a succession of duels between units, and at each step the two first units remaining in the population are considered. If the sum of their probabilities is lower than 1 (rejection step), one of the unit is randomly discarded while the other gets the sum of their probabilities. If the sum of their probabilities is greater than 1 (selection step), one of the unit is randomly selected while the other goes on with the residual probability. For illustration, the complete probability tree for ordered pivotal sampling on a small population is given in Appendix A.

The pivotal sample is selected in at most $N - 1$ steps. Pivotal sampling is a

Algorithm 2 Ordered pivotal sampling with parameter π in the population U

1. Initialize with $\pi(0) = \pi_N$.
2. At step $t = 1, \ldots, T$:
   (a) Initialize with $\pi(t) = \pi(t-1)$.
   (b) Take $k < l$ the two first units in the population such that
   $$\pi_k(t-1) \notin \{0,1\} \quad \text{and} \quad \pi_l(t-1) \notin \{0,1\}$$
   (c) If $\pi_k(t-1) + \pi_l(t-1) \leq 1$ (rejection step), then do:
   $$\begin{aligned} \{\pi_k(t), \pi_l(t)\} &= \begin{cases} \{\pi_k(t-1) + \pi_l(t-1), 0\} & \text{with prob. } p(t) \\ \{0, \pi_k(t-1) + \pi_l(t-1)\} & \text{with prob. } 1 - p(t), \end{cases} \\ \text{where } p(t) &= \frac{\pi_k(t-1)}{\pi_k(t-1) + \pi_l(t-1)}. \end{aligned}$$
   (d) If $\pi_k(t-1) + \pi_l(t-1) > 1$ (selection step), then do:
   $$\begin{aligned} \{\pi_k(t), \pi_l(t)\} &= \begin{cases} \{1, \pi_k(t-1) + \pi_l(t-1) - 1\} & \text{with prob. } p(t) \\ \{\pi_k(t-1) + \pi_l(t-1) - 1, 1\} & \text{with prob. } 1 - p(t) \end{cases} \\ \text{where } p(t) &= \frac{1 - \pi_l(t-1)}{2 - \pi_k(t-1) - \pi_l(t-1)}. \end{aligned}$$
3. The algorithm stops at step T when all the components of $\pi(T)$ are 0 or 1. Take $I = \pi(T)$.

particular case of the cube method (Deville and Tillé, 2004), which enables to perform balanced sampling, i.e. to select samples such that the Horvitz-Thompson estimator exactly matches the known totals for some auxiliary variables. Pivotal sampling has found uses in spatial sampling, since it enables to spread well the sample over space: see for example Grafström et al. (2012) for the so-called local pivotal method, Chauvet and Le Gleut (2019) for the so-called pivotal tesselation method, or Benedetti et al. (2017) for a recent review on spatial sampling methods. Pivotal sampling is also of use in Monte Carlo methods (Gerber et al., 2019).

5 Properties of Chromy sampling

We first prove in Theorem 1 that Chromy sampling and ordered pivotal sampling are equivalent, which is the main result of the paper. The proof is lengthy, and given in Appendix C.

Theorem 1. *Ordered pivotal sampling and Chromy sampling with the same parameter π induce the same sampling design.*

By using the characterization of Chromy sampling given in Theorem 1, the mean-square consistency of the Horvitz-Thompson estimator stated in equation (5.1) of Theorem 2 is a direct consequence of Theorem 2 in Chauvet (2017). The central-limit theorem stated in equation (5.2) is a direct consequence of Theorem 1 in Chauvet and Le Gleut (2019).

Theorem 2. *Suppose that the sample S_c is selected by means of $Chr(\pi; U)$. If assumption (H2) holds, then*

$$E\left\{N^{-1}(\hat{t}_{y\pi} - t_y)\right\} \;=\; O(n^{-1}). \tag{5.1}$$

If in addition assumptions (H1) and (H3) hold, then

$$\frac{\hat{t}_{y\pi} - t_y}{\sqrt{V(\hat{t}_{y\pi})}} \;\underset{\mathcal{L}}{\longrightarrow}\; \mathcal{N}(0,1), \tag{5.2}$$

where $\underset{\mathcal{L}}{\longrightarrow}$ stands for the convergence in distribution.

It also follows from Theorem 1 that Chromy sampling is a negatively associated sampling design (Joag-Dev et al., 1983). This implies that the

Sen-Yates-Grundy conditions are satisfied. From Theorem 2 in Bertail and Clémençon (2019), the Horvitz-Thompson also satisfies a Bennett/Bernstein-type exponential inequality.

From Theorem 5.1 in Chauvet (2012), and from the computation given in Deville (1998), it is possible to give an explicit expression for the second-order inclusion probabilities under Chromy sampling. This is the purpose of Theorem 3.

Theorem 3. *Let k and l be two distinct units in U. If k and l are two non cross-border units that belong to the same microstratum U_i, then*

$$\pi_{kl} = 0,$$

if k and l are two non cross-border units that belong to distinct microstrata U_i and U_j, respectively, where $i < j$, then

$$\pi_{kl} = \pi_k \pi_l \left\{1 - c(i,j)\right\},$$

if $k = k_{i-1}$ and l is a non cross-border unit that belongs to the microstratum U_j where $i \leq j$, then

$$\pi_{kl} = \pi_k \pi_l \left[1 - b_{i-1}(1-\pi_k)\left\{\pi_k(1-b_{i-1})\right\}^{-1} c(i,j)\right],$$

if $l = k_{j-1}$ and k is a non cross-border unit that belongs to the microstratum U_i where $i < j$, then

$$\pi_{kl} = \pi_k \pi_l \left\{1 - (1-\pi_l)(1-b_{j-1})(\pi_l b_{j-1})^{-1} c(i,j)\right\},$$

if $k = p_{i-1}$ and $l = p_{j-1}$, where $i < j$, then

$$\pi_{kl} = \pi_k \pi_l \left[1 - b_{i-1}(1-b_{j-1})(1-\pi_k)(1-\pi_l)\left\{\pi_k \pi_l b_{j-1}(1-b_{i-1})\right\}^{-1} c(i,j)\right],$$

where $c(i,j) = \prod_{l=i}^{j-1} c_l$, $c_l = a_l b_l \left\{(1-a_l)(1-b_l)\right\}^{-1}$ and with $c(i,i) = 1$.

