Discussion Paper

A widely applicable sampling method for spreading the survey burden in business surveys

Harm Jan Boonstra
Marc Smeets

April 28, 2026

Abstract

Due to the heavy (perceived) survey burden on businesses, Statistics Netherlands has started a pilot to prevent as much as possible individual businesses being sampled disproportionally often. Small businesses with a disproportionate high survey burden (“hotspots”) are, as far as possible, excluded from a new sample or replaced by other units. An important drawback of this pragmatic approach is that there is a risk of biased estimates. Statistics Netherlands operates a sampling module for business surveys, which allows to spread the survey burden among businesses without the risk of bias. A reduction of the number of hotspots could be achieved by spreading the survey burden across more surveys using the sampling module. However, to spread the survey burden across surveys the current Survey Burden Sampling (SBS) method in the sampling module requires that stratifications of these surveys are aligned, which is difficult in practice. This paper investigates a new sampling method, Rotating Random Number Sampling (RRNS), similar to SBS, but with the advantage that stratifications do not need to be aligned. We examine different versions of RRNS in a simulation study, where the survey burden is spread across multiple surveys including both cross-sectional and rotating panel surveys and also dynamic populations are considered. When RRNS is applied to cross-sectional surveys using different stratifications, RRNS improves the spread of the survey burden and results in fewer hotspots than SBS. With harmonized stratifications, RRNS and SBS give similar results. For panel surveys, RRNS seems to work less well than SBS, probably because RRNS does not track the built-up survey burden and is therefore unable to account for the panel membership duration. It is recommended to add RRNS to the sampling module and to apply RRNS widely across cross-sectional surveys. For panel surveys, RRNS needs to be further developed before it can be implemented.

Keywords: rotating panels, permanent random numbers, dynamic populations, hotspots.

Contents

1 Introduction

1.1 Background

A common challenge in the production of official statistics is the requirement to reduce the survey burden placed on businesses. National statistical institutes use two main strategies to control the survey burden. The first strategy reduces the total survey burden by minimizing questionnaire length or frequency and maximizing the use of administrative data. The second strategy involves spreading this burden evenly across respondents through sample coordination (Ernst 1999; Giesen et al. 2018; Erikson et al. 2023). Sample coordination aims to control the size of the overlap between successively drawn samples in contrast to the case where the samples are drawn independently. Positive coordination attempts to maximize the overlap to obtain more precise estimates of change, which is used in panel surveys. 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 spreading the response burden (Bradburn 1978).

Statistics Netherlands (SN) uses a sampling module for the business surveys, which allows to spread the survey burden within groups of surveys (Smeets and Boonstra 2018). In this way, the total survey burden is spread among the individual businesses in the common sampling frame of the surveys in the group. Despite the applied spread of survey burden, in practice there are still businesses that are regularly heavily sampled. This is mainly because of the (often necessary) high sampling fractions and detailed stratifications of the surveys. Another reason is that the sampling module is not widely applied across the different surveys. In order to spread the survey burden across different surveys effectively, the sampling module requires stratifications of these surveys to be aligned with each other. In practice, coordinating sample designs turns out to be very difficult.

The heavy survey burden faced by individual businesses has led to complaints from the business community (Strategische commissie betere regelgeving bedrijven 2020). In response, SN has started a pilot to prevent, as far as possible, businesses from being sampled disproportionally often (Gommans et al. 2022; Ippel et al. 2023). This pilot focuses primarily on small and medium-sized businesses with 0-19 working persons. Businesses with 0-9 working persons (size classes 0-3) 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 (size class 4) are labeled as “hotspots” if they have been selected for more than 4 surveys within the last 12 months. Various strategies were developed during the pilot to remove the resulting hotspots from the sample as far as possible or to potentially replace hotspots with other businesses before the final sample is released. These are pragmatic approaches, in which hotspots are removed from the sample in consultation with the statistical departments.

At the moment, the sampling module allows to spread the survey burden within groups of surveys by assigning a Permanent Random Number (PRN) to the businesses and by tracking the cumulated survey burden of the businesses based on the number of surveys for which a business was selected. It is also possible to take into account that the survey burden differs depending on the survey, for example, based on the time it takes to complete a questionnaire. More specifically, the businesses are sorted increasingly first by the cumulated survey burden and subsequently by the PRNs. The first businesses in this ordering are then selected for the sample. This method is called Survey Burden Sampling (SBS) and is described in full detail in Boonstra and Smeets (2015, 2018) and Smeets and Boonstra (2018).

