Targetted double control of burden in multiple surveys

Alina Matei[*], Paul A. Smith[†], Marc J.E. Smeets[‡], Jonas Klingwort[§]

Abstract

Sample coordination methods aim to increase (in positive coordination) or decrease (in negative coordination) the size of the overlap between samples. The samples considered can be from different occasions of a repeated survey and/or from different surveys covering a common population. Negative coordination is used to control the response burden in a given period, because some units do not respond to survey questionnaires if they are selected in many samples. Usually, methods for sample coordination do not take into account any measure of the response burden that a unit has already expended in responding to previous surveys. We introduce such a measure into a new method by adapting a spatially balanced sampling scheme, based on a generalization of Poisson sampling, together with a negative coordination method. The goal is to create a double control of the burden for these units: once by using a measure of burden during the sampling process and once by using a negative coordination method. We evaluate the approach using Monte-Carlo simulation and investigate its use for controlling for selection 'hot-spots' in business surveys in Statistics Netherlands.

Keywords: coordinated sampling, negative coordination, survey burden, burden hot-spots.

1 Introduction

Sample coordination methods seek to alter the size of the overlap(s) between two or more samples relative to the case where all the samples are selected independently. Positive coordination refers to the case where the overlap is larger than under independent sampling, and is generally used to reduce the variance of measures of change between successive periods of repeating surveys, though it can also be used to link together information from two separate surveys. Negative coordination is when the overlap is smaller than under independent sampling, and is used particularly to reduce the number of surveys in which a particular unit is selected in a given period, and therefore to control the perceived burden of responding (Bradburn, 1978). Bottone et al. (2021) have shown that increased perceived burden is associated with higher attrition and partial response, and also with lower data quality.

In fact the effect of negative coordination is to spread out a fixed overall burden across either or both of more units and more time, so the total burden is the same, but the risk of any particular unit having a large burden within a short period is reduced - ideally to zero, though this is not always possible in practice, because detailed stratification can result in some parts of the population being relatively heavily sampled. Negative coordination has been widely applied in business surveys where sampling fractions tend to be large, and we focus on examples from surveys of businesses and institutions, though the approach can be applied in various situations.

[*]Institute of Statistics, University of Neuchâtel, Switzerland. alina.matei@unine.ch
[†]S3RI & Dept. of Social Statistics and Demography, University of Southampton, UK. p.a.smith@soton.ac.uk
[‡]Statistics Netherlands, Netherlands. mje.smeets@cbs.nl
[§]Statistics Netherlands, Netherlands. j.klingwort@cbs.nl

There are many available methods for sample coordination, and Matei and Smith (2023) give an overview of them. An important approach in sample coordination is the use of permanent random numbers (PRNs). PRNs were introduced by Brewer et al. (1972) to coordinate Poisson samples, and have been widely used since as the basis of other, different methods (see Ohlsson, 1995, for an overview). They have also formed the basis of a number of sample coordination systems (which are also reviewed by Matei and Smith, 2023), which generally have to deal with a series of practical issues as well as following the theory of a particular sample coordination method. Coordination systems operate most flexibly where there are many units, and so benefit small units particularly. But they also operate to spread the response burden evenly for units of all sizes.

The consequence is that although a sample coordination system reduces the overlaps between samples and therefore the current burden, there are still some units which appear more frequently than others across the range of samples being coordinated. This is at least partly driven by the use of stratified designs with many small strata, some of which have large inclusion probabilities. Statistics Netherlands operates a sample coordination system called the Survey Burden System (Smeets and Boonstra, 2018) which uses a PRN coordination approach which takes account of the accumulated burden of the units. This system does indeed result in relatively only a few units being included in multiple samples, but these 'hot-spots' are a challenge because of the burden they represent for particular units and the consequent effects on response and relations with respondents.

A related situation is identified by Landry (2011) in Statistics Canada's Survey of Employment, Payroll and Hours (SEPH). Landry suggests the use of a take-none (cut-off) stratum, mainly to control the response burden on the smallest businesses. Although cut-off sampling is widely used, there is a risk of bias in estimates from such samples, so a better approach to burden control allowing unbiased estimation would be preferable.

In this paper we therefore seek a method which follows the requirements of negative sample coordination, but allows for some additional control for units which have particular characteristics. We adapt the approach of spatially correlated Poisson (SCP) sampling (Grafström, 2012) to this problem, by introducing a measure of the response burden in the sampling process.

In the remainder of the paper we introduce the framework and notation for sample coordination in Section 2, outline the procedure for spatially correlated Poisson sampling in Section 2.4 and develop the methodology for sample coordination with targetted double control in Section 3. We evaluate our proposal in Section 3.2; we provide Monte-Carlo simulation studies using the MU284 population from Särndal et al. (1992) in Section 3.3, and real data on the business population in the Netherlands to assess its ability to deal with hot-spots in Section 3.4. Section 4 concludes with a discussion.

2 Sample coordination and SCP sampling

2.1 Framework and notation

We consider the framework of two overlapping finite populations of units, denoted by U_1 and U_2. One selects samples s_1 from U_1 and s_2 from U_2, using the sampling designs p_1 and p_2, respectively. The set formed by the two samples can be seen as a bivariate sample $s = (s_1, s_2) \subseteq U_1 \times U_2$, having a joint sampling design p, with the marginals p_1 and p_2. The samples s_1 and s_2 are drawn dependently to alter the size of the overlap(s) between them relative to the case where the samples are selected independently. Thus, the samples s_1 and s_2 are said to be *coordinated* if

$$p(s_1, s_2) \neq p_1(s_1)p_2(s_2),$$

(see Cotton and Hesse, 1992; Mach et al., 2006).

The size of the *overlap* between s_1 and s_2, denoted by c, represents the number of units common to s_1 and s_2. It is in general a random variable having expectation

$$E(c) = \sum_{k \in U} \pi_{k,12},$$

with

$$\pi_{k,12} = P(k \in s_1, k \in s_2) = \sum_{s_1, k \in s_1} \sum_{s_2, k \in s_2} p(s_1, s_2), \forall k \in U,$$

and $U = U_1 \cup U_2$ is the so-called 'overall population'. Maximizing/minimizing $E(c)$ in positive/negative sample coordination represents an overall standard to evaluate a coordination method.

Let $\pi_{k1} = P(k \in s_1)$ and $\pi_{k2} = P(k \in s_2)$ be the first-order inclusion probabilities of unit $k \in U$ in the first and second sample, respectively. We consider that $\pi_{k1} = 0$ if $k \notin U_1$ (if s_1 and s_2 are samples for two periods of the same survey, these new units represent 'births') and $\pi_{k2} = 0$ if $k \notin U_2$ ('deaths').

Based on probability theory, the following bounds are available for the joint probability $\pi_{k,12}$, for any $k \in U$

$$ALB_k = \max(0, \pi_{k1} + \pi_{k2} - 1) \leq \pi_{k,12} \leq \min(\pi_{k1}, \pi_{k2}) = AUB_k, \quad (1)$$

where ALB stands for 'absolute lower bound' and AUB for 'absolute upper bound'. One obtains the lower and upper bounds for $E(c)$ by applying the sum over all units $k \in U$ on the left and right side of Expression (1) (see Matei and Tillé, 2005):

$$ALB = \sum_{k \in U} ALB_k \leq E(c) \leq \sum_{k \in U} AUB_k = AUB. \quad (2)$$

If s_1 is drawn before s_2, and s_2 is selected conditionally on s_1 using a probability $P(s_2|s_1)$, one can obtain any value of $\pi_{k,12} \in [ALB_k, AUB_k]$ using the conditional probabilities (Cotton and Hesse, 1992)

$$\begin{aligned} P(k \in s_2 | k \in s_1) &= \pi_{k,12}/\pi_{k1}, \\ P(k \in s_2 | k \notin s_1) &= (\pi_{k2} - \pi_{k,12})/(1 - \pi_{k1}). \end{aligned}$$

We focus below on the negative coordination of two samples. Reaching the lower bound in Expression (1) for all units $k \in U$ is equivalent to creating the 'best' possible degree of negative coordination between samples based on their overlap minimization.