It is clear from Theorem 3 that many second-order inclusion probabilities are equal to zero for Chromy sampling. It also makes possible to compute the second-order inclusion probabilities for randomized Chromy sampling. For any units $k \neq l \in U$, let us denote π_{kl}^{rc} their second-order inclusion probability

under randomized Chromy sampling. Then:

$$\pi_{kl}^{rc} \;=\; \sum_{i \in U} \frac{\pi_i}{n} \pi_{kl}^{\sigma_i}, \tag{5.3}$$

with $\pi_{kl}^{\sigma_i}$ the joint selection probabilities of units k and l with the permutation σ_i, i.e. when Chromy sampling is applied to the population U^{σ_i} with parameter π^{σ_i}. A `SAS IML` subroutine to compute the second-order inclusion probabilities in (5.3) is available as Supplementary Material.

We evaluate in the simulation study performed in Section 6 a variance estimator making use of second-order inclusion probabilities computed from equation (5.3). For illustration, we give in this Section a small example. We consider a population U of size $N = 8$, with the parameter

$$\pi = (0.2, 0.4, 0.7, 0.4, 0.6, 0.6, 0.3, 0.8)^\top,$$

which leads to a sample of size $n = 4$. From equation (5.3), we obtain the following matrix of second-order inclusion probabilities (rounded to three decimal places):

$$(\pi_{kl}^{rc}) \;=\; \begin{pmatrix} 0.200 & 0.041 & 0.133 & 0.075 & 0.116 & 0.108 & 0.046 & 0.081 \\ & 0.400 & 0.171 & 0.142 & 0.224 & 0.227 & 0.099 & 0.297 \\ & & 0.700 & 0.209 & 0.410 & 0.415 & 0.207 & 0.555 \\ & & & 0.400 & 0.118 & 0.224 & 0.113 & 0.319 \\ & & & & 0.600 & 0.293 & 0.165 & 0.474 \\ & & & & & 0.600 & 0.065 & 0.469 \\ & & & & & & 0.300 & 0.205 \\ & & & & & & & 0.800 \end{pmatrix}.$$

We also selected 10^6 samples by means of the `SURVEYSELECT` procedure with the option `METHOD=PPS_SEQ`, which leads to randomized Chromy sampling. These 10^6 samples are used to obtain a simulation-based approximation of the matrix of second-order inclusion probabilities, which is given below (rounded

to three decimal places):

$$(\pi^{rc}_{kl,sim}) \;=\; \begin{pmatrix} 0.200 & 0.041 & 0.133 & 0.075 & 0.116 & 0.108 & 0.046 & 0.081 \\ & 0.400 & 0.171 & 0.142 & 0.223 & 0.227 & 0.099 & 0.296 \\ & & 0.701 & 0.210 & 0.410 & 0.416 & 0.208 & 0.556 \\ & & & 0.400 & 0.118 & 0.225 & 0.113 & 0.318 \\ & & & & 0.600 & 0.293 & 0.165 & 0.474 \\ & & & & & 0.600 & 0.065 & 0.469 \\ & & & & & & 0.300 & 0.205 \\ & & & & & & & 0.800 \end{pmatrix}.$$

It is clear that both matrices are almost identical.

6 Simulation study

We conducted a simulation study in order to evaluate variance estimation and interval estimation for randomized Chromy sampling. The set-up is inspired from Chauvet et al. (2017). We generate 2 populations of size $N = 500$, each consisting of an auxiliary variable x and 4 variables of interest $y_1, \ldots, y_4$. In the first population, the x-values are generated according to a Gamma distribution with shape and scale parameters 2 and 2; in the second population, the x-values are generated from a log-normal distribution with parameters 0 and 1.7. The x-values are then shaped and scaled to lie between 1 and 10.

Given the x-values, the values of the variables of interest are generated according to the following models:

$$\begin{aligned} \texttt{linear}: y_{1k} &= \alpha_{10} + \alpha_{11}(x_k - \mu_x) + \sigma_1\ \epsilon_k, \\ \texttt{quadratic}: y_{2k} &= \alpha_{20} + \alpha_{21}(x_k - \mu_x)^2 + \sigma_2\ \epsilon_k, \\ \texttt{exponential}: y_{3k} &= \exp\{\alpha_{30} + \alpha_{31}(x_k - \mu_x)\} + \sigma_3\ \epsilon_k, \\ \texttt{bump}: y_{4k} &= \alpha_{40} + \alpha_{41}(x_k - \mu_x)^2 - \alpha_{42}\exp\left\{-\alpha_{43}(x_k - \mu_x)^2\right\} + \sigma_4\ \epsilon_k, \end{aligned} \tag{6.1}$$

where μ_x is the population mean of x, and where ϵ_k follows a standard normal distribution $\mathcal{N}(0, 1)$. The population mean μ_y and the population dispersion S^2_y for the two populations and the four variables of interest are given in Table 1.

In each population, we computed inclusion probabilities proportional to the

Table 1: Population mean and population dispersion for two populations and four variables of interest

	linear		quadratic		exponential		bump	
	μ_{y1}	S^2_{y1}	μ_{y2}	S^2_{y2}	μ_{y3}	S^2_{y3}	μ_{y4}	S^2_{y4}
Population 1	10.1	13.1	11.7	69.8	10.3	18.2	12.3	84.6
Population 2	10.1	5.3	8.8	8.3	9.9	4.4	5.2	18.0

x-values, according to the formula

$$\pi_k \quad = \quad n\frac{x_k}{\sum_{l\in U} x_l}, \tag{6.2}$$

with $n = 50, 100$ or 200. The range of inclusion probabilities is given in Table 2. In some cases, equation (6.2) leads to inclusion probabilities greater than 1 for some units. In such case, the corresponding units are selected with certainty ($\pi_k = 1$), and the other probabilities are recomputed. For the first population, 5 units are selected with certainty with $n = 200$. For the second population, 2 units are selected with certainty with $n = 100$ and 9 units are selected with certainty with $n = 100$.