In Boonstra et al. (2024), a new sampling method based on rotating random numbers has been developed which is referred to as RRNS (Rotating Random Number Sampling). After a sample is drawn, the random numbers assigned to the businesses are rotated in such a way that the selected businesses have a higher random number than the businesses that have not yet been selected. The first businesses in the (ascending) order based on the random numbers are then selected for the sample. A drawback of this method is that it does not allow for different survey burden values between surveys. On the other hand, the stratifications of surveys need no longer to be aligned, implying that, in practice, RRNS can easily be applied more widely across surveys. Boonstra et al. (2024) demonstrates through a simulation study involving multiple surveys, that when the surveys do not all use the same stratification, RRNS spreads the survey burden more effectively than SBS. Furthermore, the number of drawn hotspots in strata with low sampling fractions is lower with RRNS. If the stratifications are identical and SBS is applied across the surveys, both methods result in roughly the same spread of survey burden and more or less the same amount of hotspots in the sample. In the simulation study only (repeated) cross-sectional surveys were considered and no account had been taken of population dynamics, like births, deaths and stratum movers.

In this paper, we describe and examine the new sampling method RRNS in more detail. We consider three versions of this method, which differ in the way random numbers are generated or retained. In a simulation study, we investigate which variant works best under different conditions, for example, with dynamic populations, we compare RRNS with SBS and also with Simple Random Sampling (SRS) in which samples are drawn completely independently from each other. We also examine how the method can be generalized to rotating panel surveys.

1.2 Related work

In Matei et al. (2023), a sampling method was developed that applies spreading the survey burden specifically to so-called undesirable units. These are units that one would wish to exclude from the sample, such as businesses with a high survey burden or hotspots. This method, known as Adapted Spatially Correlated Poisson (ASCP) sampling, also makes use of PRNs and is based on correlated Poisson sampling (Bondesson and Thorburn 2008). By drawing balanced samples, the creation of clusters of undesirable units within the samples is prevented. In Matei et al. (2023), ASCP sampling was also applied to hotspots, and a simulation demonstrated that ASCP sampling leads to a better spread of survey burden across hotspots than standard (coordinated) sampling methods such as Poisson sampling with PRNs (Brewer et al. 1972) and Pareto order sampling (Rosén 1997).

In Smeets and Klingwort (2023), ASCP sampling was compared with the pragmatic strategies used in the SN pilot study. A special case of ASCP sampling was also examined, in which separate substrata are defined for the hotspots and the non-hotspots, and a random sample with the same sampling fraction is drawn in each substratum. This approach also prevents clusters of hotspots in the sample. Simulations have shown that the use of substrata and ASCP sampling leads to a better spread of the survey burden for hotspots than the pragmatic methods. The pragmatic methods result in fewer hotspots on average in the sample, but it is also shown in Smeets and Klingwort (2023) and Klingwort and Smeets (2023) that these pragmatic strategies carry risks and that removing hotspots from the sample outside the sampling module can lead to serious bias in the statistical results.

Boonstra et al. (2024) investigated whether extending SBS with ASCP sampling or the use of substrata leads to a better spread of survey burden among businesses, and in particular among hotspots. By combining SBS with ASCP sampling or substrata, sampling takes into account not only the built-up survey burden but also the hotspot status of the businesses. However, various simulation studies indicated that the SBS method still provides the best spread of survey burden and generally results in slightly fewer hotspots than the combined methods that also use hotspot status. In addition, these extensions have the same drawback as SBS. Since these methods can all take into account different response loads per survey, it is required that the stratifications of surveys within a group are aligned to ensure that the spread of survey burden across surveys is effective.

The remainder of the paper is structured as follows. Section 2 describes the sampling methods considered. Section 3 details the simulation study setup. Section 4 presents the results and Section 5 provides a discussion.

2 Methods

In this section, we describe the sampling methods examined in the simulation study. Survey Burden Sampling is the method currently used by SN. This method is implemented in the sampling module of SN and in particular in R package SBS. Survey Burden Sampling is compared with a new method called Rotating Random Number Sampling (RRNS). This method was introduced in Boonstra et al. (2024) and can be seen as a simplification of Survey Burden Sampling. In the simulation, we examine three versions of this new sampling method. We also consider Simple Random Sampling (SRS) as the reference method in the simulation. SRS always draws a new (simple random) sample independently from previous samples and does therefore not spread the survey burden among individual businesses. An overview of the sampling methods considered is given in Table A.1 in the appendix.

