Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

Spreading response burden in business surveys at Statistics Netherlands: Evaluating sample coordination methods targeting highly burdened businesses

by Marc J.E. Smeets and Jonas Klingwort

Release date: June 29, 2026

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

• Statistical Information Service	1-800-263-1136
• National telecommunications device for the hearing impaired	1-800-363-7629
• Fax line	1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, the Agency has developed standards of service which its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Survey Methodology, June 2026
Vol. 52, No. 1, pp. 151-172
Statistics Canada, Catalogue No. 12-001-X

Spreading response burden in business surveys at Statistics Netherlands: Evaluating sample coordination methods targeting highly burdened businesses

Marc J.E. Smeets and Jonas Klingwort[1]

Abstract

National statistical institutes operate sample coordination systems to spread the response burden in business surveys. Despite the applied sample coordination and monitoring the response burden, some businesses might still be heavily sampled within a short period. This may lead to a peaking response burden for individual businesses, which could affect response rates and response quality. This paper proposes a new sample coordination method based on Adapted Spatially Correlated Poisson (ASCP) sampling that focuses on businesses with a high response burden. The effects on the response burden will be evaluated in two simulation studies and compared with a stratified approach, a pragmatic method in which sampling fractions are manually adjusted and with the baseline method of ignoring the response burden. For the simulations, real-world scenarios and data from Statistics Netherlands are used. The first simulation study considers a practical situation in which a given sample is adjusted with the aim to avoid the occurrence of businesses with a peaking response burden. The second simulation study analyzes the longer-term effects of the different sample coordination methods and focuses both on the reduction and spread of the response burden. The advantages and disadvantages of the different methods will be explained and discussed in detail, and recommendations for applying these methods at national statistical institutes and other survey agencies will be given.

Key Words: Adapted Spatially Correlated Poisson sampling; Pareto sampling; Poisson sampling; Response burden; Sample coordination.

1. Introduction

The requirement to reduce the (perceived) response burden in business surveys is commonly faced in official statistics (Bradburn, 1978; Bavdaž, Giesen, Černe, Loöfgren and Raymond-Blaess, 2015; Giesen, Vella, Brady, Brown, Ravindra and Vaasen-Otten, 2018; Erikson, Giesen, Houben and Saraiva, 2023). Although response burden is a difficult concept to define as it may include both objective factors such as the time spent to provide questionnaire responses and subjective factors such as what is perceived as burdensome by the respondents (Yan, Fricker and Tsai, 2019; Bottone, Modugno and Neri, 2021; Yan and Williams, 2022), there is empirical evidence that increased perceived burden is associated with higher attrition, partial response, lower data quality, and lower response rates (Lorenc, Kloek, Abrahamsson and Eckman, 2013; Bottone et al., 2021).

National Statistical Institutes (NSIs) use two main strategies to control the response burden. First, reducing the total response burden by, for example, sending out fewer questionnaires and second, spreading the response burden so that the overall response burden is distributed among individual businesses as evenly as possible. To spread the response burden, NSIs apply sample coordination (Ernst, 1999; Giesen et al., 2018; Erikson et al., 2023) and use a sample coordination system. Sample coordination aims to control the

1. Marc J.E. Smeets, Statistics Netherlands, Netherlands. E-mail: mje.smeets@cbs.nl; Jonas Klingwort, Statistics Netherlands, Netherlands. E-mail: j.klingwort@cbs.nl.

size of the overlap between successively drawn samples as opposed to the case where the samples are drawn independently. Positive coordination attempts to maximize the overlap to obtain more precise estimates of change. Negative coordination methods minimize the overlap with the purpose of reducing the number of surveys in which a particular business is selected in a given period, and therefore to spread the response burden (Bradburn, 1978). Matei and Smith (2023) give an overview of sample coordination methods and also review sample coordination systems that are in use by NSIs. In the remainder of this paper, we mean negative sample coordination when we use the term sample coordination.

Statistics Netherlands (SN) operates a sample coordination system called the Survey Burden System (SBS) (Smeets and Boonstra, 2018). This system applies sample coordination among groups of surveys that is based on permanent random numbers (PRNs) and accounts for the cumulated response burden of the businesses based on the number of surveys in which a business was selected. More specifically, the businesses are sorted increasingly first by the cumulated response burden and subsequently by the PRNs. The first businesses in this ordering are then selected for the sample. Despite the applied sample coordination, some businesses are still heavily sampled each year. This is mainly because of large sampling fractions and detailed stratification but is also due to the fact that not all surveys make optimal use of the features of the SBS. This results in a peaking response burden for some individual businesses, which can cause problems especially for small companies. The sample coordination method presented in this paper is not implemented in the SBS. This paper studies whether this method can be a useful enhancement of the existing SBS and might be implemented in a future version of the SBS.

In response to complaints from the Dutch business community (Strategische commissie betere regelgeving bedrijven, 2020), SN started a project in May 2021 to prevent, as far as possible, businesses from being sampled disproportionately often. This project focuses primarily on small businesses with 0-19 working persons. A pilot study that ran from May 2021 to May 2022 monitored how often a business was sampled within the last twelve months (Gommans, Burgerjon and Paulissen, 2022). In this pilot study, businesses with 0-9 working persons are labeled as “hotspots” if they have been selected for more than 3 different surveys within the last twelve months. Businesses with 10-19 working persons are labeled as “hotspots” if they have been selected for more than 4 surveys within the last 12 months. Generally speaking, these hotspots can be considered as non-desired units. Desired units would be usually units with a low cumulated response burden and non-desired units usually have a large cumulated response burden (Matei, Smith, Smeets and Klingwort, 2023). During the pilot, pragmatic methods were applied to control the response burden by reducing or spreading the number of hotspots as much as possible (see Subsection 2.3). The pragmatic methods were applied at Statistics Netherlands because there was demand from politics (Strategische commissie betere regelgeving bedrijven, 2020). They reflect a practical situation where pragmatic solutions were needed because time was lacking to implement sound sampling methodology. The contribution of this paper is the introduction of a new sample coordination method and also a comparison of pragmatic methods that specifically aim at the reduction of hotspots (or non-desired units in general) and methodologically sound methods that aim at the spread of the response burden. We consider the analyses and results of this study of practical importance for other national statistical agencies as these might face comparable situations.

In Matei et al. (2023), a sample coordination method has been developed by means of Adapted Spatially Correlated Poisson (ASCP) sampling that realizes a double control of the response burden for businesses labeled as non-desired units. In this context, it concerns businesses that NSIs want to spare from sampling as much as possible, for example because of a large cumulated response burden. This method uses PRNs and is based on correlated Poisson sampling (Grafström, 2012). In Matei et al. (2023) this method is applied in a simulation study based on a setting inspired from Dutch business surveys, and it is shown that the variance of the number of hotspots selected in several samples can be reduced compared to Poisson sampling with PRNs (Brewer, Early and Joyce, 1972) and Pareto order sampling (Rosén, 1997a,b).

By means of ASCP sampling, samples can be drawn to avoid clustering of non-desired units in the sample. The use of ASCP sampling in a sample coordination method could prevent the selection of large numbers of businesses with a high response burden within a given period. The sample coordination method proposed in this paper updates the hotspot status of the businesses and applies ASCP sampling to the hotspots without using PRNs. We compare this method and a stratified approach with the pragmatic methods applied in the pilot study, evaluate the methods using Monte Carlo simulation, and investigate their longer-term effects on the number of sampled hotspots. We consider that especially the comparison with the methods from the pilot study is relevant, since they are straightforward to implement for NSIs or survey agencies, but may lead to problems, as we will show.

We investigate whether the proposed method based on ASCP sampling can be applied as an alternative to the methods used in the pilot study. In the first simulation study, we examine the reduction of the number of hotspots in a given sample. In the second simulation study, we investigate the longer-term effects of the studied methods on the number of sampled hotspots where both reduction and spread is considered and verify whether the methods respect the prescribed inclusion probabilities of the underlying sample design. Data from the Dutch Research and Development (R&D) survey conducted in 2020 are used for both simulation studies.

The paper is organized as follows: In Section 2, an overview of the considered sampling methods is given and it is described how these methods can be used for sample coordination. The simulation studies are outlined in Section 3. The results are presented in Section 4. A discussion and conclusion are given in Section 5.

2. Sampling methods and sample coordination

In this section, we briefly describe Poisson, Pareto and ASCP sampling, explain how they can be used for sample coordination, propose a new sample coordination method, and outline the methods used in the pilot study.

2.1 Poisson, Pareto and ASCP sample designs

A Poisson sample from a population of N units is drawn as follows (Brewer et al., 1972): generate random numbers $r_1, \ldots, r_k, \ldots, r_N \overset{\text{iid}}{\sim} \operatorname{Unif}(0, 1)$ and select unit k in sample S if $r_k < \pi_k$. Here $\pi_k = \mathrm{P}(k \in S)$ is the inclusion probability of unit, or business, k. Note that the sample size n_S of a Poisson sample is random, with expected sample size $\mathrm{E}(n_S) = \sum_{k=1}^{N} \pi_k$.