2.2 Sample coordination with PRNs

As underlined in Section 1, sample coordination methods with PRNs are commonly used in practice. They are based on the following basic idea: one associates a uniform random number drawn independently from the $Unif(0, 1)$ distribution with each unit $k \in U$. These numbers are called 'permanent' since they are used in the selection process over time and over surveys for units which persist in the population. For a 'birth' (a new unit which appears in the population), a new PRN is assigned; for a 'death' (a unit which disappears from the population), the unit and its associated PRN are deleted from the corresponding survey frame.

Introduced by Brewer et al. (1972), Poisson sampling with PRNs is widely used in sample coordination, especially as a base for coordination systems (see Qualité, 2019, for an example applied in Switzerland). In negative coordination it is implemented as follows: first, one generates the PRNs $u_1, u_2, \ldots, u_N$ independently from the $Unif(0,1)$ distribution. Next, if $u_k < \pi_{k1}, k \in U$, unit k is included in s_1. The sample s_2 is selected in a similar manner, but using the numbers $1 - u_k$ instead of u_k: if $1 - u_k < \pi_{k2}$ then unit k is included in s_2. Using Poisson sampling with PRNs to negatively coordinate s_1 and s_2 allows the bound ALB_k in Expression (1) to be reached for any unit $k \in U$. This can be shown as follows: unit k is selected in both s_1 and s_2 if $u_k < \pi_{k1}$ and $1 - u_k < \pi_{2k}$. This is equivalent to $1 - \pi_{k2} < u_k < \pi_{k1}$. The probability that this occurs is $\max(0, \pi_{k1} - (1 - \pi_{k2}) = \max(0, \pi_{k1} + \pi_{k2} - 1) = ALB_k$.

While Poisson sampling with PRNs is a very attractive scheme for sample coordination, it has an important drawback: the resulting samples have random sizes, increasing the variance of the estimates. For this reason, fixed-size sampling designs are sometimes preferred; see, for instance, Pareto sampling with PRNs (Rosén, 1997a,b). Nevertheless, the bounds provided in Expression (1) are in general not reached using sampling designs with fixed sample size and unequal inclusion probabilities; for some empirical results based on Monte-Carlo simulation, see Grafström and Matei (2018). Compared to Poisson sampling with PRNs which allows independent selection of units, such sampling designs impose more restriction on the unit selection mainly due to the fixed sample size, and fail in general to reach the bounds provided in Expression (1). Theoretical conditions to reach the two overall bounds given in Expression (2) are provided by Matei and Tillé (2005).

2.3 Response burden and sample coordination

Response burden is a difficult concept to define; it may include objective factors such as the time spent to provide questionnaire responses and subjective factors such as what is perceived as burden by the respondents, see for instance Natkowska and Modak (2014); Bottone et al. (2021). From the statistical perspective, we use the following definition provided by Sunter (1977).

Consider several surveys $j = 1, 2, \ldots, M$ having associated populations of units $U_1, U_2, \ldots, U_M$, with $U = \cup_{j=1}^{M} U_j$. The response burden of unit k in M surveys is a random variable

$$RB_k = \sum_{j=1}^{M} \beta_j \times I_{kj},$$

with $I_{kj} = 1$ if $k \in s_j$ and 0 otherwise, and β_j is the response load imposed by the jth survey for all units selected to participate in this survey. The expected value of RB_k is given by $E(RB_k) = \sum_{j=1}^{M} \beta_j \pi_{kj}, k \in U$, where $\pi_{kj} = P(k \in s_j)$, and s_j is the jth survey sample, $j = 1, \ldots, M$. When $\beta_j = 1$, for all $j = 1, \ldots, M$, we simply obtain that

$$RB_k = \sum_{j=1}^{M} I_{kj}, \text{and } E(RB_k) = \sum_{j=1}^{M} \pi_{kj}, k \in U. \quad (3)$$

The response burden is commonly associated with a negative coordination of samples. Usually, a coordination method does not use any measure of response burden in the sampling process. If a negative coordination method is applied, the value of the response burden (seen as the realization of a random variable) diminishes for the units not included in the overlap: for instance, for two selected samples s_1, s_2, and considering the first part of Expression (3), the value of RB_k is 1 if $k \in s_1 \setminus s_2$ or $k \in s_2 \setminus s_1$ while it is 2 if $k \in s_1 \cap s_2$. Minimizing the sample overlap size implies having fewer units with the value of $R_k = 2$. However, at the overall level, sample coordination methods do not affect $E(RB_k)$, but try to control for excessive burdens and to allocate burdens in a fair way, as well as to reduce the variance of RB_k.

In what follows, we use the expression *cumulated response burden* to denote the realized value of the RB_k given in the first part of Expression (3).

2.4 Spatially correlated Poisson sampling

Spatially correlated Poisson (SCP) sampling is a particular case of correlated Poisson sampling, a family of sampling designs introduced by Bondesson and Thorburn (2008). First, we review this method for a generic sample with given first-order inclusion probabilities; next, we give the modification provided by Grafström (2012) which produces SCP samples.

Correlated Poisson sampling is a list sequential method used to draw a random sample $\widetilde{s}$ from U, with prescribed inclusion probabilities $\pi_k, k \in U$. Consider the selection probability $\pi_k^{(k-1)}$, that allows unit k to be selected in $\widetilde{s}$ at the kth iteration of the following algorithm (see Step 3b):

Step 1: set $\pi_\ell^{(0)} = \pi_\ell$, for all $\ell \in U$.

Step 2: set $k = 1$;

Step 3: while $k \leq N$ do

- 3a: generate $\widetilde{u}_k$ independently from the $Unif(0,1)$ distribution,
- 3b: if $\widetilde{u}_k < \pi_k^{(k-1)}$ then $I_k = 1$ (k is selected in $\widetilde{s}$) else $I_k = 0$ (k is not selected in $\widetilde{s}$), where I_k is the indicator variable associated to unit $k \in U$.
- 3c: update the selection probabilities for the remaining units $i = k+1, \dots, N$ according to
$$\pi_i^{(k)} = \pi_i^{(k-1)} - (I_k - \pi_k^{(k-1)}) w_k^{(i)},$$
where $w_k^{(i)}$ are some weights given by unit k to other units $i = k+1, \dots, N$.
- 3d: increment k.

The weights can be chosen freely in the following range

$$-\min\left(\frac{1-\pi_i^{(k-1)}}{1-\pi_k^{(k-1)}}, \frac{\pi_i^{(k-1)}}{\pi_k^{(k-1)}}\right) \leq w_k^{(i)} \leq \min\left(\frac{\pi_i^{(k-1)}}{1-\pi_k^{(k-1)}}, \frac{1-\pi_i^{(k-1)}}{\pi_k^{(k-1)}}\right), \tag{4}$$

in order to assure that $0 \leq \pi_i^{(k-1)} \leq 1$, $i = k+1, ..., N$. The weights $w_k^{(i)}$ may depend on $I_1, I_2, \dots, I_{k-1}$ but not on $I_k, I_{k+1}, \dots, I_N$. The choice of $w_k^{(i)}$ provides different sampling designs; for instance, Poisson sampling is obtained if all $w_k^{(i)}$ are zero, $k \in U$. Bondesson and Thorburn (2008) showed that a fixed size sampling is obtained if for each $k \in U : \sum_{i=k+1}^{N} w_k^{(i)} = 1$ and $\sum_{k \in U} \pi_k = n$, n being the sample size. The prescribed inclusion probabilities are respected since

$P(k \in \widetilde{s}) = \pi_k, k \in U$, regardless of the choice of $w_k^{(i)}$ (see Remark 1 in Bondesson and Thorburn, 2008). This is due to the fact that

$$E(\pi_k^{(k-1)}) = E(E(\pi_k^{(k-1)} \mid \pi_k^{(k-2)})) = E(\pi_k^{(k-2)}) = \cdots = \pi_k, \text{ for all } k = 1, 2, \ldots, N,$$

as underlined by Grafström (2012).