2.1 Survey Burden Sampling

Here we provide only a general description of Survey Burden Sampling (SBS). For a more detailed description we refer to Smeets and Boonstra (2018) or Boonstra and Smeets (2015, 2018). SBS is based on groups of surveys (SBS groups) with a common sampling frame and spreads the survey burden across the surveys within a group, as well as over time. The surveys in a group must use a common stratification, the so-called basic stratification. It is possible to deviate from the basic stratification in a group by using substrata, although this leads to a less optimal spread of the survey burden.

In its basic form, SBS tracks two sampling variables for each unit (or business) in the common sampling frame of an SBS group: a Permanent Random Number (PRN) and a Survey Burden (SB) variable. The PRNs are uniformly and independently drawn from $[0, 1]$ and are assigned to each unit once. The SB variable corresponds in the simplest case to the built-up survey burden, that is, the number of times a unit has been drawn in a sample, possibly weighted with a measure reflecting the survey burden caused by different surveys. Every SBS group has its own PRN and SB variables. For rotating panel surveys in an SBS group, also a panel membership indicator is kept for the units in the sampling frame.

The sampling for (repeated) cross-sectional surveys in an SBS group is done per basic stratum. The units are sorted increasingly first by the SB variable and subsequently by the PRNs. For a given sample size n, the first n units in this ordering are then selected for the sample. After drawing the sample, the SB variables of the units in the basic stratum are updated. The sampling for a rotating panel survey in an SBS group is done in a similar way and is also done per basic stratum. First, the panel indicator is used to determine which units belong to the panel. Then both the units in the panel and the units not belonging to the panel are separately sorted increasingly first by the SB variable and subsequently by the PRNs. Based on the rotation fraction it is determined how many units should be added to or removed from the panel. If m units need to be added to the panel, the first m units not belonging to the panel in the given ordering are added to the panel. If m units need to be removed from the panel, then the m last panel units in the given ordering are removed from the panel. After drawing the sample, both the panel indicator and the SB variable of the units in the stratum are updated.

Survey Burden Sampling can also take population dynamics into account, like deaths, births or (basic) stratum movers, without any units becoming systematically over- or underrepresented in the drawn samples. Deaths can easily be removed from the sampling frame. Before a new sample is drawn in an SBS group, suitable values of the PRN, the SB variable and the panel indicator are assigned to births and stratum movers. For births, a new PRN is uniformly selected from [0,1] and appropiate values for the SB variable and the panel indicator are chosen. These values are obtained by copying them from the existing unit in the basic stratum whose PRN is closest to the PRN of the new unit. For stratum movers, the relative position in the basic stratum is taken over to the new basic stratum. The relative position of unit i with rank r_i according to the given ordering of its basic stratum h with N_h units is defined as $r_i/(N_h + 1)$. This is realized by copying the SB value and the panel indicator from the existing unit in the new basic stratum which is closest to the relative position of the stratum mover according to the same ordering. A new and unique PRN is then randomly chosen from the interval of PRNs determined by adjacent units of this existing unit in the new basic stratum. Note that in case of stratum movers the value of PRN is in fact not permanent anymore.

2.2 Rotating Random Number Sampling

Rotating Random Number Sampling can be seen as a simplification of the SBS method. Here, only a random number (RN) variable is used, which is uniformly and independently drawn from $[0, 1]$ and is assigned to each unit. Prior to a draw, the units in a stratum are sorted according to these random numbers. If n_h units are to be drawn in stratum h, then the first n_h units in this sorting are selected, similar to SBS. These n_h units are then placed at the end of the stratum, and the random numbers are reassigned to all units in the stratum in ascending order. In this way, the units rotate after each draw until it is their turn again. Note that the random numbers are not permanently linked to the same units. By using the same RN variable for multiple surveys, the survey burden is spread over all these surveys. In addition to being simpler than the SBS method, a major advantage of this method is that the stratifications of the different surveys do not need to be aligned. For SBS, this is required to prevent selective sampling. A disadvantage compared to SBS is that differences in the survey burden between different surveys cannot be taken into account because no SB variable is used. Since the method can also be interpreted in terms of “rotating random numbers”, we will refer to this method as Rotating Random Number Sampling (RRNS). In particular, let $R_1, \dots, R_{N_h}$ be the current values of the RN variable sorted in ascending order, of the N_h units in stratum h. Then the units with random numbers $R_1, \dots, R_{n_h}$ are drawn for the next sample draw. After this draw the random numbers are rotated n_h positions to the right within stratum h, i.e.,

$$R_1, \dots, R_{n_h}, R_{n_h+1}, \dots, R_{N_h} \;\rightarrow\; R_{N_h-n_h+1}, \dots, R_{N_h}, R_1, \dots, R_{N_h-n_h}\,,$$

and reassigned to the units. For a next draw, whether it is using the same stratification or not, the units are again sorted by their current RN values, etc.