Table 2: Range of inclusion probabilities proportional to x for two populations and three sample sizes

	$n = 50$		$n = 100$		$n = 200$	
	Min	Max	Min	Max	Min	Max
Population 1	0.030	0.305	0.060	0.609	0.122	1.000
Population 2	0.079	0.791	0.160	1.000	0.325	1.000

We select $B = 1,000$ samples by means of randomized Chromy sampling, using the `SURVEYSELECT` procedure. For each sample and each variable of interest, we compute the Horvitz-Thompson estimator $\hat{t}_{y\pi}$, and the Sen-Yates-Grundy variance estimator

$$\hat{V}(\hat{t}_{y\pi}) \quad = \quad \frac{1}{2}\sum_{k\neq l\in S_{rc}} \frac{\pi_k\pi_l - \pi^{rc}_{kl}}{\pi^{rc}_{kl}}\left(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l}\right)^2, \tag{6.3}$$

where the second-order inclusion probabilities are given by equation (5.3). To evaluate the properties of this variance estimator, we compute the relative bias

$$RB\{\hat{V}(\hat{t}_{y\pi})\} \quad = \quad 100 \times \frac{B^{-1}\sum_{b=1}^{B}\hat{V}_b(\hat{t}_{y\pi b}) - V(\hat{t}_{y\pi})}{V(\hat{t}_{y\pi})}, \tag{6.4}$$

where $\hat{V}_b(\hat{t}_{y\pi b})$ denotes the variance estimator in the b-th sample, and where $V(\hat{t}_{y\pi})$ is the exact variance, computed by using the exact second-order inclusion probabilities given in (5.3). As a measure of stability, we use the Relative Root Mean Square Error

$$RRMSE\{\hat{V}(\hat{t}_{y\pi})\} \quad = \quad 100 \times \frac{\left[B^{-1}\sum_{b=1}^{B}\left\{\hat{V}_b(\hat{t}_{y\pi b}) - V(\hat{t}_{y\pi})\right\}^2\right]^{1/2}}{V(\hat{t}_{y\pi})}.$$

Finally, we compute the error rate of the normality-based confidence intervals with nominal one-tailed error rate of 2.5 % in each tail.

The simulation results are given in Table 3. The Sen-Yates-Grundy variance estimator is almost unbiased in all cases considered, except for the first population with $n = 50$ where the variance estimator is slightly positively biased. The Relative Root Mean Square Error diminishes as n increases, as expected. We note that the coverage rates are not well respected with $n = 50$, which is likely due to the small sample size and to the instability of the variance estimator. When the sample size increases, the coverage rates become close to the nominal level.

7 Conclusion

In this paper, we have studied Chromy's sampling algorithm. We proved that it is equivalent to ordered pivotal sampling, which enables in particular computing the second-order inclusion probabilities for the randomized Chromy algorithm programmed in the `SURVEYSELECT` procedure. The results in our simulation study confirm that the variance estimator based on the second-order probabilities computed from Deville's formulas show almost no bias for moderate sample sizes.

Table 3: Relative bias (in %), Relative Root Mean Square Error (in %) of the variance estimator and Coverage Rate of the normality-based confidence intervals for three populations and four variables of interest

		$n = 50$			$n = 100$			$n = 200$		
		RB	RMSE	Cov. Rate	RB	RMSE	Cov. Rate	RB	RMSE	Cov. Rate
Population 1	`linear`	-0.8	76	11.20	-0.5	46	7.20	-1.0	28	5.50
	`quadratic`	3.2	118	13.50	-0.6	66	8.60	-2.0	37	7.00
	`exponential`	-0.4	85	11.40	-0.9	50	7.10	-1.1	30	6.50
	`bump`	3.7	103	11.30	-0.4	59	8.30	-2.1	32	7.00
Population 2	`linear`	-0.7	56	8.20	0.6	31	7.30	0.0	17	4.80
	`quadratic`	-0.7	56	8.60	0.6	30	7.70	0.0	17	4.80
	`exponential`	-0.6	55	8.10	0.5	30	7.50	0.0	17	4.90
	`bump`	-1.4	61	11.10	0.8	35	6.90	0.9	19	4.30

The number of computations for the second-order inclusion probabilities of randomized Chromy sampling is of order N^3. Formula (5.3) is therefore tractable in case of a small population, for example when Chromy sampling is used to select a set of Primary Sampling Units in a multistage survey (e.g. Rust et al., 2019). Otherwise, we may resort to a simulation-based approximation of equation (5.3).

References

Benedetti, R., Piersimoni, F., and Postiglione, P. (2017). Spatially balanced sampling: a review and a reappraisal. *International Statistical Review*, 85(3):439–454.

Bertail, P. and Clémençon, S. (2019). Bernstein-type exponential inequalities in survey sampling: Conditional poisson sampling schemes. *Bernoulli*, 25(4B):3527–3554.

Chauvet, G. (2012). On a characterization of ordered pivotal sampling. *Bernoulli*, 18(4):1320–1340.

Chauvet, G. (2017). A comparison of pivotal sampling and unequal probability sampling with replacement. *Stat. Probabil. Lett.*, 121:1–5.

Chauvet, G., Haziza, D., and Lesage, É. (2017). Examining some aspects of balanced sampling in surveys. *Statistica Sinica*, 27:313–334.

Chauvet, G. and Le Gleut, R. (2019). Asymptotic properties of pivotal sampling with application to spatial sampling. preprint.

Chromy, J. R. (1979). Sequential sample selection methods. In *Proceedings of the Survey Research Methods Section of the American Statistical Association*, pages 401–406.

Deville, J.-C. (1998). *Une nouvelle (encore une!) méthode de tirage à probabilités inégales*. INSEE.

Deville, J.-C. and Tillé, Y. (1998). Unequal probability sampling without replacement through a splitting method. *Biometrika*, 85(1):89–101.

Deville, J.-C. and Tillé, Y. (2004). Efficient balanced sampling: the cube method. *Biometrika*, 91(4):893–912.

Fuller, W. A. (1970). Sampling with random stratum boundaries. *Journal of the Royal Statistical Society: Series B*, 32(2):209–226.

Gerber, M., Chopin, N., and Whiteley, N. (2019). Negative association, ordering and convergence of resampling methods. *Ann. Statist.*, 47(4):2236–2260.

Grafström, A., Lundström, N. L., and Schelin, L. (2012). Spatially balanced sampling through the pivotal method. *Biometrics*, 68(2):514–520.

Joag-Dev, K., Proschan, F., et al. (1983). Negative association of random variables with applications. *The Annals of Statistics*, 11(1):286–295.

Mills, S. D., Henriksen, L., Golden, S. D., Kurtzman, R., Kong, A. Y., Queen, T. L., and Ribisl, K. M. (2018). Disparities in retail marketing for menthol cigarettes in the united states, 2015. *Health & place*, 53:62–70.

Radwin, D., Conzelmann, J. G., Nunnery, A., Lacy, T. A., Wu, J., Lew, S., Wine, J., and Siegel, P. (2018). 2015-16 national postsecondary student aid study (npsas: 16): Student financial aid estimates for 2015-16. first look. nces 2018-466. *National Center for Education Statistics*.