Sometimes fixed-size sample designs are preferred. An example of such a sample design is Pareto sampling (Rosén, 1997b), which is also used in the pilot study and for the stratified approach (see below). For given random numbers $r_1, \ldots, r_k, \ldots, r_N \overset{\text{iid}}{\sim} \operatorname{Unif}(0, 1)$ and inclusion probabilities $\pi_1, \ldots, \pi_k, \ldots, \pi_N$, a Pareto sample of size n is obtained by selecting the n units with the smallest values of

$$\rho_k = \frac{r_k / (1 - r_k)}{\pi_k / (1 - \pi_k)}.$$

ASCP sampling is a special case of correlated Poisson sampling (Bondesson and Thorburn, 2008). Correlated Poisson sampling is a list sequential method used to draw a random sample S from a population U with prescribed inclusion probabilities π_k for $k \in U$. Consider the selection probability $\pi_\ell^{(\ell-1)}$, for $1 \leq \ell \leq N$, that allows unit ℓ to be selected in S in the ℓth iteration of the following algorithm (see Step 3b):

1. set $\pi_k^{(0)} = \pi_k$, for all $k \in U$.
2. set $\ell = 1$;
3. while $\ell \leq N$ do
   (a) generate r_ℓ independently from the $\operatorname{Unif}(0, 1)$ distribution,
   (b) if $r_\ell < \pi_\ell^{(\ell-1)}$ then $I_\ell = 1$ (ℓ is selected in S) else $I_\ell = 0$ (ℓ is not selected in S), where I_ℓ is the indicator variable associated to unit $\ell \in U$.
   (c) update the selection probabilities for the remaining units $i = \ell + 1, \ldots, N$ according to
   $$\pi_i^{(\ell)} = \pi_i^{(\ell-1)} - (I_\ell - \pi_\ell^{(\ell-1)})\, w_\ell^{(i)},$$
   where $w_\ell^{(i)}$ are some weights given by unit ℓ to other units $i = \ell + 1, \ldots, N$.
   (d) increase ℓ by 1.

The choice of $w_\ell^{(i)}$ in Step 3c provides different sample designs; for example, Poisson sampling is obtained if all $w_\ell^{(i)} = 0$ for $\ell \in U$. Bondesson and Thorburn (2008) derive conditions on the choice of $w_\ell^{(i)}$, such that $0 \leq \pi_i^{(\ell)} \leq 1$ and fixed-size samples are obtained. They also show that under these conditions, the prescribed inclusion probabilities are respected.

Grafström (2012) applied correlated Poisson sampling to a population where each unit has associated geographical coordinates. The underlying idea is that contiguous units may provide similar information and that a sample contains more information if the sample avoids clustering of contiguous units. By using weights $w_\ell^{(i)}$ in Step 3c that meet the conditions from Bondesson and Thorburn (2008), such that unit ℓ

gives maximal weight to the unit closest to ℓ in (Euclidean) distance, among the units $i = \ell + 1, \ldots, N$, Spatially Correlated Poisson (SCP) sampling is provided. This method realizes geographic coverage of the entire survey population by spreading the sampled units across the geographic area.

Instead of applying SCP sampling to geographical coordinates, SCP sampling can also be applied to other characteristics of units, for example, the cumulated response burden or weights based on the completion time of the questionnaire, resulting in well-spread samples with respect to these characteristics. Note that these characteristics should not depend on response indicators, as this may lead to selectivity in the drawn samples, that is, selective groups of respondents may become overrepresented in the sample. Matei et al. (2023) apply SCP sampling to a measure of the cumulated response burden and call this Adapted Spatially Correlated Poisson (ASCP) sampling.

Here, we introduce a binary variable that represents whether a business is a hotspot unit or not. Using this binary variable as a measure for the cumulated response burden leads to a special case of ASCP sampling, in which spatially balanced samples are drawn with respect to the space generated by this variable. The binary variable for unit k has value 1 if k is a hotspot and 0 otherwise. ASCP samples can be drawn using the function `scps()` in the R package `BalancedSampling` (Grafström, Lisic and Prentius, 2024).

We also consider a stratified approach of ASCP sampling for stratified designs with equal inclusion probabilities within the strata. Let $1, \ldots, h, \ldots, H$ denote the strata of population U, where stratum h has sample size n_h and population size N_h. The inclusion probabilities are given by the stratum sampling fractions, i.e., $\pi_k = f_h = n_h / N_h$ for the units k in stratum h. By defining a separate substratum for the hotspots and non-hotspot units in stratum h and by drawing random samples in both substrata with sampling fraction f_h also a spatially balanced sample is obtained. We use Pareto sampling to draw the samples in the substrata.

2.2 Sample coordination

The sampling methods of Subsection 2.1 can be used for sample coordination, where commonly permanent random numbers (PRNs) are used (Ohlsson, 1995). The idea is that the businesses in the population keep the assigned random numbers $r_1, \ldots, r_k, \ldots, r_N \overset{\text{iid}}{\sim} \text{Unif}(0, 1)$ in the selection process over time. Negative sample coordination is then achieved by using r_k and $1 - r_k$ alternately as random numbers for $k \in U$ when drawing successive samples. For example, Poisson sampling with PRNs (Brewer et al., 1972) works as follows: for given inclusion probabilities $\pi_1, \ldots, \pi_k, \ldots, \pi_N$, unit k is selected in the first sample if $r_k < \pi_k$ and is selected in the next sample if $1 - r_k < \pi_k$, and so on.

In a similar way, sample coordination methods are constructed from the other sampling methods of Subsection 2.1. Combining Pareto sampling with PRNs leads to Pareto order sampling (Rosén, 1997a,b). Matei et al. (2023) apply sample coordination by using ASCP sampling in combination with PRNs. The PRNs r_k and $1 - r_k$ are then used in Step 3b of the algorithm in Subsection 2.1 and Step 3a is skipped. This method is introduced as the targeted double control strategy. It allows a double control of the response burden of non-desired units, once by applying sample coordination by means of the PRNs to minimize the overlap between samples drawn in successive surveys and once by drawing spatially balanced samples to avoid clustering of non-desired units in the sample.

We propose another sample coordination method based on ASCP sampling that does not use PRNs. The idea is as follows: keep the hotspot status of the businesses up-to-date over time and use ASCP sampling to draw spatially balanced samples with respect to the hotspot status as described in Subsection 2.1. We call this method the "ASCP hotspot strategy". By this method, the response burden is spread in a more realistic way for businesses, because it will lead to longer periods in which businesses are not contacted. In contrast, by the methods that use PRNs as described above, businesses are alternately selected and not selected in successive samples. For this reason we do not consider the PRN-based sample coordination methods such as the targeted double control strategy in the simulation studies.

2.3 Pragmatic methods applied in the pilot study

The data collection department of SN draws the samples for the Dutch business surveys. For most of the surveys this is done by means of the sample coordination system (SBS). During the pilot study, the data collection group tracked which units in the drawn samples became hotspots. If possible and in collaboration with the statistical departments, hotspots were manually removed from the samples before starting the fieldwork using the following methods (Gommans et al., 2022):

1. Remove hotspots from the drawn sample manually. As a result of this, fewer questionnaires are sent out and the realized sample size is smaller than the prescribed sample size.
2. Replace hotspots by other units manually. In this way the prescribed sample size is maintained as much as possible. The following strategies were applied:
   (a) Consider an existing sample that contains hotspots. Based on the number of hotspots, the number of expected hotspots for a new sample can be calculated. Adding the expected number of hotspots to the required sample size, redraw a new and larger sample first and then manually remove the hotspots and any additional units from the larger sample to end up with the required sample size. Repeat these steps iteratively until as many hotspots as possible are replaced. It may require a few iteration steps before the final sample is obtained.
   (b) Remove all hotspots from the sampling frame manually and draw a sample with the initially prescribed sample size from the remaining units in the sampling frame.
   (c) Replace hotspots in a rotating panel by other units manually during the panel rotation.

Note that these pragmatic methods effectively adjust the inclusion probabilities of the individual units. This means that the realized inclusion probabilities in the final sample are no longer equal to the prescribed inclusion probabilities in the sample design. In this case the sample is no longer a random sample and using a design-based estimator introduces a risk of biased estimates. By calibrating the observed response to the population totals it can be partially corrected for this bias.

The applied strategies in the pilot study aim to control the response burden either by reducing (method 1 above) or spreading (methods 2a, 2b and 2c above) the number of sampled hotspots. Spreading here refers to allocating the total sample size differently over all units in the population to replace hotspots with non-hotspots.

This paper focuses on sample coordination applied to cross-sectional surveys. We want to compare ASCP sampling and Using substrata with the strategies of the pilot study, so only methods 2a and 2b are relevant in this study.