For (rotating) panel surveys, the RRNS method works in a similar way. The first draw of a panel survey is done exactly in the same way as a draw for a non-panel survey. For an existing panel, a subsequent draw is taken according to the steps outlined below. Let H denote the total number of strata, U_h the set of N_h population units in stratum h, R_{hk} the value of the RN variable for unit k in stratum h, I_{hk} the indicator of whether unit k belongs to the current panel ($I_{hk} = 1$) or not ($I_{hk} = 0$) in stratum h, and ν_h the rotation fraction for stratum h. Then, for every stratum $h = 1, \ldots, H$:

1. Sort the units in U_h by (i) I_{hk} (decreasing), (ii) R_{hk} (increasing).
2. Let m_h be the number of units in the current panel in stratum h. If rotation is applied, define $u_h = \text{round}(\nu_h m_h)$, otherwise $u_h = 0$. Remove the u_h last units with $I_{hk} = 1$ from the panel. Rounding is done randomly upwards or downwards depending on the exact value.
3. Adjust the panel to obtain the required sample size n_h, which can be different from the current panel size m_h due to population dynamics or changed requirements:
   – If $m_h - u_h < n_h$, add the first $n_h - (m_h - u_h)$ units with $I_{hk} = 0$ to the panel.
   – If $m_h - u_h > n_h$, remove additional $(m_h - u_h) - n_h$ units from the panel (last units with $I_{hk} = 1$).
4. After the previous steps the units in U_h consist of four groups ordered as <remaining in><rotated out><rotated in><remaining out>, where some of the groups may be empty, and each (non-empty) group is still sorted by increasing RN. The sampling variables R_{hk} and I_{hk} are now updated as follows. Set $I_{hk} = 0$ for every unit in <rotated out> and set $I_{hk} = 1$ for every unit in <rotated in>. Then reorder the blocks as <remaining out><rotated out><rotated in><remaining in>, without changing the ordering of units within each block, and reassign all N_h RN random numbers sorted from smallest to largest to the units in this order from left to right.

Steps 1 to 3 are the same steps as used for updating the panel under the SBS sampling method. The last step is clearly different, as instead of updating a built-up burden variable, the random numbers are reassigned to the units in the specified order. By assigning the smallest RN random numbers to the units that remain out of the panel, these units are likely to be the first units selected in a next sample draw, whether it is a sample for the same survey, or another. In step 4, an alternative reordering would be <remaining out><rotated in><rotated out><remaining in>, so with the units rotated into the panel positioned before those rotated out. This would make the probability of a unit just rotated out of the panel to be selected for a subsequent survey slightly smaller. However, the difference would normally be small as the numbers of units rotated in and out are usually very small relative to the population size. Note that with $\nu_h = 1$ one obtains the non-panel version of the RRNS method.

The description so far has assumed that random numbers for the RRNS method are drawn once and never change, except that they are rotated over the sampling units after each draw. However, one might also regenerate the rotating numbers after each draw. In this paper, we consider three versions of Rotating Random Number Sampling (RRNS):

– RRNSa: random numbers are generated only initially and never change, except that they get reassigned to different units after each RRNS sample draw.
– RRNSb: the random numbers are completely regenerated after each draw, uniformly and independently drawn from $[0, 1]$.
– RRNSc: the random numbers are equidistantly regenerated after each draw, i.e., they become $(r_i - u)/N_h$ where r_i runs from 1 to the stratum population size N_h, and u is a single random number uniformly drawn from $[0, 1]$.

The RRNS method deals with population dynamics in basically the same way as the SBS method, except that it is somewhat simpler because no built-up survey burden variable needs to be assigned for births or stratum movers. Deaths can be simply removed from the sampling frame, along with their random number (and possibly panel indicators). For each birth a new random number is generated uniformly and independently from $[0, 1]$, whereas stratum movers keep the random numbers attached to them at that moment. Panel indicators need to be assigned for births in stratum h and reassigned for stratum movers moving to stratum h, to prevent such units from becoming under- or overrepresented in the panel. This is done for each birth or stratum mover by copying the panel indicator(s) from the existing unit in stratum h with random number nearest to that of the birth or stratum mover. If there are no existing units in stratum h the panel indicator values for all births and stratum movers are set to 0. Unlike SBS, no separate 'relative position' variable is needed for RRNS as the random numbers themselves already define a relative position comparable over strata.