Rust, S. M., Myers, A. E., D'Angelo, H., Queen, T. L., Laska, M. N., and Ribisl, K. M. (2019). Tobacco marketing at snap- and wic-authorized retail food stores in the united states. *Health Education & Behavior*, 46(4):541–549.

Särndal, C.-E., Swensson, B., and Wretman, J. (1992). *Model assisted survey sampling*. Springer Series in Statistics.

Schmitt, C. L., Curry, L., Boudewyns, V., Williams, P. A., Glasgow, L., Van Hersh, D., Willett, J., and Rogers, T. (2018). Peer reviewed: Relationships between theoretically derived short-term outcomes and support for policy among the public and decision-makers. *Preventing chronic disease*, 15.

A Chromy sampling and pivotal sampling on an example

We consider the population $U = \{1, 2, 3, 4, 5\}$ with $\pi = (0.4, 0.8, 0.5, 0.6, 0.7)$. The complete probability tree for Chromy sampling with parameter π is given in Figure 2. For example, at the first step the unit 1 is selected with probability 0.4, and discarded with probability 0.6. If unit 1 is selected, then $\sum_{l=1}^{1} I_l = V_1^I + 1$. Since unit 2 is a cross-border unit, we follow Step 2.b of Algorithm 1 and unit 2 is selected at the next step with probability $V_2^F / V_1^F = 0.2/0.4 = 1/2$.

The complete sampling design is

$$S_c = \begin{cases} \{1,2,4\} & \text{with proba.} \quad 3/35, \\ \{1,2,5\} & \text{with proba.} \quad 4/35, \\ \{1,3,4\} & \text{with proba.} \quad 3/56, \\ \{1,3,5\} & \text{with proba.} \quad 1/14, \\ \{1,4,5\} & \text{with proba.} \quad 3/40, \\ \{2,3,4\} & \text{with proba.} \quad 9/56, \\ \{2,3,5\} & \text{with proba.} \quad 3/14, \\ \{2,4,5\} & \text{with proba.} \quad 9/40. \end{cases} \tag{A.1}$$

Now, we consider ordered pivotal sampling on the same population U with the same parameter π. The complete probability tree for pivotal sampling is given in Figure 3. For example, at the first step, the units 1 and 2 fight with respective probabilities 0.4 and 0.8. With probability $(1 - 0.8)/(2 - 0.4 - 0.8) = 1/4$, unit 1 is selected and unit 2 gets the residual probability 0.2, and with the complementary probability unit 2 is selected and unit 1 gets the residual probability 0.2. In the first case, unit 2 faces unit 3 with respective probabilities 0.2 and 0.5. With probability $0.2/(0.2 + 0.5) = 2/7$, unit 2 gets the sum of the probabilities and unit 3 is discarded, and with the complementary probability unit 3 gets the sum of the probabilities and unit 2 is discarded.

It follows from straightforward computations that the complete sampling design is the same as for Chromy sampling.