Grafström (2012) applied this method in spatial sampling, where the units in the population have associated geographical coordinates which can be used to compute distances between them. His goal was to draw a balanced sample, so that the selected units are spread over the space under study. To avoid clustering of similar units and to obtain well-spread samples, Grafström (2012) used positive weights $w_k^{(i)}$ in Expression (4) chosen such that unit k gives maximal weight to the unit closest to k in (Euclidean) distance, among the units $i = k+1, \ldots, N$, then as much weight as possible to the second closest unit, etc. with the restriction that $\sum_{i=k+1}^{N} w_k^{(i)} = 1$ is fulfilled, for any $k \in U$, and respecting the upper bound for each weight $w_k^{(i)}$ in Expression (4). This method, called the *maximal weight strategy*, provides spatially correlated Poisson sampling, which is a balanced spatial sampling design of fixed sample size, assuming that $\sum_{k \in U} \pi_k = n$. Moreover, if $\pi_k, k \in U$ are proportional to a size measure, SCP sampling becomes a πps sampling design of fixed sample size.

Grafström and Matei (2018) employed SCP sampling with PRNs to coordinate samples in a manner similar to Poisson sampling with PRNs (Brewer et al., 1972). Thus, for negative coordination, Step 3a of the previous algorithm is executed only once for s_1, in order to associate a PRN u_k with each unit $k \in U$, and $\widetilde{u}_k$ is replaced by u_k. Next, to select s_1 one uses u_k instead of $\widetilde{u}_k$ and the corresponding selection probabilities in Step 3a; to select s_2, one uses $1 - u_k$ instead of $\widetilde{u}_k$ and the corresponding selection probabilities in Step 3a. SCP sampling with PRNs is implemented in the function 'scps_coord' of the R package 'BalancedSampling' (Grafström et al., 2022).

3 Targetted double control strategy

3.1 Description of the strategy

Due to the cumulated response burden, some units do not answer the survey questionnaires if they are selected in many samples (Lorenc et al., 2013). We assume that, over time, some such units with a large cumulated burden become 'notorious' non-respondents. The same occurs with 'hot-spots' in business surveys in Statistics Netherlands (see Section 3.4). We denote these units as *non-desired units*, and we want to exclude them as much as possible from future selections, while respecting their prescribed inclusion probabilities. It is possible to classify the units of U into two categories: desired (usually with low cumulated response burden) and non-desired units (usually with large cumulated response burden).

Our goal is to produce a double control of the response burden for the non-desired units: first by using a measure of the response burden in the sampling process, and second by using a method for negative sample coordination.

To create a targetted double control of non-desired units, we modify spatially correlated Poisson sampling (Grafström, 2012), and use a negatively coordinated sampling scheme. We describe below how to adapt SCP sampling for the coordination framework with targetted double control.

As explained in Section 2.4, in spatial sampling, the units in U have associated geographical coordinates. It is thus possible to compute distances between units, usually using Euclidean distance.

Units close in distance provide in general similar information. Spatially balanced sampling allows the selection of units that are spread over the space, and thus avoids collecting similar information.

We propose to replace the matrix of geographical coordinates in SCP sampling by the vector formed by a measure based on the response burden of each unit. Other information, such as the inclusion probabilities, can also be included, and the vector becomes a matrix. This could lead to an extra spread with respect of the inclusion probabilities. Similar units usually have inclusion probabilities close to each other and non-desired units often have larger inclusion probabilities.

The Euclidean distances between units are next computed using this new vector or matrix. If a non-desired unit is selected in the current sample, SCP sampling avoids selecting a similar unit. Next, we use a negative coordination method for samples. We call this method the *targetted double control strategy*, while a SCP sample used in this strategy is called an *adapted SCP sample* (ASCP sample). For adapted SCP sampling, a measure of the response burden for a unit can simply be defined as 1 if the unit is a non-desired unit and 0, otherwise; this represents a proxy for large and small cumulated response burdens respectively, and defines a measure of the unit status. Other measures of response burden can be used. For instance, we employ the cumulated response burden in Section 3.4.

These two options (unit status and cumulated response burden respectively) are used in the algorithm given below which describes the adapted SCP sampling for two negatively coordinated samples s_1 and s_2:

Step a: based on previous information (previous surveys), create the vector or matrix which includes the information about the unit status (desired or non-desired) or the cumulated response burdens of the units in U;

Step b: select s_1 using the corresponding inclusion probabilities, by applying SCP sampling with the maximum weight strategy. The Euclidean distances between units are computed using the vector or matrix created in Step a. If the cumulated response burdens of the units are used in the vector or matrix, update them, as well as the vector or matrix, after the selection of s_1, according to the definition given in Expression (3);

Step c: select s_2, negatively coordinated with s_1, using the corresponding inclusion probabilities, by applying SCP sampling with the maximum weight strategy. The Euclidean distances between units are computed using the same vector or matrix as for selection of s_1 if the unit status is used, or the updated ones if the cumulated response burdens are used.

If s_1 and s_2 are negatively coordinated using PRNs, s_1 is drawn using $u_k, k \in U$, while s_2 using $1 - u_k, k \in U$ (u_k replaces $\widetilde{u}_k$ in Step 3b of the algorithm given in Section 2.4).

Remark 1 *As indicated in Section 2.3, the response burden is a random variable. The Euclidean distances in the previous algorithm are computed conditionally on the realized value of $RB_k, k \in U$.*

3.2 Effectiveness of targetted double control strategy

We provide in the next two sections results of Monte-Carlo simulation to show the effectiveness of the proposed strategy, and use two methods to test its performances:

- Method 1: two samples are negatively coordinated using PRNs;

- Method 2: two samples are negatively coordinated but without using PRNs. One new sample s_2 is drawn and negatively coordinated with an existing sample s_1 (s_1 is fixed and the inclusion probabilities are known). It is possible that s_1 was selected using PRNs, but these numbers are not available for the second selection. Thus, conditional on s_1, a new random number u_{k2} is associated with unit $k \in U$: if $k \in s_1$, one generates u_{k2} independently from the $Unif(0, \pi_{k1})$ distribution; otherwise one generates u_{k2} independently from the $Unif(\pi_{k1}, 1)$ distribution. Next, k is selected in s_2 using the number $1 - u_{k2}$, and the probability $\pi_{k2}, k \in U$.

In practice, Method 1 is applied in the general case when it is possible to draw new samples for both survey 1 and 2. Method 2 is applied if it is not possible to draw a new sample s_1 for practical reasons. For example, if the businesses have already received a questionnaire for survey 1, or when a new survey is added to a coordination system and the PRNs of the businesses cannot be extracted from the system.

Five measures provided below are used to quantify the performance of the proposed strategy. Measures 1 and 2 focus on the selection of non-desired units. Measure 1 quantifies the number of pairs (s_1, s_2) with given numbers of non-desired units in common between s_1 and s_2 (that is, the number of non-desired units in the overlap), and it is the most important for the study of the proposed strategy application. We expect that the proposed strategy will reduce the variance of the number of non-desired units in the overlap compared to its competitors and independent sampling. Measure 2 is related to ALB_k, with k being a non-desired unit. Ideally, one wants to reach the lower bound ALB_k given in Expression (1) for $P(k \in s_1, k \in s_2) = \pi_{k,12}$, for any $k \in U$. As underlined in Section 2.2, we are able to reach it when both s_1 and s_2 are selected using Poisson sampling with PRNs (so when Method 1 is applied). Method 2 does not always allow it to be reached. We hope that the proposed strategy provides values of the estimated $P(k \in s_1, k \in s_2)$ of the non-desired units close to this bound, showing that such units have small chances to be selected in the two samples. In some cases, however, a direct comparison with ALB_k, is not possible (see for example Method 2, framework 2 in Section 3.3), but we expect that the proposed strategy provides values of the estimated $P(k \in s_1, k \in s_2)$ lower than its competitors, or at least similar.