Methods 2a and 2b are two different methods leading to a sample in which the hotspots are replaced by other units as much as possible. In method 2a this is accomplished by (iteratively) increasing the sampling fractions of non-hotspot units and drawing preliminary samples until the desired sample is obtained. In method 2b this is accomplished by sampling only once from the thinned out sampling frame, but with the initially prescribed inclusion probabilities. In both methods the sampling fractions of the non-hotspot units are therefore increased during the sampling process. Methods 2a and 2b are considered as similar methods in this paper and we only apply method 2b in the simulation studies in Section 3. We refer to the sample coordination method in which the hotspot status is updated over time and method 2b is applied as "Increasing fractions".

3. Outline of the simulation studies

This section describes the outline of both simulation studies. The purpose is to investigate whether ASCP sampling can be used as an alternative to the pragmatic method of increasing fractions (Section 2). Simulation study 1 focuses on the reduction of the number of hotspots in a given sample drawn by the SBS. Simulation study 2 focuses on the longer-term effects on the sampled hotspots and examines whether all methods respect the prescribed sampling fractions of the sample design. This simulation study informs on both the reduction and the variation over time of the number of hotspots.

Both simulation studies use data from the Dutch Research & Development (R&D) survey conducted in 2020. The R&D survey is an annual survey with businesses as sampling units and a stratified sample design with equal inclusion probabilities within the strata (CBS, 2022). This implies that the considered methods can be applied to each stratum separately. The strata are defined by a combination of industrial classification according to SBI ("Standaard Bedrijfsindeling" – Standard Industrial Classification) and business size class (SC). The SBI is based on NACE (nomenclature statistique des activités économiques dans la Communeauté européenne, the standard European industrial classification) and SC is based on the number of working persons. The sampling frame consists of the sampling part, take-all strata, and take-none strata. There are 99 SBI combinations and 6 size classes used in the stratification. The R&D survey uses a sample for SC 4 to SC 9 (10 working persons or more). Since hotspots are defined in SC 0 to SC 4 (Section 1), only hotspots in the sampling for SC 4 can be considered in the simulations.

For both simulation studies we use the sample design that was used for the R&D survey in 2020. At that time, Statistics Netherlands did not take any action to remove or replace the hotspots from the sample, only the hotspot status was monitored. Accordingly, this sample to be used for the simulation studies reflects a realistic scenario occurring in practice. We only consider the sampling part, because spreading the response burden has no effect on the samples when the sampling fraction is 0 or 1 (take-none or take-all strata).

Let U be the population of the R&D survey containing the businesses in SC 4 that are in the sampling part of the sampling frame of the R&D 2020 survey, and let S be the sample selected from U that was drawn with the SBS for the R&D 2020 survey. The size of U is $N =$ 33,044 and sample S consists of $n =$ 926 units. There are $N_{\mathrm{HS}} = 211$ hotspots in U and $n_{\mathrm{HS}} = 108$ hotspots in S. Table A.1 in the accompanying working paper by Smeets and Klingwort (2023) shows the $H = 83$ strata, together with N_h, n_h, $f_h = n_h / N_h$, the number of hotspots in the population $N_{\mathrm{HS},h}$, the population fraction of hotspots $F_{\mathrm{HS},h} = N_{\mathrm{HS},h} / N_h$, the expected number of sampled hotspots $\mathrm{E}(n_{\mathrm{HS},h}) = n_h F_{\mathrm{HS},h}$ and the variance of the number of sampled hotspots $\mathrm{V}(n_{\mathrm{HS},h}) = n_h F_{\mathrm{HS},h}(1 - F_{\mathrm{HS},h})(N_h - n_h)/(N_h - 1)$. Note that $n_{\mathrm{HS},h}$ is hypergeometrically distributed with parameters N_h, $N_{\mathrm{HS},h}$ and n_h, from which the expectation and variance of $n_{\mathrm{HS},h}$ follow.

3.1 Simulation study 1: A practical situation

The first simulation study addresses the situation where the different methods are applied to a given sample, the initial sample, with the aim of reducing the number of hotspots in this sample. The initial sample is assumed to be drawn using the sample coordination system and is considered fixed in this simulation study. Next, in each run, a second sample is obtained using, respectively, the following methods:

1. *Ignoring hotspots*: take the initial sample and keep all the hotspots in the sample.
2. *Increasing fractions*: manually replace hotspots in the initial sample by other units by increasing the sampling fractions of non-hotspots (method 2b, see Subsection 2.3).
3. *ASCP sampling*: apply ASCP sampling taking into account the binary variable (1 if a hotspot, 0 otherwise) in the sampling process (see Subsection 2.1).
4. *Using substrata*: use two separate substrata for the hotspots and non-hotspots (Subsection 2.1).

The method of ignoring hotspots does not change the initial sample and, therefore, does not reduce the number of hotspots in the initial sample. We used this method as a reference.

We only consider strata that contain at least one hotspot in the population, because otherwise all methods produce samples without hotspots. So, we consider the strata in SC 4 with $0 < f_h < 1$ and $N_{\mathrm{HS},h} > 0$. Let U' be the population that meets these three requirements. Let $S' \subset S$ be the part in S containing the selected units from the strata that meet these three requirements, where S is the given sample drawn by the SBS. Sample S' is used as the initial sample. Population U' consists of $H = 48$ strata, has size $N' =$ 17,773 and S' consists of $n' = 598$ units and $n'_{\mathrm{HS}} = 108$ hotspots.

In this simulation study, sample coordination is applied in a different way as described in Subsection 2.2. The hotspot status is not updated because the initial sample and the second sample refer to the same period, and the aim is to replace the initial sample with the second sample. Sample coordination is applied by generating random numbers that lead to the initial sample when ignoring the hotspots. The other methods use these random numbers to draw the second sample. The methods are implemented in the algorithm below. Pareto sampling (Rosén, 1997b) is described in Subsection 2.1.

Algorithm 1.

1. Generate uniform random numbers $0 < r_k < 1$ for all $k \in U'$, such that $r_k < \pi_k$ if $k \in S'$ and $r_k \geq \pi_k$ otherwise. Compute

$$\rho_k = \frac{r_k / (1 - r_k)}{\pi_k / (1 - \pi_k)} \quad \text{for all} \quad k \in U'.$$

2. Apply the sampling methods:
   - *Ignoring hotspots*: draw a Pareto sample S_1 with given r_k and π_k by selecting the n_h units with the smallest values of ρ_k.
   - *Increasing fractions*: draw a Pareto sample S_2 with given r_k and sample size n_h where the $N_{\mathrm{HS},h}$ hotspots in stratum h are removed from the population. The sample is obtained by selecting the n_h units in the thinned population with the smallest values of ρ_k. If $n_h > (N_h - N_{\mathrm{HS},h}),$ then "Ignoring hotspots" is applied and no hotspots are removed from the sample in stratum $h.$
   - *ASCP sampling*: draw an ASCP sample S_3 with given r_k and π_k and hotspot status.
   - *Using substrata*: Draw separate Pareto samples in the substrata of hotspots and non-hotspots with given r_k and let S_4 be the combination of these samples. Sample size $n_{h,1}$ in the substratum of hotspots in stratum h is obtained by rounding $f_h N_{\mathrm{HS},h}$ randomly and sample size $n_{h,2}$ in the substratum of non-hotspots is $n_{h,2} = n_h - n_{h,1}.$ Random rounding is applied with probabilities proportional to $f_h N_{\mathrm{HS},h}$ and $1 - f_h N_{\mathrm{HS},h}$ of rounding up and down, respectively. The sample in the substratum of hotspots is obtained by selecting the $n_{h,1}$ units with the smallest values of ρ_k in the substratum of hotspots and for the non-hotspots, the $n_{h,2}$ units with the smallest values of ρ_k in the substratum of non-hotspots are selected.

In the algorithm, all methods are applied per stratum leading to a sample per stratum. The samples S_1, S_2, S_3 and S_4 are then obtained by combining these samples per stratum. Note that the random numbers in Step 1 are generated in such a way that $S_1 = S'$ in every simulation run. The method Increasing fractions is not applied if $n_h > (N_h - N_{\mathrm{HS},h}).$ In this case, the number of hotspots in the stratum is at least $n_h - (N_h - N_{\mathrm{HS},h})$ and could be reduced to this minimum. For practical reasons, it was decided not to do so.