3 Simulation setup

We perform simulations in which we repeatedly draw series of T consecutive samples. Drawing these series is repeated R times, each time with an independent initialization of the random numbers. The SB variables are initialized to 0. To investigate how the methods perform under population dynamics, we also simulate births and deaths of population units as well as stratum changes. Subsection 3.1 describes the considered scenarios and Subsection 3.2 discusses the different simulation measures used to compare the sampling methods.

3.1 Scenarios

3.1.1 Single stratified cross-sectional survey for a range of sampling fractions

We first consider the situation in which the different sampling methods are applied to a single stratified cross-sectional survey, with the sampling fractions in the strata varying from low to high. We assume $H = 9$ strata, $h = 1, \dots, H$, with an initial population size of $N_h = 100$ for each stratum and we draw stratified samples with increasing sampling fractions, given by $f_h = 0.1, \dots, 0.9$. The simulation is run over $T = 60$ time periods. In each period a new sample is drawn using one of the sampling methods. To obtain reliable estimates for the simulation measures described in Subsection 3.2, this setup is repeated $R = 120$ times. So, to sum up, we have 120 runs of simulated series of sample draws, each of length 60 periods.

The population dynamics are simulated as follows. In each period of each simulation run, the number of births and deaths simulated, both equal 2% of the current population size of each stratum. In addition, the number of units moving from stratum 1 to 2, 2 to 3, ... , 8 to 9 are set to 1% of the current population size in each stratum. The stratum changes in the other directions, i.e., 2 to 1, 3 to 2, ..., 9 to 8, consist of 0.5% of the current population size of the strata. In this way a population is created whose overall size remains reasonably stable at 900 units. No movements between strata more than two in number apart are simulated, to mimic the situation in realistic business surveys where strata typically move between consecutive size classes.

3.1.2 Multiple surveys, five strata

In order to test the methods in a multi-survey setting, we also set up a simulation study in which consecutive samples are drawn repeatedly for 4 different surveys, each with its own sample design. Here we consider two stratification variables. The most detailed one, Strat1, subdivides the population into 5 strata, and the other stratum variable, Strat2, combines these detailed strata into 3 strata. The first Strat2 stratum combines the first two Strat1 strata, the second Strat2 stratum combines Strat1 strata 3 and 4, and Strat1 stratum 5 is identical to Strat2 stratum 3. An initial population of size 1000 units is generated. See Table 3.1 for this setup as well as the sampling fractions used for the 4 surveys. Note that surveys 1 and 4 draw samples according to the detailed stratification of Strat1, whereas surveys 2 and 3 use the broad stratification of Strat2.

In the simulation with multiple surveys, we apply survey burden sampling in two variants. SBS denotes the variant in which all surveys form their own group and survey burden sampling is applied to the surveys individually. SBS1 denotes the variant in which the four surveys form one SBS group with Strat1 as basic stratification and the survey burden is spread across all four surveys.

<table>
<tr><th>Strat1</th><th>N1</th><th>Fraction1</th><th>Fraction4</th><th>Strat2</th><th>N2</th><th>Fraction2</th><th>Fraction3</th></tr>
<tr><td>1</td><td>700</td><td>0.05</td><td>0.06</td><td rowspan="2">1</td><td rowspan="2">850</td><td rowspan="2">0.04</td><td rowspan="2">0.07</td></tr>
<tr><td>2</td><td>150</td><td>0.12</td><td>0.09</td></tr>
<tr><td>3</td><td>90</td><td>0.25</td><td>0.25</td><td rowspan="2">2</td><td rowspan="2">130</td><td rowspan="2">0.18</td><td rowspan="2">0.20</td></tr>
<tr><td>4</td><td>40</td><td>0.60</td><td>0.75</td></tr>
<tr><td>5</td><td>20</td><td>1</td><td>1</td><td>3</td><td>20</td><td>0.80</td><td>0.75</td></tr>
</table>

Table 3.1 Initial stratum population sizes for stratum variables Strat1 and Strat2, and sampling fractions for 4 surveys used in the simulation study.

The simulation is run over $T = 100$ time periods. In each period a sample for each of the four surveys is drawn according to the stratified design as described. To be able to reliably estimate several measures that indicate whether the realized samples adhere to the sample design as well as various measures of survey burden, this setup is repeated $R = 120$ times. In summary, we have 120 runs of simulated series of sample draws, each of length 100 periods.