Measures 3, 4 and 5 concern the overall performance of a negative coordination method. A value of Measure 3 close to ALB indicates an important degree of negative coordination of the two samples. Measures 4 and 5 are measures of the overlap (relative) variance. As before, we expect that our strategy is able to provide lower values than its competitors for Measures 3, 4, and 5, or at least similar.

The five measures are:

- Measure 1: number of pairs of samples (s_1, s_2), with s_1 and s_2 containing in common a number of given non-desired units through simulations;
- Measure 2: values of the estimated $P(k \in s_1, k \in s_2)$ of the non-desired units through simulations;
- Measure 3: the Monte-Carlo expected overlap

$$E_{sim}(c) = \frac{1}{m} \sum_{\ell=1}^{m} c_\ell^{1,2},$$

where $c_\ell^{1,2} = |s_{1\ell} \cap s_{2\ell}|$, and $s_{1\ell}$, $s_{2\ell}$, are the samples drawn in the ℓ^{th} run, and $|s_{1\ell} \cap s_{2\ell}|$ represents the number of common units of $s_{1\ell}$ and $s_{2\ell}$, m is the number of runs; for Method

2, when s_1 is fixed (as in some simulations below), $c_\ell^{1,2} = |s_1 \cap s_{2\ell}|$, where $s_{2\ell}$, is the sample drawn in the ℓ^{th} run, where $|s_1 \cap s_{2\ell}|$ represents the number of common units of s_1 and $s_{2\ell}$;

- Measure 4: the Monte-Carlo variance of the overlap

$$V_{sim}(c) = \frac{1}{m-1} \sum_{\ell=1}^{m} (c_\ell^{1,2} - E_{sim}(c))^2;$$

- Measure 5: the Monte-Carlo coefficient of variation of the overlap

$$CV_{sim}(c) = \frac{\sqrt{V_{sim}(c)}}{E_{sim}(c)}.$$

3.3 Simulation with MU284 population

We consider as U region 2 of the well-known MU284 data (see Appendix B in Särndal et al., 1992). Samples s_1 and s_2 with expected sizes $n_1 = 10$ and $n_2 = 6$ are selected respectively from this region which contains in total $N = 48$ units. No births or deaths are used. The inclusion probabilities π_{k1} and $\pi_{k2}, k \in U$ are respectively proportional to variables P75 and P85, the population size of Swedish municipalities in 1975 and 1985. We call this setting the 'MU284 population'. Six units from U (with labels 4, 12, 21, 22, 32 and 44) are declared non-desired units for both s_1 and s_2. As a measure of the response burden of a unit k we use a binary variable to compute the Euclidean distances in adaptive SCPS : 1, if unit k is a non-desired unit, and 0 otherwise (the unit status).

We provide below results of Monte-Carlo simulations using Methods 1 and 2 and different sampling schemes. In the simulations 100,000 runs are used.

In Method 1, both s_1 and s_2 are random in each run and have the same type (the same sampling scheme is used), but different inclusion probabilities. They are negatively coordinated with PRNs in the following first 5 cases below, while the 6th case refers to independent sample selections. In each run of the Monte-Carlo simulation, we draw:

1. two Poisson samples;
2. two Pareto samples;
3. two adapted SCP samples; for s_1, s_2, only π_1 and respectively π_2 are used to compute the Euclidean distances between the units (this case is indicated ASCP_π in the following tables; ASCP stands for adapted SCP);
4. two adapted SCP samples; for s_1, π_1 and the values of the unit status are used to compute the Euclidean distances; for s_2, π_2 and the values of the unit status are used to compute the Euclidean distances (this case is indicated by ASCP_π_inf in the following tables);
5. two adapted SCP samples; for s_1, s_2, the values of the unit status are used to compute the Euclidean distances (this case is indicated by ASCP_inf in the following tables);
6. two independent adapted SCP samples (without negative coordination); for s_1, s_2 , the values of the unit status are used to compute the Euclidean distances (this case is indicated by 'ASCP_inf without coordination' in the following tables).

Pareto sampling is introduced in the Monte-Carlo simulation because it provides fixed sample sizes, as well as the adaptive SCPS. The inclusion probabilities are used in the adaptive SCPS in cases 3 and 4 above to compare with case 5, which only uses the binary information to compute the Euclidean distances.

Method 2 is applied in two different frameworks in order to make the connection with the application given in Section 3.4 which uses real data:

1. framework 1: both s_1 and s_2 are random in each run; s_1 is a Poisson sample, while s_2 is a sample of the type enumerated for Method 1 (Poisson, Pareto, etc.);

2. framework 2: the first sample s_1 is fixed ($s_1 = \{4, 6, 12, 18, 22, 35, 44\}$) and is a Poisson sample, and only s_2 is random in each run; s_2 is a sample of the type enumerated for Method 1 (Poisson, Pareto, etc.).

Note that Measure 2 ($\widehat{P}(k \in s_1, k \in s_2)$) can be compared to $ALB_k, k \in U$ for Method 1 and Method 2 (but only in framework 1). In Method 2, framework 2, s_1 is fixed and we are not able to reconstruct by simulation $\sum_{s_1, k \in s_1} \sum_{s_2, k \in s_2} p(s_1, s_2) = P(k \in s_1, k \in s_2)$. Thus, we only estimate $\sum_{s_2, k \in s_2} p(s_1, s_2)$. A conditional ALB_k cannot be used in this case. The corresponding tables below do not include values of ALB_k.

The tables in Section 3.3.1 present the results for Method 1, while for Method 2, they are given in Section 3.3.2 (framework 1) and Section 3.3.3 (framework 2). Using Measure 1 (see Tables 1, 4 and 7), the possible number of non-desired units common to s_1 and s_2 is between 0 and 6 for Method 1 and Method 2, framework 1, and between 0 and 4 for Method 2, framework 2. Both Poisson and Pareto sampling reach the maxima of the ranges. In contrast, adaptive SCPS (even without PRNs) shrinks the distribution of the possible number of non-desired units in common, and avoids the selection of a large number of non-desired units, resulting in a decreased variance of their number, compared with its competitors.

The use of the binary variable (without any supplementary information, such as the inclusion probabilities) in the adaptive SCPS seems to be the best choice to compute the Euclidean distances between units (the case ASCP_inf). For this setting, in general, ASCP_inf performs the best, and selects mostly 1 (see Table 7) or 2 non-desired units in common (see Tables 1 and 4), that is, fewer than the other two methods. However, no pairs (s_1, s_2) provide 0 non-desired units in common in Method 1 or Method 2, framework 2 (see Tables 1 and 7), as shown by Poisson and Pareto sampling.

For Measure 2, in Method 1 and Method 2 (framework 1), Poisson sampling is able to reach the ALB_k for any non-desired unit k, as expected (Tables 2 and 5). ASCP_inf and Pareto sampling perform similarly in the case of these two methods, and provide values of $\widehat{P}(k \in s_1, k \in s_2)$ equal to ALB_k or slightly larger. In Method 2, framework 2, $\widehat{P}(k \in s_1, k \in s_2)$ cannot be compared to ALB_k, because of the way the samples are simulated. For this method, ASCP_inf presents values of Measure 2 in agreement with Poisson and Pareto sampling, and no sampling method is the best.

In Tables 3 and 6, Pareto sampling shows lower values for the expected overlap (Measure 3; excepting Poisson sampling which reaches ALB as expected), indicating a very good overall degree of negative sample coordination. However, it displays a large estimated variance of the overlap (Measure 4), comparable to that of Poisson sampling. Compared to Pareto sampling, ASCP_inf shows a larger value for Measure 3 in Method 1 and Method 2, framework 1, but substantially reduces the values of Measures 4 and 5, indicating a better precision in estimating the overlap between s_1 and s_2. In Method 2, framework 2, ASCP_inf again performs the best for all Measures 3, 4 and 5 (see Table 9).