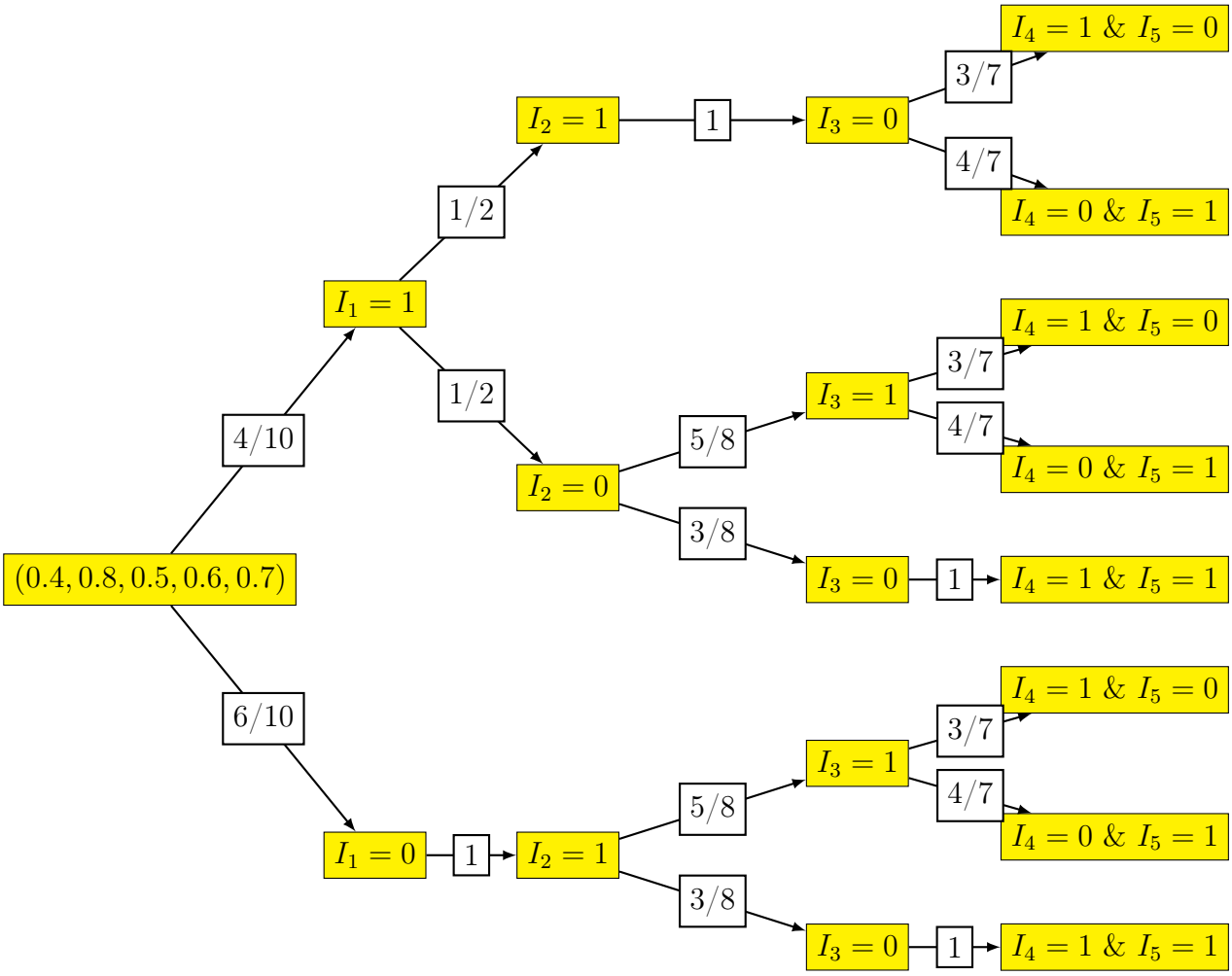

Figure 2: Probability tree for Chromy sampling on a population U of size $N = 5$

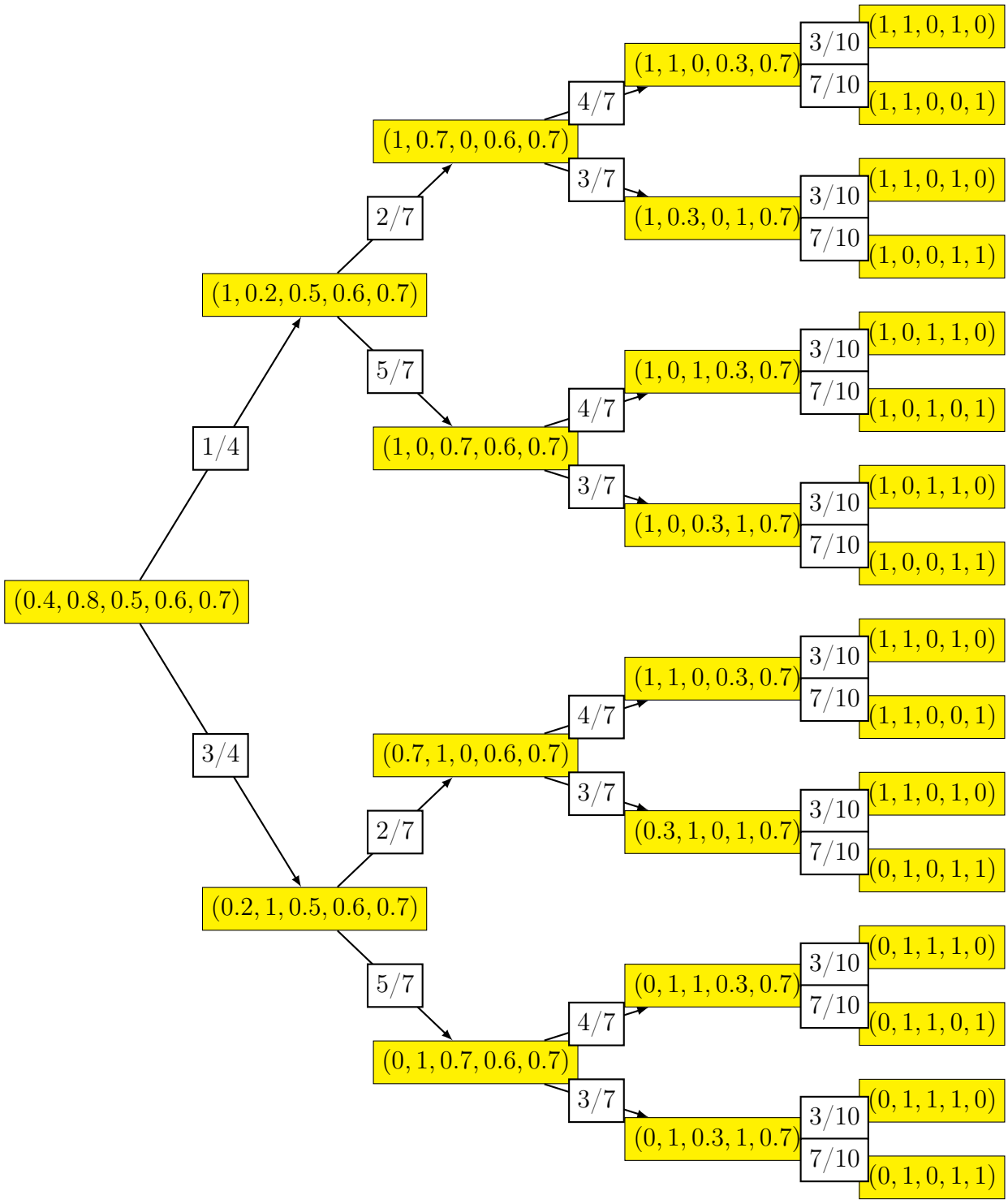

Figure 3: Probability tree for ordered pivotal sampling on a clustered population U_c

B Proof of Proposition 1

The proof is by induction. For $k = 1$, we have $I_1 \in \{0, 1\}$ and $V_1^I = 0$, so the property holds. Suppose that the property holds at $k - 1$, namely

$$V_{k-1}^I \leq \sum_{l=1}^{k-1} I_l \leq V_{k-1}^I + 1. \tag{B.1}$$

First note that if k is a cross-border unit, we have $V_k^I = V_{k-1}^I + 1$ and if not, we have $V_k^I = V_{k-1}^I$.

- Suppose that $\sum_{l=1}^{k-1} I_l = V_{k-1}^I$. If k is a non cross-border unit, we have

$$V_{k-1}^I \leq \sum_{l=1}^{k-1} I_l + I_k \leq V_{k-1}^I + 1 \quad \Leftrightarrow \quad V_k^I \leq \sum_{l=1}^{k} I_l \leq V_k^I + 1.$$

  If k is a cross-border unit, we obtain from Algorithm 1 that $I_k = 1$, and $\sum_{l=1}^{k} I_l = V_{k-1}^I + 1 = V_k^I$.

- Suppose that $\sum_{l=1}^{k-1} I_l = V_{k-1}^I + 1$. If k is a non cross-border unit, then from Algorithm 1, we have $I_k = 0$ and $\sum_{l=1}^{k} I_l = V_{k-1}^I + 1 = V_k^I$. If k is a cross-border unit, we obtain from $I_k \in \{0, 1\}$

$$V_{k-1}^I + 1 \leq \sum_{l=1}^{k-1} I_l + I_k \leq V_{k-1}^I + 2 \quad \Leftrightarrow \quad V_k^I \leq \sum_{l=1}^{k} I_l \leq V_k^I + 1.$$

C Proof of Theorem 1

The proof proceeds in two main steps. We first prove that Chromy sampling may be alternatively seen as the result of a two-stage sampling procedure, inside a population U_c of clusters which is introduced in Section C.1. We then consider in Section C.2 the ordered sample, which is given by the selected units ranked with respect to the natural order in the population, and we give the transition probabilities between the selected units. These results are used in Section C.3 to prove Theorem 1.