In the simulation study, we draw samples $S_1^{(j)}$, $S_2^{(j)}$, $S_3^{(j)}$ and $S_4^{(j)}$ by repeating Steps 1 and 2 in each run $j = 1, \ldots, R.$ In total, we consider $R =$ 1,000 runs, where each run represents a single sample draw. For every sample $S_m^{(j)}$, drawn by method m in run $j,$ the sample size n and the number of sampled hotspots n_{HS} are computed, both for the whole sample and for the sample from stratum $h.$ We write

$$n_m^{(j)} = n\left(S_m^{(j)}\right), \quad n_{\mathrm{HS},m}^{(j)} = n_{\mathrm{HS}}\left(S_m^{(j)}\right),$$
$$n_{h,m}^{(j)} = n_h\left(S_m^{(j)}\right), \; n_{\mathrm{HS},h,m}^{(j)} = n_{\mathrm{HS},h}\left(S_m^{(j)}\right),$$

for $m = 1, \ldots, 4,$ $h = 1, \ldots, H$ and $j = 1, \ldots, R.$

In the simulation study, the realized sample sizes are verified, and the reduction of the number of sampled hotspots by the different methods is examined. For this purpose, the distributions of the number of sampled hotspots in the simulation study are investigated. For the whole sample, frequency tables are made that present the number of samples in the simulation study with a certain number of sampled hotspots. For the strata, the distributions of the number of sampled hotspots $n_{\mathrm{HS},h,m}^{(j)},$ for $j = 1, \ldots, R,$ are examined per stratum. The results are given in Section 4.

3.2 Simulation study 2: Longer-term effects

The second simulation study investigates the longer-term effects on the number of sampled hotspots and verifies whether all units are drawn with the prescribed inclusion probabilities. The same four methods as given in Simulation study 1 are considered (Ignoring hotspots, Increasing fractions, ASCP sampling and Using substrata). Now we apply sample coordination as described in Subsection 2.2 where the hotspot status is updated over time. The reference method of ignoring hotspots now amounts to independent sampling.

We simulate a series of draws over time, where the periods represent the months $t = 1, \ldots, T.$ In each month t it is determined which units in the population are hotspots and the methods are applied to prevent sampling of hotspots as much as possible. For this simulation study, we consider all strata in SC 4 with $0 < f_h < 1.$ So the samples are drawn from the population U that is described in Section 3.

This simulation study does not follow the same idea of a given initial sample. Here, new samples are drawn in both the first and the following periods. The sampling fractions of the sample S define the inclusion probabilities in the sampling process, where S is the given sample drawn by the SBS. The methods are implemented in the algorithm below, where only the first step (generating the random numbers) differs from Algorithm 1.

Algorithm 2.

1. Generate random numbers $r_1, \ldots, r_k, \ldots, r_N \overset{\text{iid}}{\sim} \mathrm{Unif}(0,1)$ and compute
$$\rho_k = \frac{r_k / (1 - r_k)}{\pi_k / (1 - \pi_k)} \text{ for all } k \in U.$$
2. Draw samples $S_1,$ $S_2,$ S_3 and S_4 by applying Step 2 of Algorithm 1, for given $r_k,$ π_k and $\rho_k.$

Also in Algorithm 2, the methods in Step 2 are applied per stratum. The overall samples $S_1,$ $S_2,$ S_3 and S_4 are obtained by combining the samples drawn from the strata. In this simulation study the hotspot status is no longer fixed, but is updated every month. We use the same definition as in the pilot study, that is, a unit in SC 4 is a hotspot when it is drawn for more than 4 surveys in the past 12 months.

For each method we simulate a series with a length of 25 years, so T = 300 months. We draw samples $S_1^{(t)},$ $S_2^{(t)},$ $S_3^{(t)}$ and $S_4^{(t)}$ by repeating the Steps 1 and 2 for each month $t = 1, \ldots, T.$ For every sample $S_m^{(t)},$

drawn by method m in month $t,$ the sample size n and the number of sampled hotspots n_{HS} are computed, both for the whole sample and for the samples per stratum h:

$$\begin{aligned} n_m^{(t)} &= n\left(S_m^{(t)}\right), & n_{\text{HS},m}^{(t)} &= n_{\text{HS}}\left(S_m^{(t)}\right), \\ n_{h,m}^{(t)} &= n_h\left(S_m^{(t)}\right), & n_{\text{HS},h,m}^{(t)} &= n_{\text{HS},h}\left(S_m^{(t)}\right), \end{aligned}$$

for $m = 1, \ldots, 4,$ $h = 1, \ldots, H$ and $t = 1, \ldots, T.$

To study the longer-term effects of the sampling methods on the number of sampled hotspots, we look at the development of the number of sampled hotspots in the time series for the whole population and for the strata separately, so $n_{\text{HS},m}^{(t)}$ and $n_{\text{HS},h,m}^{(t)},$ for $t = 1, \ldots, T$ and method $m = 1, \ldots, 4.$

We finally compute the realized sampling fractions for the subpopulations of hotspots and non-hotspots per stratum. The realized sampling fractions are computed by averaging over time, in the following way:

$$\begin{aligned} \hat{f}_{\text{HS},h,m} &= \sum_{t=1}^{T} n_{\text{HS},h,m}^{(t)} \Big/ \sum_{t=1}^{T} N_{\text{HS},h}^{(t)}, \\ \hat{f}_{\text{nHS},h,m} &= \sum_{t=1}^{T} n_{\text{nHS},h,m}^{(t)} \Big/ \sum_{t=1}^{T} N_{\text{nHS},h}^{(t)}. \end{aligned}$$

Here $N_{\text{HS},h}^{(t)}$ and $N_{\text{nHS},h}^{(t)}$ are, respectively, the number of hotspots and non-hotspots in the population in month t and stratum h and $n_{\text{nHS},h,m}^{(t)}$ is the number of sampled non-hotspots in month t and stratum h by method $m.$ We compute estimates for the 95% margins of the realized sampling fractions by assuming the normal distribution as following

$$2\sqrt{\frac{f_h(1-f_h)}{\sum_{t=1}^{T} N_{\text{HS},h}^{(t)}}} \quad \text{and} \quad 2\sqrt{\frac{f_h(1-f_h)}{\sum_{t=1}^{T} N_{\text{nHS},h}^{(t)}}},$$

for the subpopulations of hotspots and non-hotspots, respectively. Here $f_h = n_h / N_h$ is the prescribed sampling fraction of stratum $h.$ It is evaluated whether the prescribed sampling fractions fall within the 95% confidence intervals based on these margins. Note, that these margins are only an approach and do not account for the discrete nature of the numbers. Furthermore, these margins only consider the variation over time and not in any given month. The results are given in Section 4.

4. Results

This section discusses the results of both simulation studies. The corresponding tables and figures are included here.

4.1 Simulation study 1: A practical situation

By definition, all considered methods draw samples with a fixed sample size (Subsection 3.1). The realized overall sample sizes are therefore equal to the prescribed sample sizes for all methods and in all simulation runs.

Table 4.1 shows the distribution of the number of sampled hotspots for all considered methods. By definition, Ignoring hotspots does not reduce the number of hotspots and leads to 108 hotspots in every sample. Increasing sampling fractions leads to the lowest number of 36 sampled hotspots in each simulation run. The number of hotspots does not vary, because the initial sample and the hotspot labels of the units do not change. Both ASCP sampling and Using substrata reduce the number of hotspots in the initial sample. Using substrata leads to fewer hotspots (with mode 55) than ASCP sampling (with mode 72), but with a larger range over the simulation runs. Although the largest reduction is obtained by the pragmatic method of increasing fractions, we will address the drawbacks of this method in the second simulation study.

Figure 4.1 shows the mean number of sampled hotspots per stratum averaged over the 1,000 simulation runs. The color indicates the population fraction of hotspots $F_{HS,h}$. By increasing the sampling fractions the number of sampled hotspots is reduced to zero in most of the strata. When using substrata, in 26 of the 48 strata the mean number of sampled hotspots is close to zero. For ASCP sampling this is only the case in 11 strata, mainly in the lower part of the displayed strata. In general, we can say that ASCP sampling leads to fewer sampled hotspots than Ignoring hotspots, but still more than Using substrata and Increasing fractions.

Table 4.1
Number of sampled hotspots over 1,000 samples in Simulation study 1 (row sums are equal to 1,000)

	Number of sampled hotspots											
	36	**53**	**68**	**69**	**70**	**71**	**72**	**73**	**74**	**75**	**76**	**108**
Ignoring												1,000
Increasing	1,000											
ASCP			3	18	88	223	275	221	131	35	6	

(a) The methods Ignoring hotspots, Increasing fractions and ASCP sampling.

	Number of sampled hotspots															
	47	**48**	**49**	**50**	**51**	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**	**61**	**62**
Substrata	1	3	7	26	60	97	134	145	172	123	102	67	31	20	7	5

(b) The methods Using substrata.