3.3.1 Method 1: both s_1 and s_2 are random in each run; s_1, s_2 are samples of the same type, but with different inclusion probabilities

Table 1: MU284 population, Method 1, Measure 1: number of pairs (s_1, s_2) by possible number of non-desired units in common over 100,000 runs (so the row sums are equal to 100,000).

	Possible number of non-desired units in common						
Design	0	1	2	3	4	5	6
Poisson	7469	27170	36007	21879	6536	919	20
Pareto	5225	25946	38965	23534	5784	541	5
ASCP_π	49	8646	69411	21145	749	0	0
ASCP_π_inf	0	3614	93807	2575	4	0	0
ASCP_inf	0	3443	93406	3140	11	0	0
ASCP_inf without coordination	0	831	45796	53127	246	0	0

Table 2: MU284 population, Method 1, Measure 2: $\widehat{P}(k \in s_1, k \in s_2)$, with k being a non-desired unit, 100,000 runs

	$\widehat{P}(k \in s_1, k \in s_2)$					
Non-desired unit k	4	12	21	22	32	44
Poisson	0.63	0.03	0.26	0.35	0.34	0.34
Pareto	0.63	0.07	0.27	0.35	0.34	0.34
ASCP_π	0.63	0.12	0.30	0.35	0.36	0.38
ASCP_π_inf	0.63	0.03	0.27	0.35	0.35	0.35
ASCP_inf	0.63	0.06	0.26	0.35	0.35	0.35
ASCP_inf without coordination	0.64	0.25	0.38	0.42	0.42	0.42
ALB_k	0.63	0.03	0.26	0.35	0.34	0.34

Table 3: MU284 population, $ALB = 1.96$, Method 1, Measures 3, 4, 5: $E_{sim}(c)$, $Var_{sim}(c)$, $CV_{sim}(c)$, 100,000 runs

Design	$E_{sim}(c)$	$Var_{sim}(c)$	$100 \times CV_{sim}(c)$
Poisson	1.96	1.13	54.44
Pareto	2.00	0.99	49.53
ASCP_π	2.18	0.34	26.53
ASCP_π_inf	2.03	0.09	15.08
ASCP_inf	2.01	0.08	13.70
ASCP_inf without coordination	3.06	0.68	26.91

3.3.2 Method 2, framework 1: both s_1 and s_2 are random in each run; s_1 is a Poisson sample

Table 4: MU284 population, Method 2, framework 1, Measure 1: number of pairs (s_1, s_2) by possible number of non-desired units in common, both s_1 and s_2 are random in each run over 100,000 runs (so the row sums are equal to 100,000).

	Possible number of non-desired units in common						
Design	0	1	2	3	4	5	6
Poisson	114060	51724	125551	157299	108236	37832	5298
Pareto	5995	26389	37893	23216	5931	569	7
ASCP_π	1053	16383	45920	34313	2331	0	0
ASCP_π_inf	1025	16950	48830	33175	20	0	0
ASCP_inf	1663	20987	44084	33266	0	0	0
ASCP_inf without coordination	315	6178	34797	58418	292	0	0

Table 5: MU284 population, Method 2, framework 1, Measure 2: both s_1 and s_2 are random in each run; 100,000 runs

	$\widehat{P}(k \in s_1, k \in s_2)$					
Non-desired unit k	4	12	21	22	32	44
Poisson	0.63	0.03	0.26	0.35	0.34	0.34
Pareto	0.63	0.06	0.26	0.35	0.34	0.34
ASCP_π	0.63	0.13	0.34	0.35	0.36	0.40
ASCP_π_inf	0.63	0.03	0.35	0.35	0.37	0.41
ASCP_inf	0.63	0.06	0.26	0.35	0.37	0.41
ASCP_inf without coordination	0.64	0.25	0.38	0.43	0.42	0.42
ALB_k	0.63	0.03	0.26	0.35	0.34	0.34

Table 6: MU284 population, $ALB = 1.96$, Method 2, framework 1, Measures 3, 4, 5: $E_{sim}(c)$, $Var_{sim}(c)$, $CV_{sim}(c)$, both s_1 and s_2 are random in each run; 100,000 runs

Design	$E_{sim}(c)$	$Var_{sim}(c)$	$100 \times CV_{sim}(c)$
Poisson	1.96	1.12	54.41
Pareto	1.98	1.03	51.03
ASCP_π	2.24	0.63	35.38
ASCP_π_inf	2.18	0.56	34.28
ASCP_inf	2.10	0.61	37.20
ASCP_inf without coordination	3.05	0.82	29.78

3.3.3 Method 2, framework 2: $s_1 = \{4, 6, 12, 18, 22, 35, 44\}$ is fixed and is a Poisson sample, while s_2 is random in each run

Table 7: MU284 population, Method 2, framework 2, Measure 1: number of pairs (s_1, s_2) by possible number of non-desired units in common, $s_1 = \{4, 6, 12, 18, 22, 35, 44\}$ is fixed in each run over 100,000 runs (so the row sums are equal to 100,000).

	Possible number of non-desired units in common				
Design	0	1	2	3	4
Poisson	11970	38609	36849	12033	539
Pareto	14497	47328	33006	5085	84
ASCP_π	2449	69047	24291	4183	30
ASCP_π_inf	0	81609	17689	702	0
ASCP_inf	0	84227	15773	0	0
ASCP_inf without coordination	0	18869	60320	20802	9

Table 8: MU284 population, Method 2, framework 2, Measure 2: $s_1 = \{4, 6, 12, 18, 22, 35, 44\}$ is fixed in each run; 100,000 runs

	$\widehat{P}(k \in s_1, k \in s_2)$					
Non-desired unit k	4	12	21	22	32	44
Poisson	0.64	0.05	0	0.41	0	0.40
Pareto	0.59	0.04	0	0.33	0	0.33
ASCP_π	0.64	0.13	0	0.32	0	0.21
ASCP_π_inf	0.64	0.05	0	0.36	0	0.14
ASCP_inf	0.64	0.09	0	0.32	0	0.11
ASCP_inf without coordination	0.65	0.37	0	0.50	0	0.50

Table 9: MU284 population, Method 2, framework 2, Measures 3, 4, 5: $E_{sim}(c)$, $Var_{sim}(c)$, $CV_{sim}(c)$, $s_1 = \{4, 6, 12, 18, 22, 35, 44\}$ is fixed in each run; 100,000 runs

Design	$E_{sim}(c)$	$Var_{sim}(c)$	$100 \times CV_{sim}(c)$
Poisson	1.51	0.76	57.99
Pareto	1.29	0.60	60.19
ASCP_π	1.30	0.35	45.12
ASCP_π_inf	1.19	0.17	34.47
ASCP_inf	1.16	0.13	31.48
ASCP_inf without coordination	2.31	0.64	34.66

3.4 Simulation with business surveys: Application to 'hot-spot units' at Statistics Netherlands

Statistics Netherlands (CBS) operates a sample coordination system for business surveys. Despite the sample coordination, each year several businesses are still heavily sampled, mainly because of the number of drawn samples, different stratification schemes and large sampling fractions. This results in a large cumulated response burden for these specific businesses. Therefore, CBS started monitoring the number of surveys for which a business was sampled within the last twelve months to identify so-called 'hot-spot units' (equivalent to non-desired units, see Section 3.1). CBS classifies businesses with 0–9 employees (size classes 0–3) as hot-spot units if they are sampled ≥ 3 times within the last twelve months. Businesses with 10–19 employees (size class 4) are classified as 'hot-spot units' if they are sampled ≥ 4 times within the last twelve months. For larger businesses, no hot-spot units were defined despite the large sample fractions that are required for these businesses. Businesses with more than 50 employees are usually sampled with inclusion probability 1. These large businesses usually have dedicated staff to fill in the questionnaires. The impact on daily business is therefore lower for these businesses. Moreover, sample coordination is generally not suitable for businesses with a sampling fraction of 1.