C.1 Clustered population

The N sampling units in the population U are grouped to obtain a population $U_c = \{u_1, \ldots, u_{2n-1}\}$ of clusters. There are the clusters associated to the cross-border units ($n-1$ singletons), denoted as u_{2i} with associated probability $\phi_{2i} = \pi_{k_i}$. There are the n clusters of non cross-border units that are between two consecutive integers, denoted as u_{2i-1} with associated probability $\phi_{2i-1} = V_{k_i-1} - V_{k_{i-1}}$. The vector of inclusion probabilities in the population U_c is denoted as $\phi = (\phi_1, \ldots, \phi_{2n-1})'$. For illustration, useful quantities for population U_c are presented in Figure 4.

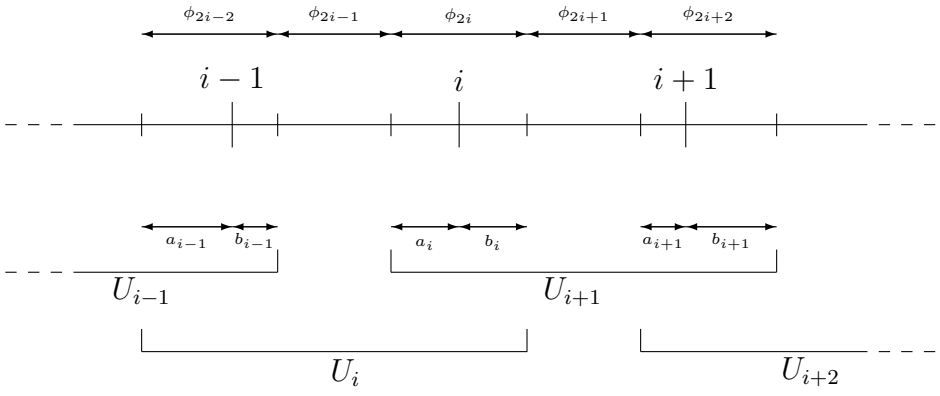

Figure 4: Inclusion probabilities and cross-border units in microstrata U_i and U_{i+1} for population U_c

Proposition 2. *Chromy sampling with parameter π in U may be performed by two-stage sampling, with:*

1. *a first-stage selection of a sample S_c of n clusters by means of Chromy sampling with parameter ϕ in the population U_c,*
2. *an independent second-stage selection inside each $u_i \in S_c$ of a sample S_i of size 1, with unit $k \in u_i$ selected with a probability π_k/ϕ_i.*

Proof. It is sufficient to prove that the transition probabilities given in Algorithm 1 are the same under Chromy sampling with parameter π and under the two-stage sampling procedure. We use the following notation: for any unit $k \in U$, recall that I_k is the sample membership indicator under Chromy sampling with parameter π; for any cluster $u_i \in U_c$, $J_{c,i}$ is the sample membership indicator under Chromy sampling with parameter ϕ; for any unit $k \in U$, J_k is the sample membership indicator under the two-stage procedure. We first note that, by definition of the two-stage procedure, we have

$$J_{c,i} \;=\; \sum_{k \in u_i} J_k \text{ for any } u_i \in U_c. \tag{C.1}$$

We now consider the case when $k \in U$ is a cross-border unit, k_i say. The corresponding cluster is u_{2i}, and in such case

$$J_{k_i} = 1 \;\Leftrightarrow\; J_{c,2i} = 1. \tag{C.2}$$

We obtain successively

$$\begin{aligned}
Pr\left(J_{k_i} = 1 \left| \sum_{l=1}^{k_i-1} J_l \right.\right) &= Pr\left(J_{k_i} = 1 \left| \sum_{j=1}^{2i-1} J_{c,j} \right.\right) \text{ from equation (C.1)} \\
&= Pr\left(J_{c,2i} = 1 \left| \sum_{j=1}^{2i-1} J_{c,j} \right.\right) \text{ from equation (C.2)} \\
&= \begin{cases} 1 & \text{if } \sum_{j=1}^{2i-1} J_{c,j} = i-1, \\ \frac{b_i}{1-a_i} & \text{if } \sum_{j=1}^{2i-1} J_{c,j} = i, \end{cases} \\
&\quad \text{from Step 2.b in Algorithm 1} \\
&= Pr\left(I_{k_i} = 1 \left| \sum_{l=1}^{k_i-1} I_k \right.\right), \qquad (C.3)
\end{aligned}$$

where the last line is obtained again from Algorithm 1.

We now consider the case when k is not a cross-border unit, and belongs to the cluster u_{2i-1}, say. We begin by computing the quantities

$$Pr\left(J_k = 1 \left| \sum_{j=1}^{2i-2} J_{c,j}, \sum_{l=k_{i-1}+1}^{k-1} J_l \right.\right).$$

Note that from Proposition 1, $\sum_{j=1}^{2i-2} J_{c,j}$ may only take the values $i-1$ and i, and since we select at most one unit l inside u_{2i-1}, $\sum_{l=k_{i-1}+1}^{k-1} J_l$ may only take the values 0 and 1. If $\sum_{l=k_{i-1}+1}^{k-1} J_l = 1$, we have $J_k = 0$ since we select at most one unit l inside u_{2i-1}. Therefore

$$Pr\left(J_k = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1, \sum_{l=k_{i-1}+1}^{k-1} J_l = 1 \right.\right) \;=\; 0. \tag{C.4}$$

If $\sum_{j=1}^{2i-2} J_{c,j} = i$, Algorithm 1 implies that the cluster u_{2i-1} and therefore k may not be selected. Therefore

$$Pr\left(J_k = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i, \sum_{l=k_{i-1}+1}^{k-1} J_l = 0 \right.\right) \;=\; 0. \tag{C.5}$$

For the same reason, we may not have simultaneously $\sum_{j=1}^{2i-2} J_{c,j} = i$ and $\sum_{l=k_{i-1}+1}^{k-1} J_l = 1$. Finally, we have

$$\begin{aligned} & Pr\left(J_k = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1, \sum_{l=k_{i-1}+1}^{k-1} J_l = 0 \right.\right) \\ = \;& \frac{Pr\left(J_k = 1, \sum_{l=k_{i-1}+1}^{k-1} J_l = 0 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1 \right.\right)}{Pr\left(\sum_{l=k_{i-1}+1}^{k-1} J_l = 0 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1 \right.\right)} \\ = \;& \frac{Pr\left(J_k = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1 \right.\right)}{Pr\left(\sum_{l=k_{i-1}+1}^{k-1} J_l = 0 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1 \right.\right)}, \end{aligned} \tag{C.6}$$

where the last line in (C.6) follows from the fact that if $J_k = 1$, we necessarily

have $\sum_{l=k_{i-1}+1}^{k-1} J_l = 0$. We compute the numerator and the denominator in (C.6) separately. The numerator is

$$\begin{aligned} & Pr\left(J_k = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1 \right.\right) \\ = \; & Pr\left(J_{c,2i-1} = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1 \right.\right) Pr\left(J_k = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1, J_{c,2i-1} = 1 \right.\right) \\ = \; & Pr\left(J_{c,2i-1} = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1 \right.\right) Pr\left(J_k = 1 \left| J_{c,2i-1} = 1 \right.\right). \end{aligned}$$