One of the objectives of the simulation study is to see how the methods behave under population dynamics. Therefore, we simulate births and deaths of population units, as well as stratum changes. In each period of each simulation run, the number of births and deaths simulated both equal 0.6%, 0.4%, 0.8%, 0.8% and 1%, respectively, of the current population size of each detailed stratum 1 to 5. In addition, the number of units that move from (detailed) stratum 1 to 2, 2 to 3, 3 to 4 and 4 to 5 are set to 1% of the current population size of each stratum 1 to 4, respectively. The stratum movements in the other directions, 2 to 1, 3 to 2, 4 to 3 and 5 to 4 are taken to be 0.5% of the current population size of strata 1 to 4, respectively. No movements between strata more than two in number apart are simulated, to mimic the situation in realistic business surveys where strata typically move between consecutive size classes. This setup yields a dynamic population whose overall size remains reasonably stable at 1000 units, and where detailed strata 2 to 5 slowly increase in size over time, at the cost of the size of stratum 1.

3.1.3 Multiple surveys including a rotating panel survey, five strata

The next simulation considers the combination of a rotating panel survey with two non-panel surveys. The rotation fraction of the rotating panel survey is taken to be 10%, so that approximately 10% of the panel units are rotated out each (simulated) period and roughly the same number of non-panel members are added to the panel. In addition to this active form of panel rotation, the panel is further (slightly) updated in each period to adjust for population dynamics, so that newborn units and units that moved in from other strata are appropriately represented. Apart from the panel survey the simulation setup is similar to that of Section 3.1.2. We also make a distinction between SBS and SBS1, where SBS spreads the survey burden only over time for the separate surveys and SBS1 spreads the survey burden over time and across all three surveys.

An initial population of size 1000 units is generated, and details of the initial stratum population sizes and sampling fractions are given in Table 3.2, where in this case the first survey with sampling fractions given in the Fraction1 column corresponds to the panel survey. Fraction2 and Fraction3 are the sampling fractions of the cross-sectional surveys, which have been renumbered and do not necessarily correspond to the same surveys as in Table 3.1. The same population dynamics as described in Section 3.1.2 are used for this simulation too, and the simulation size is also the same with $R = 120$ simulation runs, each simulating sample draws for the three surveys in each of $T = 100$ consecutive periods.

<table>
<tr><th>Strat1</th><th>N1</th><th>Fraction1</th><th>Fraction3</th><th>Strat2</th><th>N2</th><th>Fraction2</th></tr>
<tr><td>1</td><td>700</td><td>0.05</td><td>0.06</td><td rowspan="2">1</td><td rowspan="2">850</td><td rowspan="2">0.04</td></tr>
<tr><td>2</td><td>150</td><td>0.12</td><td>0.09</td></tr>
<tr><td>3</td><td>90</td><td>0.25</td><td>0.25</td><td rowspan="2">2</td><td rowspan="2">130</td><td rowspan="2">0.18</td></tr>
<tr><td>4</td><td>40</td><td>0.60</td><td>0.75</td></tr>
<tr><td>5</td><td>20</td><td>1</td><td>1</td><td>3</td><td>20</td><td>0.80</td></tr>
</table>

Table 3.2 Initial stratum population sizes for stratum variables Strat1 and Strat2, and sampling fractions for 3 surveys, of which the first is a rotating panel survey.

3.2 Simulation measures

We compare the sampling methods using a number of measures. First, we check whether the sampling methods draw units in the simulation with the specified inclusion probabilities. Next, we examine the effects on spreading the survey burden over the units in the simulation.

3.2.1 Realized inclusion probabilities and coverage

The realized inclusion probability of unit k in the simulation can be estimated for each period $1 \leq t \leq T$ in the following way:

$$\hat{\pi}_k(t) = \frac{1}{R_{kt}} \sum_{r=1}^{R_{kt}} \iota\{k \in S_r(t)\}. \tag{1}$$

Here, $\iota\{k \in S_r(t)\} \in \{0, 1\}$ indicates whether unit k in simulation run r and in period t is selected in sample $S_r(t)$ and $R_{kt} = \sum_{r=1}^{R} \iota\{k \in U_r(t)\}$ denotes the number of simulation runs in which unit k exists in the simulated population $U_r(t)$ in period t. If unit k is drawn with inclusion probability π_k the indicator $\iota\{k \in S_r(t)\}$ follows a Bernoulli distribution with expectation π_k and variance $\pi_k(1 - \pi_k)$. It follows from Equation (1) that $\hat{\pi}_k(t)$ has expectation π_k and variance $\pi_k(1 - \pi_k)/R_{kt}$. In order to obtain one estimate for each unit k, we average the probabilities $\hat{\pi}_k(t)$ over the T periods:

$$\hat{\pi}_k = \frac{1}{T} \sum_{t=1}^{T} \hat{\pi}_k(t). \tag{2}$$

The estimate $\hat{\pi}_k$ should have expectation π_k and, in the case of independent sample draws, variance $\pi_k(1 - \pi_k)/ \sum_{t=1}^{T} R_{kt}$.