For the application we consider the population of Dutch businesses in 2021 with 0–19 employees ($N = 1810581$), the Structural Business Survey (SBS) with sample size $n_{\mathrm{SBS}} = 54491$, the Investment Survey (INV) with $n_{\mathrm{INV}} = 30090$ and the Finance Monitor (FIN) with $n_{\mathrm{FIN}} = 6977$. In this population there are 1693 hot-spot units. Table 10 shows the distribution of businesses by size class (columns) and number of samples (rows). The hot-spot units are highlighted in italics. In 2021, there were no businesses with 0-19 employees that were selected in more than nine surveys by the Dutch coordination system.

Table 10: Distribution of all Dutch businesses with 0–19 employees in 2021 by size class and number of surveys; hot-spot units are highlighted in italics

Number of	Size class					
surveys	0	1	2	3	4	Total
0	381591	1051551	202137	39544	8047	1682870
1	4832	40580	35529	16640	11732	109313
2	85	985	2509	3196	7194	13969
3	*8*	*34*	*125*	*423*	2736	3326
4	0	*4*	*6*	*46*	*785*	841
5	0	*1*	*1*	*3*	*167*	172
6	0	0	*1*	*1*	*56*	58
7	0	0	0	*1*	*22*	23
8	0	0	0	0	*7*	7
9	0	0	0	0	*2*	2
Total	386516	1093155	240308	59854	30748	1810581

All three surveys are annual surveys with a stratified sample design with equal probability within strata. The strata are defined by a combination of industrial classification according to NACE[1] and size class. SBS and INV are coordinated by the Dutch coordination system (see Smeets and

[1]Nomenclature statistique des activits conomiques dans la Communeaut europenne, the standard European industrial classification.

Boonstra, 2018) and use the same stratification. FIN is independent from SBS and INV, but the system coordinates the samples for FIN from year to year. FIN only selects businesses with ≥ 2 employees (size class 2 and larger) and uses different combinations of NACE codes than SBS and INV to define the strata.

We consider the following scenarios using Methods 1 and 2 (see Section 3.2) and different combinations of the surveys:

1. Method 1, where both samples s_1 and s_2 are drawn for SBS. Both samples use the SBS allocation of 2021 (scenario 1).
2. Method 1, where s_1 is drawn for SBS and s_2 is drawn for INV. Both samples use the allocation of the corresponding survey of 2021 (scenario 2).
3. Method 2, where s_2 is drawn for SBS, conditional on the existing SBS sample s_1 of 2021, i.e., s_1 is fixed. Sample s_2 uses the SBS allocation of 2021 (scenario 3).
4. Method 2, where s_2 is separately drawn for SBS, INV and FIN, conditional on sample s_1, that is obtained by combining the existing samples of SBS, INV and FIN of 2021, i.e., s_1 is fixed. The samples for s_2 use the allocation of the corresponding survey of 2021. It is possible that a unit is drawn for each of the three separate samples (scenario 4).

In this simulation, method 2 (in scenarios 3 and 4) is only considered under framework 2 (see Section 3.3). Scenarios 1 and 3 represent coordination over time for one survey. Scenario 2 represents coordination over time and over two surveys with common stratification. Scenario 4 represents coordination over time and over two surveys with different stratifications. Because of the large population sizes and since the majority of the strata do not contain any hot-spots at all, in all scenarios a selection of strata is used for the simulation. First, in scenarios 1 and 3 the take-all strata ($\pi_{k1} = \pi_{k2} = 1$) are excluded and in scenarios 2 and 4 both the take-all ($\pi_{k2} = 1$) and take-none strata ($\pi_{k2} = 0$) are excluded. In scenario 4, a stratum is only excluded if $\pi_{k2} = 0$ or 1 in this stratum for all three surveys. Second, strata are selected based on the population size and the expected number of sampled hot-spot units, such that the total population size N is around 1000. In scenarios 1 and 3, strata with less than 300 businesses in the population and at least 10 expected hot-spots in the sample are selected. In scenario 2, strata with less than 500 businesses and at least six expected hot-spots are selected. In scenario 4, all strata with at least one expected hot-spot in the sample for all three surveys are selected. Table 11 gives the population size and sample information for the considered scenarios. The populations of businesses and hot-spot units are a subset of the populations shown in Table 10. The sample size of the combined samples in scenario 4 is denoted by n_{comb}.

Table 11: Population and sample information for the scenarios

					sample s_1		sample s_2		
scenario	method	N	strata	hot-spots	n_{SBS}	n_{comb}	n_{SBS}	n_{INV}	n_{FIN}
1	1	894	5	97	617		617		
2	1	1053	8	161	659			487	
3	2	894	5	97	617		617		
4	2	418	34	62		418	384	278	37

Both the information of whether a particular business is considered a hot-spot (binary information as used in Section 3.3) and the number of surveys for which the business was sampled within the last 12 months are measures of its cumulated response burden. In the simulation with CBS business surveys, we use both measures to compute the Euclidean distance between the units. However, because of the stratified sampling with equal inclusion probabilities within the strata, the inclusion probabilities of the surveys are not used to compute the Euclidean distances between the units in this simulation, as in Section 3.3.

We provide results of a Monte-Carlo simulation for the considered scenarios and different sampling schemes. The sampling schemes are applied per stratum. Due to the substantial computational burden, 5000 runs are used. In each run of the Monte-Carlo simulation we draw:

- two Poisson samples for Method 1 and one Poisson sample for Method 2;
- two Pareto samples for Method 1 and one Pareto sample for Method 2;
- two adapted SCP samples for Method 1 and one adapted SCP sample for Method 2; for s_1, s_2 the measure of cumulated response burden based on hot-spot status is used to define the Euclidean distances between the units (indicated by ASCP_inf);
- two adapted SCP samples for Method 1 and one adapted SCP sample for Method 2; for s_1, s_2 the measure of cumulated response burden given by the number of surveys is used to compute the Euclidean distances between the units (indicated by ASCP_inf_svy);
- two independent adapted SCP samples (without negative coordination) for Method 1 and one independent adapted SCP sample for Method 2; for s_1, s_2 the measure of cumulated response burden based on the hot-spot status is used to compute the Euclidean distance between the units (indicated by ASCP_inf without coordination).

In the tables and figures below, we use the following notation: ASCP_inf indicates results based on adapted SCP sampling with hot-spot status used to define the Euclidean distances between the units and negative coordination, ASCP_inf_svy for adapted SCP sampling with cumulated response burden and negative coordination (svy stands for survey), while 'ASCP_inf without coordination' for adapted SCP sampling and independent sample selection. The results of scenarios 1, 2, and 3 are shown in Figure 1 and Table 12. The results of scenario 4 are shown in Figure 2 and Table 13 (for ASCP_inf hot-spot status is used as the burden measure and for ASCP_inf_svy the number of surveys is used). In scenarios 1, 2 and 3 all measures give similar results to the simulation with the MU284 population. The results of Measure 2 are in line with the results presented for the MU284 population and are not shown here to save space. In these scenarios, adapted SCP sampling with the measure of cumulated response burden based on hot-spot status is the best sampling strategy. This is, because ASCP_inf leads to the smallest variation of the overlaps, not only for all businesses but also for the hot-spot units. This implies that the overall response burden is most evenly spread by ASCP_inf. In scenario 4 the differences between the sampling schemes are smaller. This is caused by FIN using different strata than SBS and INV, which leads to small strata when adapted SCP sampling is applied as a coordination method for the three surveys together. When s_2 is drawn for SBS or INV, the ASCP_inf_svy sampling scheme is slightly better than ASCP_inf. When in scenario 4 sample s_2 is drawn for the SBS, the adapted SCP sampling and independent sampling perform similarly. This has to do with the selection of the strata in this scenario. The inclusion probabilities of SBS and INV are large in these strata, while the inclusion probabilities of FIN are small. Moreover, 28 strata are take-all strata for SBS and 17 strata are take-all for INV, while FIN has no take-all strata in this selection. When Method 2 is applied as coordination method and if $k \in s_1$, the random number u_{k2} is generated from $Unif(0, \pi_{k1})$. If π_{k1} is close to 1 then u_{k2} is generated from a distribution that is approximately equal to $Unif(0, 1)$. Generating u_{k2} from $Unif(0, 1)$ implies independent sampling.