From Algorithm 1, we have $Pr\left(J_{c,2i-1} = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1 \right.\right) = \frac{1-b_{i-1}-a_i}{1-b_{i-1}}$, and from the definition of the two-stage procedure $Pr\left(J_k = 1 \left| J_{c,2i-1} = 1 \right.\right) = \frac{\pi_k}{1-b_{i-1}-a_i}$. This leads to

$$Pr\left(J_k = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1 \right.\right) = \frac{\pi_k}{1-b_{i-1}}. \tag{C.7}$$

From the definition of the two-stage procedure, the denominator in (C.6) is

$$\begin{aligned} Pr\left(\sum_{l=k_{i-1}+1}^{k-1} J_l = 0 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1 \right.\right) &= 1 - Pr\left(\sum_{l=k_{i-1}+1}^{k-1} J_l = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1 \right.\right) \\ &= 1 - \frac{\sum_{l=k_{i-1}+1}^{k-1} \pi_l}{1-b_{i-1}}. \end{aligned} \tag{C.8}$$

From (C.6), (C.7) and (C.8), we obtain

$$\begin{aligned} Pr\left(J_k = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} = i-1, \sum_{l=k_{i-1}+1}^{k-1} J_l = 0 \right.\right) &= \frac{\pi_k}{1-b_{i-1}-\sum_{l<k\in u_{2i-1}} \pi_l} \\ &= \frac{V_k^F - V_{k-1}^F}{1-V_{k-1}^F}. \end{aligned} \tag{C.9}$$

From (C.4), (C.5) and (C.9), we obtain

$$\begin{aligned} Pr\left(J_k = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} + \sum_{l=k_{i-1}+1}^{k-1} J_l = i-1 \right.\right) &= \frac{V_k^F - V_{k-1}^F}{1 - V_{k-1}^F}, \\ Pr\left(J_k = 1 \left| \sum_{j=1}^{2i-2} J_{c,j} + \sum_{l=k_{i-1}+1}^{k-1} J_l = i \right.\right) &= 0. \end{aligned} \tag{C.10}$$

From equation (C.1), we have

$$\sum_{j=1}^{2i-2} J_{c,j} + \sum_{l=k_{i-1}+1}^{k-1} J_l = \sum_{l=1}^{k_{i-1}} J_l + \sum_{l=k_{i-1}+1}^{k-1} J_l = \sum_{l=1}^{k-1} J_l, \tag{C.11}$$

and from (C.10), this implies that

$$\begin{aligned} Pr\left(J_k = 1 \left| \sum_{l=1}^{k-1} J_l \right.\right) &= \begin{cases} \frac{V_k^F - V_{k-1}^F}{1 - V_{k-1}^F} & \text{if } \sum_{l=1}^{k-1} J_l = i-1, \\ 0 & \text{if } \sum_{l=1}^{k-1} J_l = i, \end{cases} \\ &= Pr\left(I_k = 1 \left| \sum_{l=1}^{k-1} I_l \right.\right), \end{aligned} \tag{C.12}$$

where the last line in (C.12) follows from Step 2.b in Algorithm 1. This completes the proof. □

C.2 Ordered sample

We use the same notation as in Section C.1, and consider a sample S_c selected in U_c by means of Chromy sampling with parameter ϕ. In this case, we let $V_{c,i} = \sum_{j=1}^{i} \phi_j$ denote the cumulated inclusion probabilities up to unit u_i, with $V_{c,0} = 0$. The integer part of $V_{c,i}$ is denoted as $V_{c,i}^I$, and the difference between $V_{c,i}$ and its integer part $V_{c,i}^I$ is the fractional part, denoted as $V_{c,i}^F$.

Let us denote by $X_1 < \ldots < X_n$ the selected units, ranked with respect to the natural order in the population U. The transition probabilities between the ranked selected units are given in Proposition 3.

Proposition 3. *Let S_c denote a sample selected by Chromy sampling with parameter ϕ in the population U_c. The transition probabilities between the*

ordered sampled units $X_1 < \ldots < X_n$ *are:*

$$Pr(X_{i+1} = u_j | X_i = u_{2i-2}) = \begin{cases} \frac{b_i}{1-a_i} & \text{if } j = 2i, \\ \frac{(1-b_i-a_{i+1})(1-a_i-b_i)}{(1-a_i)(1-b_i)} & \text{if } j = 2i+1, \\ \frac{a_{i+1}(1-a_i-b_i)}{(1-a_i)(1-b_i)} & \text{if } j = 2i+2, \end{cases}$$

$$Pr(X_{i+1} = u_j | X_i = u_{2i-1}) = \begin{cases} \frac{b_i}{1-a_i} & \text{if } j = 2i, \\ \frac{(1-b_i-a_{i+1})(1-a_i-b_i)}{(1-a_i)(1-b_i)} & \text{if } j = 2i+1, \\ \frac{a_{i+1}(1-a_i-b_i)}{(1-a_i)(1-b_i)} & \text{if } j = 2i+2, \end{cases}$$

$$Pr(X_{i+1} = u_j | X_i = u_{2i}) = \begin{cases} \frac{1-b_i-a_{i+1}}{1-b_i} & \text{if } j = 2i+1, \\ \frac{a_{i+1}}{1-b_i} & \text{if } j = 2i+2. \end{cases}$$