For every sampling method, the coverage in the simulation is computed as the fraction of units in the strata $1 \leq h \leq H$ for which the realized inclusion probability lies within an approximate 95%

confidence interval:

$$\text{coverage}_h = \frac{1}{N_h^*} \sum_{k=1}^{N_h^*} \iota \left\{ |\hat{\pi}_k - \pi_k| < 2\sqrt{\pi_k(1-\pi_k) / \sum_{t=1}^{T} R_{kt}} \right\}. \tag{3}$$

In Equation (3) the sum is taken over all units k that have been in stratum h at least once during the simulation and N_h^* is the total number of these units in the simulation.

We also compute the realized sampling fractions per basic stratum h, given by

$$\hat{f}_h = \sum_{r=1}^{R} \sum_{t=1}^{T} n_{hr}(t) \Bigg/ \sum_{r=1}^{R} \sum_{t=1}^{T} N_{hr}(t), \tag{4}$$

where $N_{hr}(t)$ and $n_{hr}(t)$ are respectively the number of units in the simulated population and the number of sampled units in stratum h in simulation run r and in period t. The estimated fractions $\hat{f}_h$ have expectation $f_h = \pi_k$ and variance $\mathsf{V}(f_h) = \pi_k(1-\pi_k)/\sum_{r=1}^{R}\sum_{t=1}^{T} N_{hr}(t)$ under independence. We also consider the margin of the approximate 95% confidence interval of $\hat{f}_h$, given by

$$\text{error}_h = 2\sqrt{\mathsf{V}(\hat{f}_h)}. \tag{5}$$

The realized inclusion probabilities, the sampling fractions and the coverage are computed for the overall population and also for the subpopulations of births and stratum movers.

3.2.2 Measures of survey burden

The survey burden of a unit can be measured in different ways. We consider the total survey burden of a unit and the mean duration of a survey leave of a unit. For the simulations with multiple surveys we also look at stacking, i.e., the simultaneous inclusion of units in multiple surveys. And for panels the mean durations of the panel memberships of the units are considered. And finally, we examine the percentage of hotspots in the population and the number of sampled hotspots as a function of time. In the simulation, we consider a unit to be a hotspot when this unit has been selected at least 5 times for one of the surveys in the past twelve periods. The total survey burden and the mean duration of survey leaves and panel memberships are only computed for units that live in a single stratum over the entire simulation period, the so-called stable units. So we exclude births, deaths and stratum movers from these calculations. For the other measures, we consider all units in the population.

The total survey burden of a unit is defined as the number of times a unit is sampled in the series of T periods and is computed for unit k and simulation run r:

$$\text{SB}_{kr} = \sum_{t=1}^{T} \iota\{k \in S_r(t)\}. \tag{6}$$

We are interested in the spread of the total survey burden over the stable units in the simulation. The spread of the total survey burden in stratum h is computed as follows:

$$\sigma\text{SB}_h = \frac{1}{R}\sum_{r=1}^{R} \sigma_{r,\text{SB}}, \text{ with } \sigma_{r,\text{SB}} = \sqrt{\frac{\sum_{k=1}^{N'_{hr}} (\text{SB}_{kr} - \mu_{r,\text{SB}})^2}{N'_{hr} - 1}} \text{ and } \mu_{r,\text{SB}} = \frac{1}{N'_{hr}} \sum_{k=1}^{N'_{hr}} \text{SB}_{kr}. \tag{7}$$

Here, N'_{hr} is the number of stable units in stratum h in simulation run r.

The duration of a survey leave is defined as the number of consecutive periods between two draws in any survey. We consider the mean duration of the survey leaves of unit k in the entire simulation period, given by

$$\mathsf{SL}_{kr} = \text{the mean duration of the survey leaves of unit } k \text{ in run } r. \tag{8}$$

Regarding the survey leave, we are interested both in the mean duration in the simulation and in the spread of the durations over the units in the simulation. The mean duration of the survey leave in stratum h is given by

$$\mathsf{SL}_h = \frac{1}{R}\mu_{r,\mathsf{SL}}, \text{ with } \mu_{r,\mathsf{SL}} = \frac{1}{N'_{hr}}\sum_{k=1}^{N'_{hr}} \mathsf{SL}_{kr}, \tag{9}$$

and the spread by

$$\sigma\mathsf{SL}_h = \frac{1}{R}\sum_{r=1}^{R} \sigma_{r,\mathsf{SL}}, \text{ with } \sigma_{r,\mathsf{SL}} = \sqrt{\frac{\sum_{k=1}^{N'_{hr}}(\mathsf{SL}_{kr} - \mu_{r,\mathsf{SL}})^2}{N'_{hr} - 1}}. \tag{10}$$

The percentage of units in stratum h having a hotspot status in the population is computed as follows:

$$\mathsf{pc}_h^{\mathsf{HS}} = \frac{1}{RT}\sum_{r=1}^{R}\sum_{t=1}^{T} \mathsf{pc}_{hr}^{\mathsf{HS}}(t), \tag{11}$$

where $\mathsf{pc}_{hr}^{\mathsf{HS}}(t)$ is the percentage of hotspot units in basic stratum h in run r and period t.