Figure 1: Scenarios 1, 2, 3; Measure 1: number of pairs of samples (s_1, s_2) (bullets) by possible number of hot-spot units in common (y-axis), 5,000 runs. The size of the bullets is an indication of the number of sampled pairs.

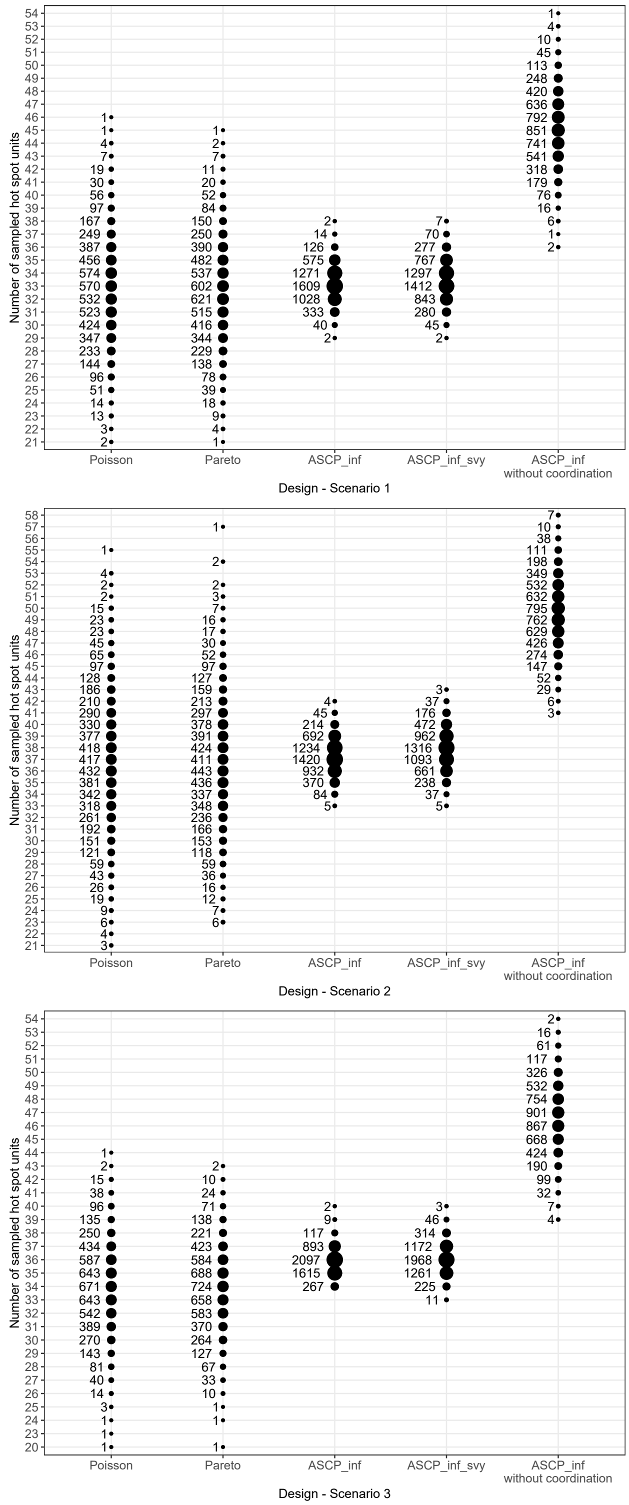

Figure 2: Scenario 4; Measure 1: number of pairs of samples (s_1, s_2) (bullets) by possible number of hot-spot units in common (y-axis), 5,000 runs. The size of the bullets is an indication of the number of sampled pairs.

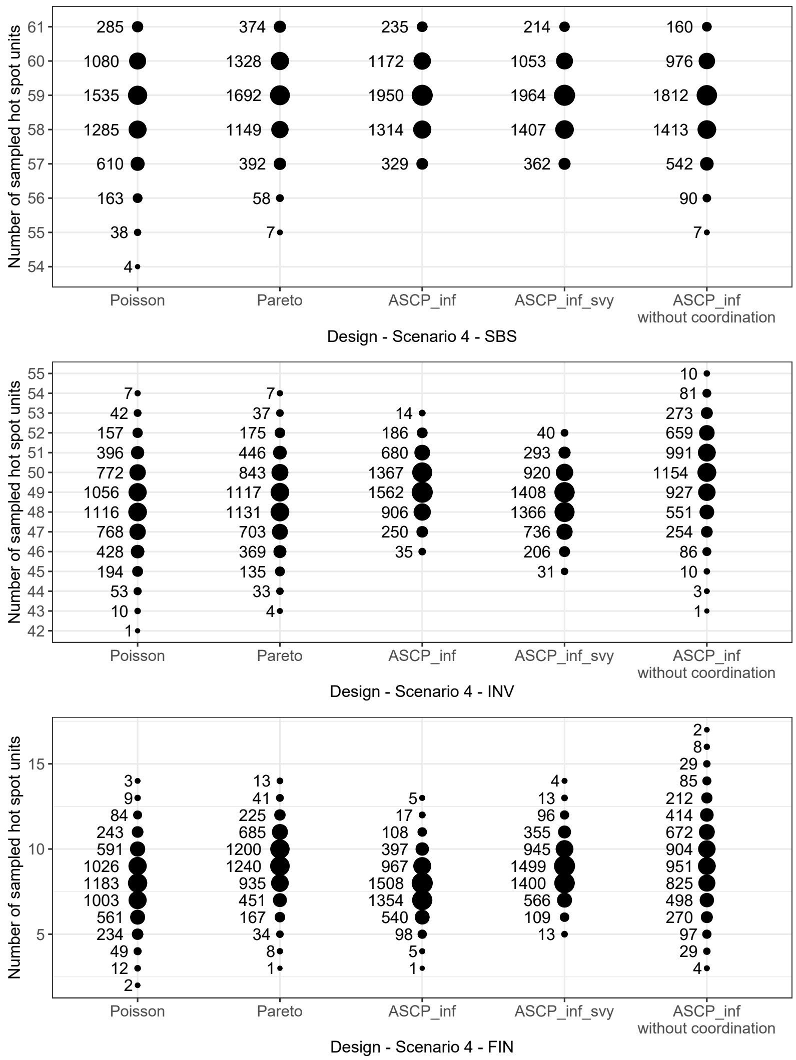

Table 12: Scenarios 1, 2 and 3; Measures 3, 4, 5: $E_{sim}(c), Var_{sim}(c), CV_{sim}(c)$, 5,000 runs

Design – Scenario 1	$E_{sim}(c)$	$Var_{sim}(c)$	$100 \times CV_{sim}(c)$
Poisson	352	109.0	2.97
Pareto	352	20.3	1.28
ASCP_inf	357	0.7	0.23
ASCP_inf_svy	365	2.5	0.45
ASCP_inf without coordination	455	35.2	1.31
Design – Scenario 2	$E_{sim}(c)$	$Var_{sim}(c)$	$100 \times CV_{sim}(c)$
Poisson	211	136.0	5.53
Pareto	211	49.9	3.35
ASCP_inf	212	0.6	0.38
ASCP_inf_svy	214	2.4	0.73
ASCP_inf without coordination	300	35.4	1.98
Design – Scenario 3	$E_{sim}(c)$	$Var_{sim}(c)$	$100 \times CV_{sim}(c)$
Poisson	352	68.7	2.36
Pareto	352	14.3	1.07
ASCP_inf	357	1.4	0.34
ASCP_inf_svy	365	3.6	0.52
ASCP_inf without coordination	455	35.5	1.31

Table 13: Scenario 4; Measures 3, 4, 5: $E_{sim}(c), Var_{sim}(c), CV_{sim}(c)$, 5,000 runs