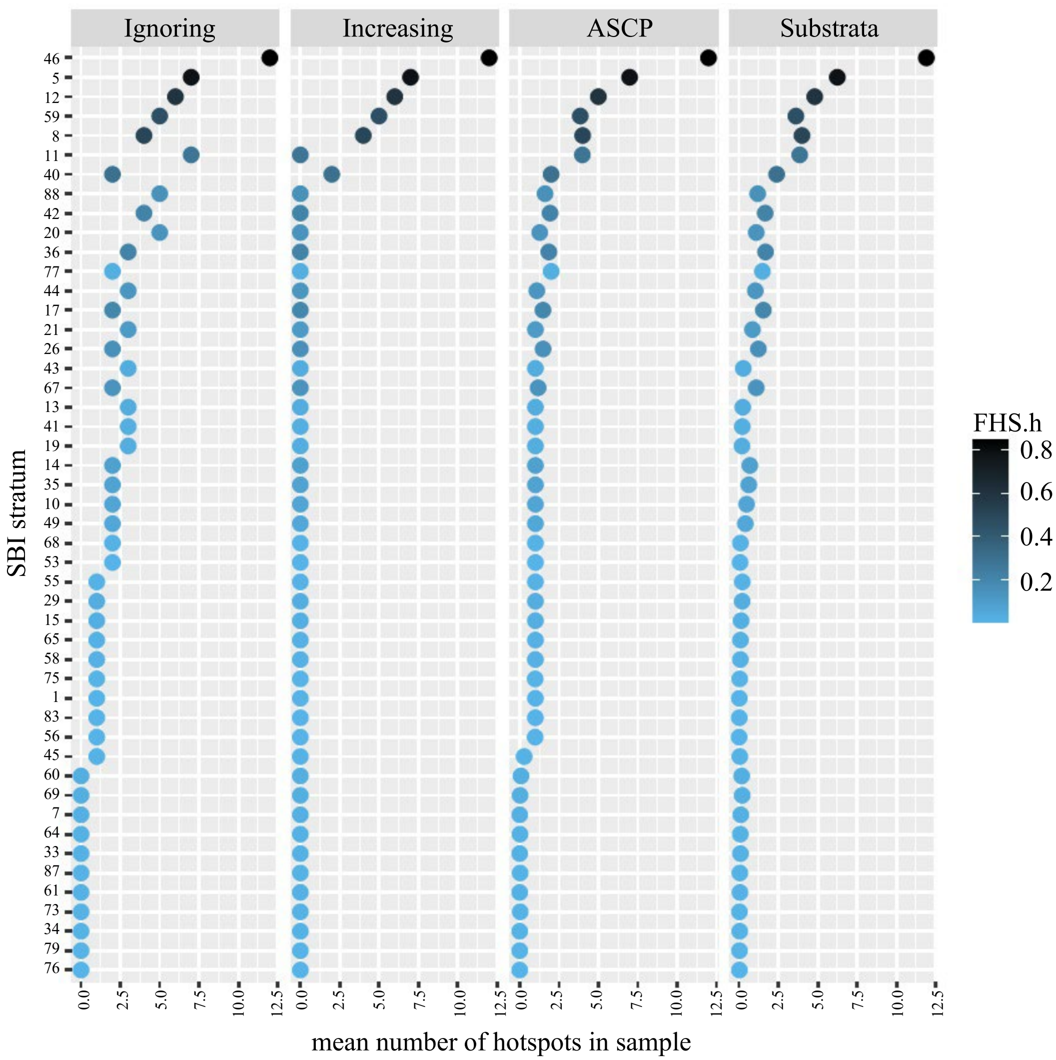
Note: The first column contains the applied methods and the following columns the number of sampled hotspots. The cells of the table show the number of samples drawn containing the number of sampled hotspots given in the corresponding column. For instance, Adapted Spatially Correlated Poisson (ASCP) sampling selects 3 samples, each one containing 68 hotspots.

Figure 4.2a displays the mean reduction of the number of sampled hotspots compared to the Ignoring hotspots method. The color indicates the expectation of the number of sampled hotspots $\mathrm{E}(n_{\mathrm{HS},h}) = n_h F_{\mathrm{HS},h}$ (Section 3). The largest reductions are achieved in strata SBI 11, 20 and 88, which is the case for all three methods. In these strata, the initial sample contains a (relatively) large number of hotspots (see Figure 4.1) and there are enough non-hotspot units in the population to replace the hotspots in the initial sample with. For these strata, we have $N_h - N_{\mathrm{HS},h} \geq n_h$ or equivalently $f_h + F_{\mathrm{HS},h} \leq 1.$ In stratum SBI 46 the number of hotspots in the initial sample is not reduced by the three sampling methods despite the high value of $\mathrm{E}(n_{\mathrm{HS},h}).$ In this stratum, we have $f_h + F_{\mathrm{HS},h} > 1,$ implying that there is not enough space in the population to replace hotspots.

Figure 4.2b compares ASCP sampling and Using substrata and shows the mean reduction of the number of sampled hotspots compared to the method of increasing fractions. The color indicates the expectation of the number of sampled hotspots $\mathrm{E}(n_{\mathrm{HS},h}) = n_h F_{\mathrm{HS},h}.$ It turns out that ASCP sampling leads to fewer hotspots only in the strata SBI 12 and 59. In both strata the number of hotspots in the preliminary sample is relatively large, but there is no space in the population to replace hotspots by non-hotspot units.

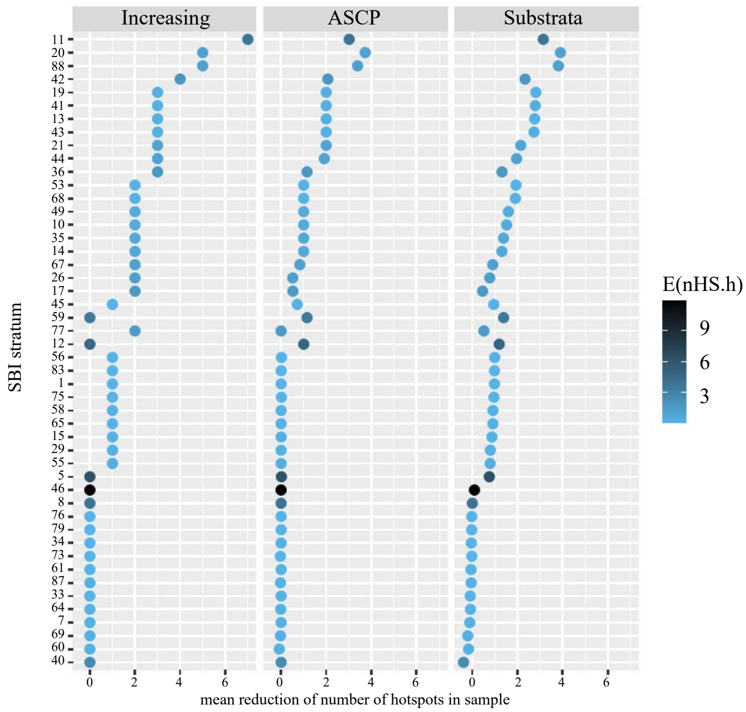
Using substrata and ASCP sampling lead to similar results. In the middle part of the displayed strata in Figure 4.2b, Using substrata leads to a mean reduction close to 0 hotspots, while ASCP sampling selects 1 hotspot more than the method of increasing fractions. In these strata the sampling fractions are small and there are only few hotspots in the population, so that $\mathrm{E}(n_{\mathrm{HS},h})$ is close to 0. Using substrata draws samples from the hotspot stratum with a sample size equal to the randomly rounded number of $\mathrm{E}(n_{\mathrm{HS},h})$ (Section 3, Algorithm 1), which is mostly 0 in these strata. ASCP sampling draws samples from the complete stratum with the prescribed sampling fraction. Rounding is also applied for ASCP sampling, but the effect on the realized sampling fractions turns out to be smaller than in the substrata approach. This is due to the fact that sampling is applied at the unit level for ASCP sampling and at the substratum level in the substratum approach.

Figure 4.1 Mean number of sampled hotspots per stratum in Simulation study 1

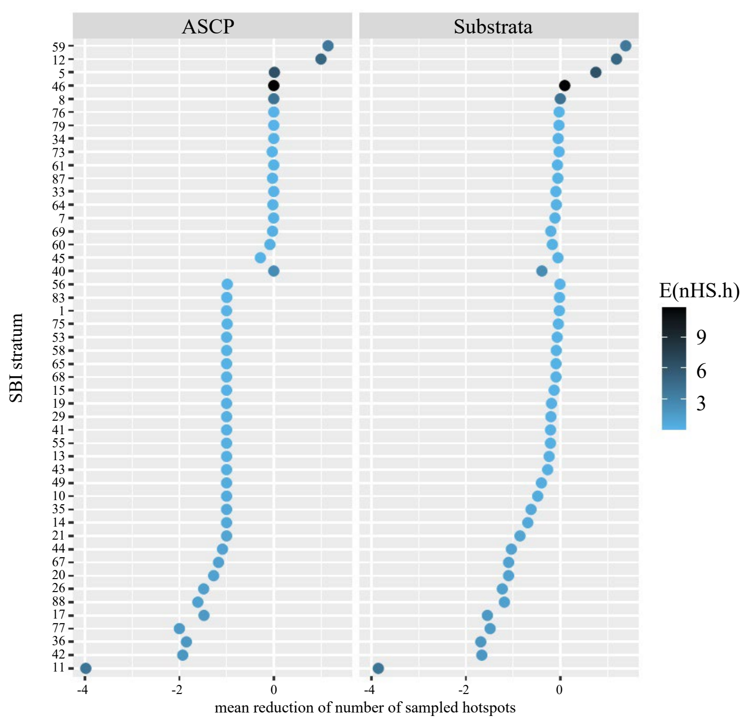


Notes: -The x-axis shows the mean number of sampled hotspots and the y-axis the 2-digit SBI classification. Color indicates the population fraction of hotspots $F_{\mathrm{HS},h}$ in stratum h.
-"Standaard Bedrijfsindeling" – Standard Industrial Classification (SBI).