Proof. We first consider the case when $X_i = u_{2i-2}$, which is equivalent to $\sum_{j=1}^{2i-2} J_{c,j} = i$. Therefore,

$$\begin{aligned} Pr(X_{i+1} = u_j | X_i = u_{2i-2}) &= Pr\left(X_{i+1} = u_j \left| \sum_{j=1}^{2i-2} J_{c,j} = i \right. \right) \\ &= Pr\left(X_{i+1} = u_j \left| \sum_{j=1}^{2i-1} J_{c,j} = i \right. \right), \quad \text{(C.13)} \end{aligned}$$

where the second line in (C.13) follows from the fact that, from Step 2.a of Algorithm 1, $\sum_{j=1}^{2i-2} J_{c,j} = i$ implies that $J_{c,2i-1} = 0$. If $j = 2i$:

$$\begin{aligned} Pr(X_{i+1} = u_{2i} | X_i = u_{2i-2}) &= Pr\left(X_{i+1} = u_{2i} \left| \sum_{j=1}^{2i-1} J_{c,j} = i \right. \right) \\ &= Pr\left(J_{c,2i} = 1 \left| \sum_{j=1}^{2i-1} J_{c,j} = i \right. \right) \\ &= \frac{V^F_{c,2i}}{V^F_{c,2i-1}} = \frac{b_i}{1-a_i}. \quad \text{(C.14)} \end{aligned}$$

If $j = 2i + 1$:

$$\begin{aligned} Pr(X_{i+1} = u_{2i+1} | X_i = u_{2i-2}) &= Pr\left(X_{i+1} = u_{2i+1} \left| \sum_{j=1}^{2i-1} J_{c,j} = i \right. \right) \\ &= Pr\left(J_{c,2i+1} = 1, J_{c,2i} = 0 \left| \sum_{j=1}^{2i-1} J_{c,j} = i \right. \right) \\ &= Pr\left(J_{c,2i+1} = 1 \left| \sum_{j=1}^{2i} J_{c,j} = i \right. \right) Pr\left(J_{c,2i} = 0 \left| \sum_{j=1}^{2i-1} J_{c,j} = i \right. \right) \\ &= \left(\frac{V_{c,2i+1}^F - V_{c,2i}^F}{1 - V_{c,2i}^F} \right) \left(1 - \frac{b_i}{1 - a_i} \right) \\ &= \frac{(1 - b_i - a_{i+1})(1 - a_i - b_i)}{(1 - a_i)(1 - b_i)}, \end{aligned} \tag{C.15}$$

where the last but one line in (C.15) follows from Step 2.a of Algorithm 1 and from equation (C.14). If $j = 2i + 2$, we obtain similarly:

$$\begin{aligned} Pr(X_{i+1} = u_{2i+2} | X_i = u_{2i-2}) &= Pr\left(X_{i+1} = u_{2i+2} \left| \sum_{j=1}^{2i-1} J_{c,j} = i \right. \right) \\ &= Pr\left(J_{c,2i+2} = 1, J_{c,2i+1} = 0, J_{c,2i} = 0 \left| \sum_{j=1}^{2i-1} J_{c,j} = i \right. \right) \\ &= Pr\left(J_{c,2i+2} = 1 \left| \sum_{j=1}^{2i+1} J_{c,j} = i \right. \right) Pr\left(J_{c,2i+1} = 0 \left| \sum_{j=1}^{2i} J_{c,j} = i \right. \right) \\ &\times Pr\left(J_{c,2i} = 0 \left| \sum_{j=1}^{2i-1} J_{c,j} = i \right. \right) \\ &= 1 \times \left(1 - \frac{1 - a_{i+1} - b_i}{1 - b_i} \right) \left(1 - \frac{b_i}{1 - a_i} \right) \\ &= \frac{a_{i+1}(1 - a_i - b_i)}{(1 - a_i)(1 - b_i)}. \end{aligned} \tag{C.16}$$

This gives the first equation in Proposition 3.

Now, we consider the case when $X_i = u_{2i-1}$. We have

$$\begin{aligned} X_i = u_{2i-1} &\Rightarrow \sum_{j=1}^{2i-2} J_{c,j} = i-1 \text{ and } J_{c,2i-1} = 1 \\ &\Rightarrow \sum_{j=1}^{2i-1} J_{c,j} = i. \end{aligned} \tag{C.17}$$

Since at each step of Chromy sampling, the conditional probabilities only depend on the number of units already selected, this leads to

$$Pr(X_{i+1} = u_j | X_i = u_{2i-1}) = Pr\left(X_{i+1} = u_j \left| \sum_{j=1}^{2i-1} J_{c,j} = i \right. \right), \tag{C.18}$$

which is identical to equation (C.13). Therefore, the second equation in Proposition 3 follows.

Finally, we consider the case when $X_i = u_{2i}$, which is equivalent to $\sum_{j=1}^{2i} J_{c,j} = i$. If $j = 2i+1$:

$$\begin{aligned} Pr(X_{i+1} = u_{2i+1} | X_i = u_{2i}) &= Pr\left(X_{i+1} = u_{2i+1} \left| \sum_{j=1}^{2i} J_{c,j} = i \right. \right) \\ &= Pr\left(J_{c,2i+1} = 1 \left| \sum_{j=1}^{2i} J_{c,j} = i \right. \right) \\ &= \frac{V^F_{c,2i+1} - V^F_{c,2i}}{1 - V^F_{c,2i}} = \frac{1 - a_{i+1} - b_i}{1 - b_i}. \end{aligned} \tag{C.19}$$

If $j = 2i + 2$:

$$\begin{aligned} Pr(X_{i+1} = u_{2i+2} | X_i = u_{2i}) &= Pr\left(X_{i+1} = u_{2i+2} \left| \sum_{j=1}^{2i} J_{c,j} = i \right. \right) \\ &= Pr\left(J_{c,2i+2} = 1, J_{c,2i+2} = 0 \left| \sum_{j=1}^{2i} J_{c,j} = i \right. \right) \\ &= Pr\left(J_{c,2i+2} = 1 \left| \sum_{j=1}^{2i+1} J_{c,j} = i \right. \right) Pr\left(J_{c,2i+1} = 0 \left| \sum_{j=1}^{2i} J_{c,j} = i \right. \right) \\ &= 1 \times \left(1 - \frac{1 - a_{i+1} - b_i}{1 - b_i} \right) = \frac{a_{i+1}}{1 - b_i}. \end{aligned} \tag{C.20}$$

This completes the proof. □

C.3 Proof of Theorem 1

From Lemma 3.1 in Chauvet (2012) and our Proposition 2, Chromy sampling and ordered pivotal sampling have the same two-stage characterization. It is therefore sufficient to prove that they lead to the same sampling design when sampling with parameter ϕ in the clustered population U_c. However, from equations (4.2)-(4.4) in Chauvet (2012) and our Proposition 3, both Chromy sampling and ordered pivotal sampling have the same transition probabilities between ordered sampled units, which means that the induced sampling designs are identical.