Design – Scenario 4 SBS	$E_{sim}(c)$	$Var_{sim}(c)$	$100 \times CV_{sim}(c)$
Poisson	375	11.10	0.890
Pareto	378	0.00	0.000
ASCP_inf	379	0.91	0.252
ASCP_inf_svy	379	0.87	0.246
ASCP_inf without coordination	379	0.92	0.254
Design – Scenario 4 INV	$E_{sim}(c)$	$Var_{sim}(c)$	$100 \times CV_{sim}(c)$
Poisson	271	36.00	2.210
Pareto	274	0.00	0.000
ASCP_inf	273	1.93	0.508
ASCP_inf_svy	273	1.99	0.516
ASCP_inf without coordination	274	1.57	0.584
Design – Scenario 4 FIN	$E_{sim}(c)$	$Var_{sim}(c)$	$100 \times CV_{sim}(c)$
Poisson	21	12.20	17.00
Pareto	24	0.00	0.00
ASCP_inf	21	3.65	8.95
ASCP_inf_svy	21	3.71	9.13
ASCP_inf without coordination	24	5.11	9.40

4 Discussion

The strategy developed in Section 3.1 provides an approach to negative coordination of samples which fulfils the requirement to reduce the overlap size between two or more samples, and additionally reduces the variance of the number of nondesirable units (units with particular characteristics) in the overlap compared to its competitors. Thus, a double control of the response burden is targetted at this specific set of units. The coordination strategy results in a more even spread of the response burden of these units. The targetting can be achieved through an indicator variable, or through a continuous variable demonstrating the size of a unit for the characteristic of interest (such as the cumulative response burden).

We consider several variants of the approach, depending on the kind of information used to designate non-desired units. In general ASCP_inf performs the best because it uses only binary information and non-desired units are therefore as similar as possible to each other on this characteristic, and as different as possible from other units. This approach is therefore better at avoiding samples with clusters of non-desired units. Other variants may however be better in situations where there is a gradation of non-desirability. In our simulations, the proposed strategy shows a smaller variance of the number of non-desired units in the overlap (especially for ASCP_inf) compared to Poisson and Pareto sampling. This is due to the spread of the units in the space generated by the measure of response burden used (spread obtained by using the algorithm given in Section 2.4), as indicated in the tables and figures related to Measure 1. On the other hand, similar results to the competitor methods were obtained for the expected overlap size, while the variance of the overlap size was smaller than for Poisson and Pareto sampling with PRNs in most cases. A single exception concerning this variance was provided by Pareto sampling with PRNs in Table 12.

Targetted double control is an effective strategy for managing situations where some businesses are selected for relatively many surveys in a short period, as demonstrated by the application to hot-spot units in Statistics Netherlands. The problem of hot-spots is not eliminated, but it is reduced because the response burden is more controlled within the constraints of the survey designs.

The targetted double control strategy can therefore be used to formalise an approach to dealing with businesses that complain that they have been selected in too many surveys. Without such a system, these are sometimes dealt with in an ad hoc way by moving them (explicitly or implicitly) to a take-none stratum, to relieve the burden. But this approach is not fair in that it can be different for businesses with the same characteristics depending on whether they complain or not. For a single sample selected at any given moment, targetted double control is ‘fair’ in that it minimises the number of such units included in the sample by spreading the selected units through the space generated by the measure of the response burden used and thus avoiding the clustering of non-desired units) and because it respects the inclusion probabilities so that unbiased estimates can be obtained. Statistics Canada have considered extending a take-none stratum to deal with small units which may receive a disproportionate burden (Landry, 2011), but we consider that using targetted double control would be a better solution in this case too. Targetted double control therefore addresses an important practical problem in the coordination of multiple samples in a finite population.

Applications of targetted double control are not restricted to selection hot-spots, however, and any kind of undesirable unit could in principle be the target of the method, as long as the undesirability property is observable or predictable from available data sources.

Acknowledgements

We are grateful to the referees for their comments.

References

Bondesson, L. and Thorburn, D. (2008). A list sequential sampling method suitable for real-time sampling. *Scandinavian Journal of Statistics*, 35:466–483.

Bottone, M., Modugno, L., and Neri, A. (2021). Response burden and data quality in business surveys. *Journal of Official Statistics*, 37(4):811–836.

Bradburn, N. (1978). Respondent burden. In *Proceedings of the Survey Research Methods Section of the American Statistical Association*, volume 35, pages 35–40. American Statistical Association Alexandria, VA.

Brewer, K., Early, L., and Joyce, S. (1972). Selecting several samples from a single population. *Australian Journal of Statistics*, 3:231–239.

Cotton, F. and Hesse, C. (1992). Tirages coordonnés d'échantillons. Technical Report E9206, Direction des Statistiques Économiques, INSEE, Paris, France.

Grafström, A. (2012). Spatially correlated Poisson sampling. *Journal of Statistical Planning and Inference*, 142(1):139–147.

Grafström, A., Lisic, J., and Prentius, W. (2022). `BalancedSampling`*: balanced and spatially balanced sampling.* R package version 1.6.3, https://CRAN.R-project.org/package=BalancedSampling.

Grafström, A. and Matei, A. (2018). Coordination of spatially balanced samples. *Survey Methodology*, 44(2):215–238.

Landry, S. (2011). Managing response burden by controlling sample selection and survey coverage. In *Proceedings of the Survey Research Methods Section of the American Statistical Association*, volume 2011, pages 2276–2284. American Statistical Association Alexandria, VA.

Lorenc, B., Kloek, W., Abrahamsson, L., and Eckman, S. (2013). Description of actual response burden and response behaviour with Swedish register. In Giesen, D., Bavdaž, M., and Bolko, I., editors, *Comparative report on integration of case study results related to reduction of response burden and motivation of businesses for accurate reporting. BLUE-ETS deliverable 8.1*, pages 11–27. doi: 10.13140/RG.2.2.24741.63207.

Mach, L., Reiss, P. T., and Şchiopu-Kratina, I. (2006). Optimizing the expected overlap of survey samples via the northwest corner rule. *Journal of the American Statistical Association*, 101(476):1671–1679.

Matei, A. and Smith, P. A. (2023). Sample coordination methods and systems for establishment surveys. In Snijkers, G., Bavdaž, M., Bender, S., Jones, J., MacFeely, S., Sakshaug, J., Thompson, K., and van Delden, A., editors, *Advances in Business Statistics, Methods and Data Collection*, chapter 27. Wiley, Hoboken.

Matei, A. and Tillé, Y. (2005). Maximal and minimal sample co-ordination. *Sankhyā*, 67(3):590–612.

Natkowska, M. and Modak, A. (2014). *Theme: Sample Co-ordination*. European Commission, Eurostat, Directors of Methodology.

Ohlsson, E. (1995). Coordination of samples using permanent random numbers. In Cox, B. G., Binder, D. A., Chinnapa, B. N., Christianson, A., Colledge, M. J., and Kott, P. S., editors, *Business Survey Methods*, chapter 9, pages 153–169. Wiley, New York, USA.

Qualité, L. (2019). Coordinated sampling: Theory, method and application at SFSO. Presentation, ENBES workshop on Coordinated Sampling for Business Surveys. https://statswiki.unece.org/display/ENBES/ENBES+Workshop+2019+%28The+Hague%29+-+Workshop+on+Coordinated+Sampling?preview=/234225830/244092344/Slides_Qualite_2019.pdf.

Rosén, B. (1997a). Asymptotic theory for order sampling. *Journal of Statistical Planning and Inference*, 62:135–158.

Rosén, B. (1997b). On sampling with probability proportional to size. *Journal of Statistical Planning and Inference*, 62:159–191.

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

Smeets, M. and Boonstra, H. J. (2018). Sampling coordination of business surveys at Statistics Netherlands. In Lorenc, B., Smith, P., Bavdaž, M., Haraldsen, G., Nedyalkova, D., Zhang, L.-C., and Zimmermann, T., editors, *The Unit Problem and Other Current Topics in Business Survey Methodology*, pages 127–137. Cambridge Scholars, Newcastle-upon-Tyne.

Sunter, A. B. (1977). Response burden, sample rotation, and classification renewal in economic surveys. *International Statistical Review*, 45:209–222.