Figure 4.2 Mean reduction of sampled hotspots per stratum in Simulation study 1



(a) Mean reduction of sampled hotspots compared to ignoring hotspots.



(b) Mean reduction of sampled hotspots compared to increasing fractions.

Notes: -The x -axis shows the mean reduction of sampled hotspots and the y -axis the 2-digit SBI classification. The color indicates the expected number of sampled hotspots $\mathrm{E}(n_{\mathrm{HS},h}) = n_h F_{\mathrm{HS},h}$ in stratum h.
-Adapted Spatially Correlated Poisson (ASCP); "Standaard Bedrijfsindeling" – Standard Industrial Classification (SBI).

In general, we can say that if there is enough space in the population to replace the hotspots, Increasing fractions leads to fewer sampled hotspots than ASCP sampling and Using substrata. In that case, a large reduction can be achieved when the number of hotspots in the initial sample is (relatively) large. Using substrata and ASCP sampling show similar results. By rounding, Using substrata results in slightly fewer hotspots in strata with a small value of $\mathrm{E}(n_{\mathrm{HS},h})$.

4.2 Simulation study 2: Longer-term effects

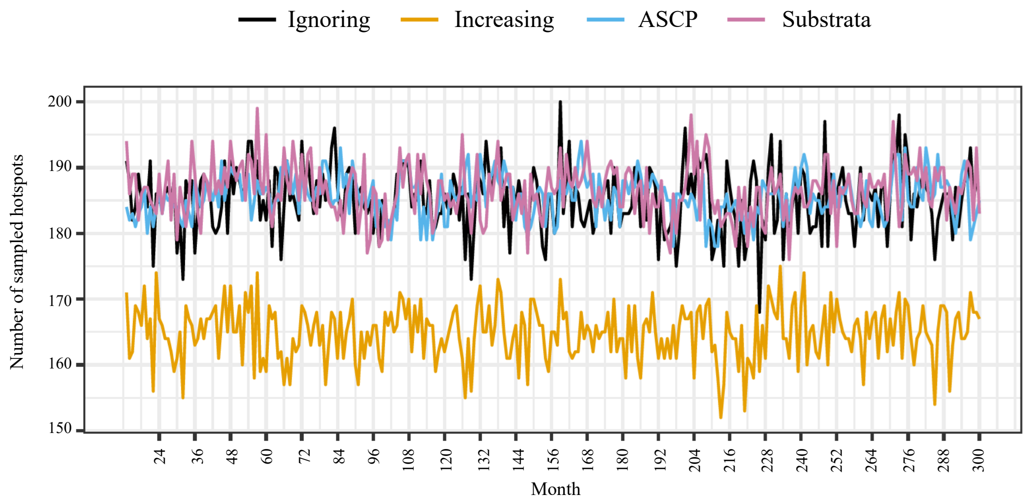
Figure 4.3a shows the development over time of the total number of sampled hotspots by the four methods in all strata together. The first 12 months are not shown because, according to the definition (Section 1), hotspots appear only after 12 months. Table 4.2a summarizes the distributions of the number of sampled hotspots over the months 13-300. The first thing to notice is that the number of hotspots generated by ASCP sampling and Using substrata is at the same level as for Ignoring hotspots with a mean of 185. There is a clear reduction in the number of sampled hotspots of about 20 when the fractions are increased. Ignoring hotspots leads to the largest standard deviation (sd), implying that the spread of response burden can be improved by Increasing fractions, ASCP sampling and Using substrata. The best spread of response burden is achieved by ASCP sampling.

Figures 4.3b and 4.3c and Tables 4.2b and 4.2c show the results in stratum selections with, respectively, small ($0.04 \leq f_h \leq 0.41$) and medium ($0.41 < f_h \leq 0.48$) sampling fractions. For strata with other sampling fractions the results of the methods are similar and are therefore not shown. When the sampling fractions are very low ($0 < f_h < 0.04$) no hotspots are created at all and the number of sampled hotspots is always zero for every method. When the sampling fractions are large ($0.48 \leq f_h < 1$) the number of sampled hotspots is at the same high level of around 160 for all methods. For large sampling fractions, the smallest sd of the number of hotspots is obtained by ASCP sampling.

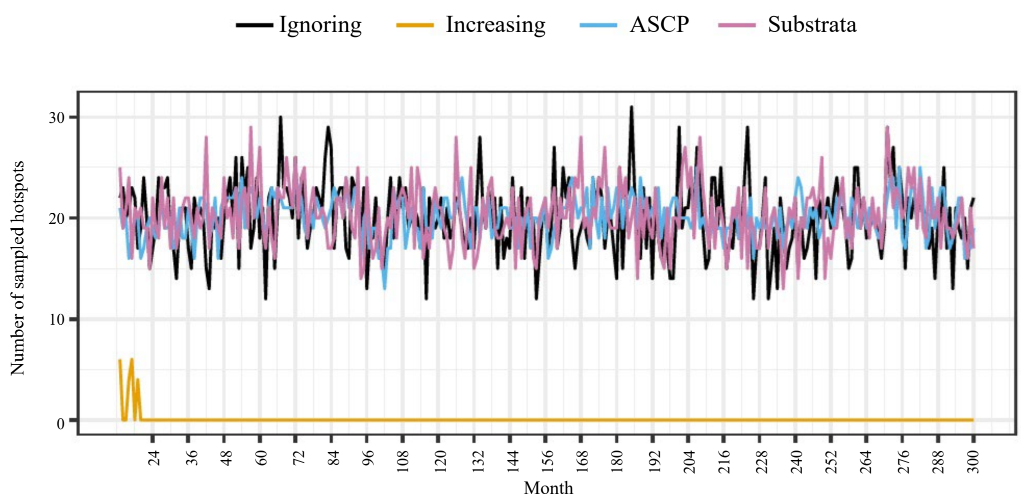
In strata with small sampling fractions ($0.04 \leq f_h \leq 0.41$) there is room in the population to replace hotspots every month. The strongest reduction is therefore achieved by increasing the fractions, where the number of sampled hotspots in each month equals zero. The numbers of sampled hotspots for ASCP sampling and Using substrata are at the same level as for Ignoring hotspots, so these methods do not reduce the number of hotspots. Compared to Ignoring hotspots and Using substrata, ASCP sampling results in the best spread of response burden, i.e., the smallest sd of the number of hotspots.

In strata with medium sampling fractions ($0.41 < f_h \leq 0.48$) the method of increasing fractions shows an interesting pattern. The numbers of sampled hotspots are roughly at the same level for all methods, while Increasing fractions leads to the largest sd. By replacing hotspots by non-hotspot units more and more units in the population are becoming hotspots. At a certain point, there is no room left to replace the hotspots. This leads to the worst spread of response burden and no reduction in the number of sampled hotspots. Also in these strata the optimal spread of response burden is obtained by ASCP sampling.

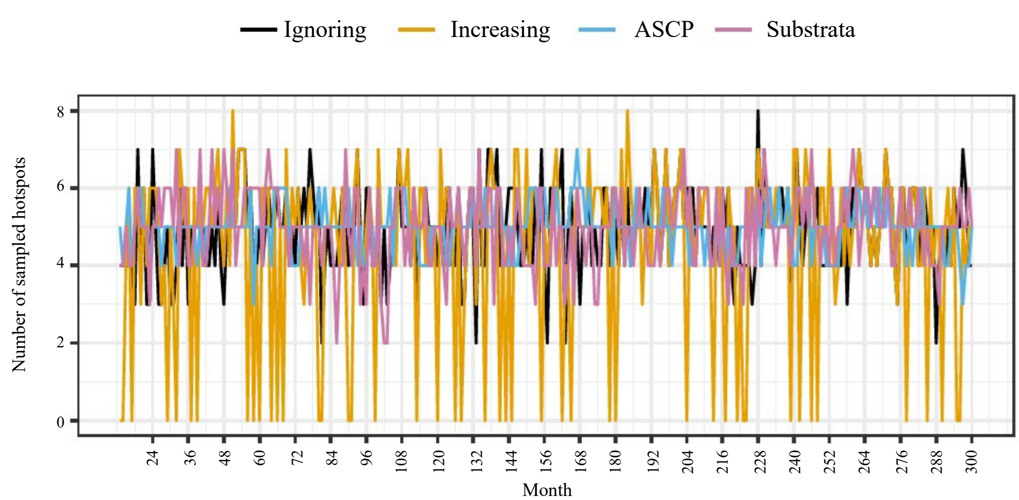
Figure 4.3 Development over time of the number of sampled hotspots in Simulation study 2



(a) All strata in the population.



(b) Strata with small sampling fractions ($0.04 \leq f_h \leq 0.41$).



(c) Strata with medium sampling fractions ($0.41 < f_h \leq 0.48$).

Note: Adapted Spatially Correlated Poisson (ASCP).

Table 4.2
Summary of the number of sampled hotspots over the months 13-300 in Simulation study 2 for all strata in the population and by sampling fraction

Method	min.	mean	max.	sd
Ignoring hotspots	168	185.2	200	4.82
Increasing fractions	152	164.9	175	4.14
ASCP sampling	178	185.6	194	3.25
Using substrata	176	186.4	199	4.06

(a) All strata.

Method	min.	mean	max.	sd
Ignoring hotspots	12	19.96	31	3.53
Increasing fractions	0	0.07	6	0.60
ASCP sampling	13	20.17	25	2.04
Using substrata	13	20.29	29	3.00

(b) Strata with small sampling fractions ($0.04 \leq f_h \leq 0.41$).

Method	min.	mean	max.	sd
Ignoring hotspots	2	4.92	8	1.14
Increasing fractions	0	4.51	8	2.19
ASCP sampling	3	5.00	7	0.64
Using substrata	2	4.98	7	0.99

(c) Strata with medium sampling fractions ($0.41 < f_h \leq 0.48$).

Note: Adapted Spatially Correlated Poisson (ASCP); standard deviation (sd).

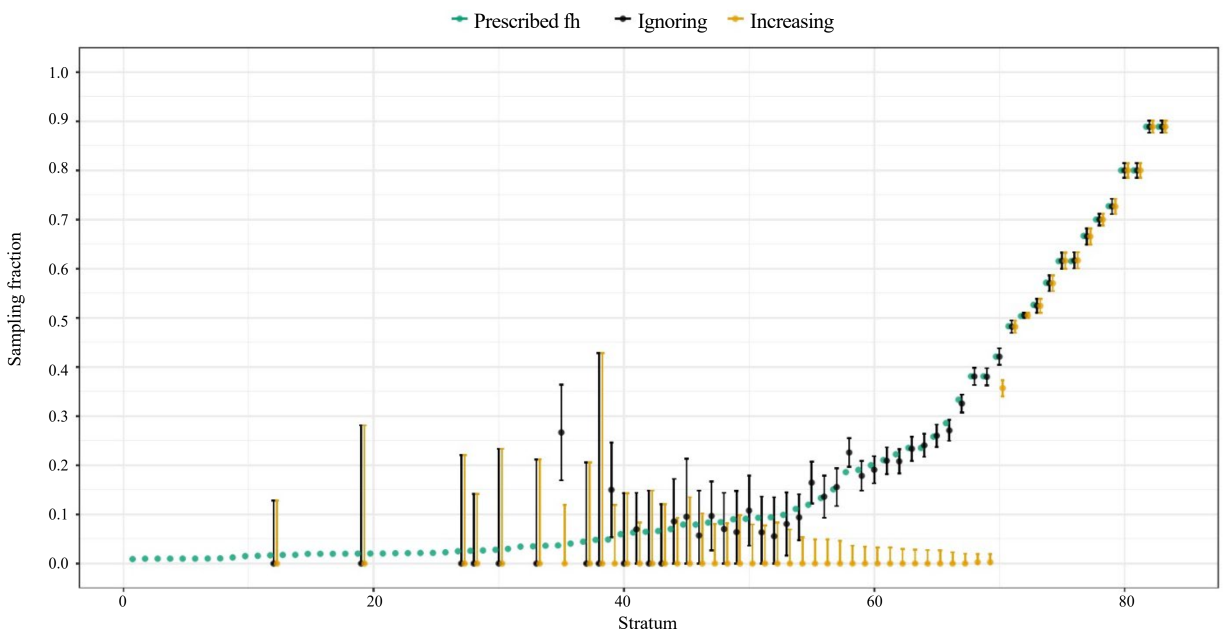
As a final result, we examine the realized sampling fractions in Simulation study 2. By definition (Section 3), the realized sampling fractions $\hat{f}_{h,m}$ per stratum h are equal to the prescribed sampling fractions f_h for all methods m. Figure 4.4 shows the realized sampling fractions $\hat{f}_{\mathrm{HS},h,m}$ in the subpopulation of hotspots together with the 95% confidence intervals (Subsection 3.2). The strata on the x-axis are ordered by the prescribed (or target) sampling fractions f_h. The prescribed sampling fractions are shown with the green points. In strata with very small sampling fractions ($0 < f_h < 0.04$) no hotspots were created in the simulation. Consequently, sampling fractions or margins cannot be computed.

When hotspots are ignored, the 95% confidence intervals do not cover the prescribed sampling fractions in four strata (indices 35, 39, 55 and 58). For increasing the fractions, the realized fractions are too low in 23 strata (indices 47-48 and 50-70). They correspond mainly to strata with medium sampling fractions ($0.08 \leq f_h \leq 0.42$). For ASCP sampling and Using substrata there is only one stratum where the fraction f_h is outside the 95% confidence interval (indices 35 and 12, respectively), which is likely due to chance.

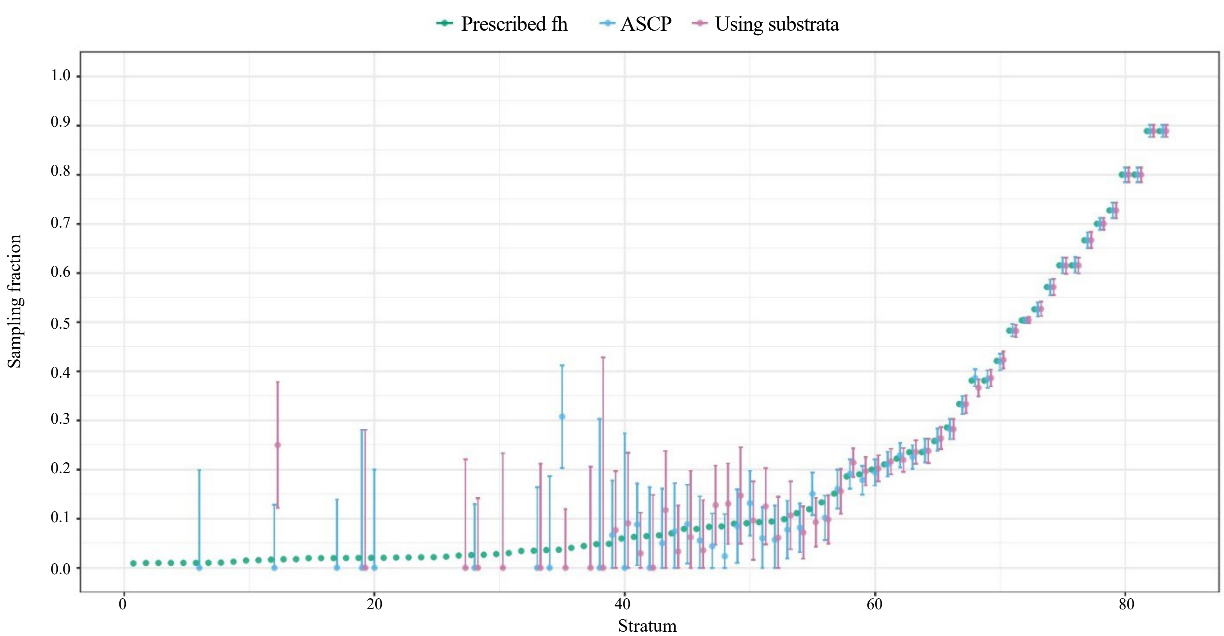
The results in the subpopulation of non-hotspots are similar and are not shown here. Increasing sampling fractions now leads to sampling fractions that are too high, mainly in strata with medium sampling fractions ($0.19 \leq f_h \leq 0.42$).

From Simulation study 2 we conclude that in a specific range of small sampling fractions a reduction in the number of sampled hotspots can be achieved by increasing the fractions, but with a risk of introducing selectivity in the drawn samples. For larger values of sampling fractions no reduction in the number of sampled hotspots is obtained by this method, but there is a risk of worsening the spread of response burden. ASCP and Using substrata do not generally reduce the expected number of sampled hotspots, but realize a more even spread of response burden without introducing selectivity in the drawn samples. ASCP sampling and Using substrata lead to similar results, where ASCP sampling results in a more even spread in all considered situations.

Figure 4.4 Realized sampling fractions $\hat{f}_{\mathrm{HS},h,m}$ with 95% confidence intervals in Simulation study 2 in the subpopulations of hotspots per stratum h and method m and the prescribed sampling fractions f_h per stratum h



(a) Ignoring hotspots and Increasing fractions.



(b) ASCP sampling and Using substrata.

Note: Adapted Spatially Correlated Poisson (ASCP).

5. Discussion and conclusion

In this paper, a new sample coordination method based on Adapted Spatially Correlated Poisson (ASCP) sampling has been introduced that focuses on small businesses with a high response burden (hotspots) and does not use Permanent Random Numbers (PRNs). This method (ASCP hotspot strategy) is compared in two simulation studies with a stratified approach (Using substrata) and a pragmatic method in which sampling fractions are manually adjusted (Increasing fractions). All methods aim to better spread the high response burden of small businesses. Regardless of the simulation results discussed below, we emphasize the difference in the inferential approach between the pragmatic method and the ASCP method. The ASCP hotspot strategy respects the prescribed inclusion probabilities of the sample design and allows for direct statistical inference. The pragmatic method, in particular, implies selectivity issues that need to be addressed before drawing statistical conclusions. Here, zero inclusion probabilities are assigned to the hotspots, and the prescribed inclusion probabilities of other units are also no longer respected. The simulations are based on the Research & Development Survey of Statistics Netherlands conducted in 2020.

In the first simulation study, a practical situation that generally occurs at NSIs or other survey agencies is considered. The methods are applied to an initial given sample with the aim to replace non-desired units (hotspots) in this sample with desired units (non-hotspot units) as much as possible. The simulation study shows that in strata with sufficient non-hotspot units in the population, i.e., when the sampling fraction does not exceed the population fraction of non-hotspot units, the largest number of hotspots can be replaced by increasing the fractions, compared to ASCP sampling and Using substrata. In that case, if there are enough units in the population to replace hotspots, then a large proportional reduction of the number of hotspots can be achieved.

The second simulation study investigates the longer-term effects of the considered methods. A series of monthly draws with a length of 25 years is simulated, where the hotspot status is updated monthly for all the businesses in the population. This simulation study shows that for a specific range of small sampling fractions a reduction in the number of sampled hotspots can be achieved by increasing the fractions. However, there is a risk of introducing selectivity in the drawn samples, i.e., that selective groups of units become over- or underrepresented in the sample, which may cause biased estimates. For other values of the sampling fractions no reduction in the number of sampled hotspots is obtained by increasing the fractions and there is a risk of worsening the spread of the response burden. This is because by replacing hotspots by non-hotspot units, more units in the population are becoming hotspots. At a certain point, there is no room left to replace the sampled hotspots by non-hotspots in the population. ASCP sampling and Using substrata do not reduce the expected number of sampled hotspots, but realize a more even spread of the response burden without introducing selectivity in the drawn samples. In this simulation study, ASCP sampling and Using substrata lead to similar results regarding the number of sampled hotspots, while ASCP sampling results in the best spread in all considered situations.

In general, we conclude that the pragmatic method of increasing the sampling fractions can substantially reduce the number of non-desired units in a given sample, when there are sufficient other units in the population. However, for the longer term there is no guarantee that the number of hotspots will remain at a low level using this method. Moreover, there is a risk of creating selectivity in the samples and increasing it through applying sample coordination. It is therefore recommended to use this method only for strata with small sampling fractions, so that only a few hotspots are replaced by other units and the risk of introducing bias is limited. The ASCP hotspot strategy shows, that it indeed reduces the variance of the number of sampled hotspots and prevents the occurrence of large numbers of hotspots in the sample. Using substrata provided similar results. Although the ASCP hotspot strategy is more generally applicable and avoids rounding problems compared to stratification, the method of using substrata might be easier to implement in an existing sample coordination system that an NSI or survey agency uses. Accordingly, for business surveys with a stratified sample design, the use of substrata might therefore be a good alternative. A drawback of using substrata could be that in case of a detailed stratification, due to rounding, the prescribed inclusion probabilities are no longer realized. We did not encounter this in the simulation studies.

We conclude this paper with some remarks and recommendations. First, the simulation studies in this paper are based on one business survey of Statistics Netherlands. However, results can be generalized to other business surveys with a stratified sample design and when successive samples are drawn for surveys with different stratifications. In the latter case a basic stratification could be defined by crossing the strata of the separate surveys. Second, if a structural reduction of the number of non-desired units such as hotspots is required, the sampling fractions of the surveys need to be adjusted. It is furthermore recommended to update the allocations and stratifications of the surveys regularly, so that the required sample size can be used as efficiently as possible. Finally, the hotspot status of the businesses at SN is used as measure for the response burden in the simulation studies. The hotspot status is based on the number of surveys for which a business has been selected within the last 12 months. Other measures for the response burden, e.g., weights based on the completion time of the questionnaire, would most likely not lead to other conclusions.

Disclaimer

The views expressed in this paper are those of the authors and do not necessarily reflect the policies of Statistics Netherlands.

Acknowledgements

The authors thank two anonymous referees and an Associate Editor for careful reading and providing useful comments on former drafts of this manuscript.

References

Bavdaž, M., Giesen, D., Černe, S.K., Loöfgren, T. and Raymond-Blaess, V. (2015). Response burden in official business surveys: Measurement and reduction practices of national statistical institutes. *Journal of Official Statistics*, 31(4), 559-588.

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. *Proceedings of the Survey Research Methods Section*, American Statistical Association, 35, 35-40. Alexandria ,VA.

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

CBS (2022). Onderzoeksomschrijvingen: Research and development. Available at: https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksomschrijvingen/research-development.

Erikson, J., Giesen, D., Houben, L. and Saraiva, P. (2023). Managing response burden for official statistics business surveys – Experiences and recent developments at Statistics Netherlands, Statistics Portugal, and Statistics Sweden. In *Advances in Business Statistics, Methods and Data Collection*, (Eds., G. Snijkers, M. Bavdaž, S. Bender, J. Jones, S. MacFeely, J.W. Sakshaug, K.J. Thompson and A. van Delden), Hoboken: John Wiley & Sons, Inc., 193-223.

Ernst, I. (1999). The maximization and minimization of sample overlap problems: A half century of results. *Proceedings of the International Statistical Institue*, 52, 168-182. Helsinki, Finland.

Giesen, D., Vella, M., Brady, C.F., Brown, P., Ravindra, D. and Vaasen-Otten, A. (2018). Response burden management for establishment surveys at four national statistical institutes. *Journal of Official Statistics*, 34(2), 397-418.

Gommans, G., Burgerjon, L. and Paulissen, R. (2022). Hotspots. Internal CBS memo, Statistics Netherlands.

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

Grafström, A., Lisic, J. and Prentius, W. (2024). *BalancedSampling: Balanced and Spatially Balanced Sampling*. R package version 2.1.1. Available at: https://CRAN.R-project.org/package=BalancedSampling.

Lorenc, B., Kloek, W., Abrahamsson, L. and Eckman, S. (2013). Description of actual response burden and response behaviour with swedish register. In *Comparative Report on Integration of Case Study Results Related to Reduction of Response Burden and Motivation of Businesses for Accurate Reporting*, (Eds., D. Giesen, M. Bavdaž and I. Bolko). *BLUE-ETS deliverable 8.1*, 11-27.

Matei, A. and Smith, P.A. (2023). Sample coordination methods and systems for establishment surveys. In *Advances in Business Statistics, Methods and Data Collection*, (Eds., G. Snijkers, M. Bavdaž, S. Bender, J. Jones, S. MacFeely, J.W. Sakshaug, K.J. Thompson and A. van Delden), Chichester, England, UK: John Wiley & Sons, Inc., 637-657.

Matei, A., Smith, P.A., Smeets, M.J.E. and Klingwort, J. (2023). Targetted double control of burden in multiple surveys. *Survey Methodology*, 49(2), 363-384. Available at: https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2023002/article/00010-eng.pdf.

Ohlsson, E. (1995). Coordination of samples using permanent random numbers. In *Business Survey Methods*, (Eds., B.G. Cox, D. Binder, B. Chinnapa, A. Christianson, M. Colledge and P. Kott), New York: John Wiley & Sons, Inc., USA, Chapter 9, 153-169.

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.

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

Smeets, M. and Klingwort, J. (2023). Sampling methods to better spread the response burden for small businesses. CBS Discussion paper, Statistics Netherlands. Available at: https://www.cbs.nl/en-gb/background/2023/38/sampling-methods-to-better-spread-the-response-burden-for-small-businesses.

Strategische commissie betere regelgeving bedrijven (2020). Onevenredige enquêtebelasting van specifieke ondernemers in het MKB. Internal CBS memo, Statistics Netherlands.

Yan, T., Fricker, S. and Tsai, S. (2019). Response burden: What is it and what predicts it? In *Advances in Questionnaire Design, Development, Evaluation and Testing*, (Eds., P. Betty, D. Collins, L. Kaye, J.L. Padilla, G. Willis and A. Wilmot), Hoboken: John Wiley & Sons, Inc., 193-212.

Yan, T. and Williams, D. (2022). Response burden – Review and conceptual framework. *Journal of Official Statistics*, 38(4), 939-961.