In addition, the simulation average number of sampled hotspots in stratum h is considered as a function of time:

$$\mathsf{n}_h^{\mathsf{HS}}(t) = \frac{1}{R}\sum_{r=1}^{R} \mathsf{n}_{hr}^{\mathsf{HS}}(t), \tag{12}$$

where $\mathsf{n}_{hr}^{\mathsf{HS}}(t)$ is the number of sampled hotspot units in stratum h in run r and period t.

For the simulations with multiple surveys, stacking is defined as the number of times that units are drawn simultaneously in two or more surveys. We consider the total number of times in the simulation that units in stratum h are drawn simultaneously in two or more surveys. Finally, for panels the overall distribution in the simulation of the durations of the panel memberships is considered for the units that exist in stratum h over the entire simulation period.

4 Results

4.1 Single stratified cross-sectional survey for a range of sampling fractions

This section discusses the results of the first simulation study, which is described in Section 3.1.1.

4.1.1 Realized sampling fractions and coverage

We first check whether the simulated sampling methods draw samples with the specified sampling fractions. Table 4.1 shows the realized overall sampling fractions according to (4) for all strata and sampling methods. The third column shows the corresponding error margin given

stratum	fraction (%)	error	SRS	SBS	RRNSa	RRNSb	RRNSc
1	10	0.076	9.993	10.000	9.990	10.003	9.999
2	20	0.096	20.003	20.003	20.007	20.002	19.995
3	30	0.109	30.007	29.994	30.004	29.999	29.997
4	40	0.116	40.007	40.002	39.996	40.005	39.996
5	50	0.119	50.000	50.000	50.003	49.994	49.987
6	60	0.116	59.999	60.006	60.006	59.997	59.998
7	70	0.109	70.001	69.997	70.003	70.003	70.000
8	80	0.094	80.002	79.999	80.000	80.004	79.999
9	90	0.067	89.998	89.997	90.002	89.996	89.995

Table 4.1 Realized overall sampling fractions (in %) for the different methods by survey and stratum

by (5). We see that the realized sampling fractions are close to the specified sampling fractions, for every sampling method and every stratum.

Tables 4.2 and 4.3 show the realized sampling fractions with the corresponding margins for the subpopulations of births and stratum movers, respectively. Due to the relatively small number of births and stratum movers in each period, we see that the realized fractions for all methods and in all strata deviate further from the specified sampling fractions compared to the overall population. For all sampling methods except RRNSa the realized fractions match the design fractions, given the estimated standard errors. For method RRNSa, the fractions deviate further from the specified fractions than expected based on the errors, but the fractions are not systematically overestimated or underestimated.

stratum	fraction (%)	error	SRS	SBS	RRNSa	RRNSb	RRNSc
1	10	0.550	10.131	10.342	8.962	10.089	9.668
2	20	0.690	19.951	20.434	18.583	20.070	19.862
3	30	0.783	30.156	29.945	28.894	29.726	29.616
4	40	0.836	40.160	39.242	37.915	40.262	39.329
5	50	0.855	50.004	50.318	50.450	50.070	49.521
6	60	0.836	59.707	60.384	62.225	60.930	58.951
7	70	0.783	70.127	70.360	68.610	70.492	70.134
8	80	0.679	79.743	80.522	81.062	80.003	80.356
9	90	0.479	89.854	90.423	91.131	89.765	90.321

Table 4.2 Realized sampling fractions (in %) for births for the different methods by survey and stratum

stratum	fraction (%)	error	SRS	SBS	RRNSa	RRNSb	RRNSc
1	10	1.046	10.036	8.668	7.664	10.371	9.641
2	20	0.827	19.953	20.391	19.589	19.835	19.985
3	30	0.913	30.118	30.774	30.625	30.069	30.019
4	40	0.970	40.137	39.598	38.726	40.519	40.480
5	50	0.990	50.054	49.966	48.966	50.044	49.975
6	60	0.972	59.191	60.284	61.987	59.447	60.540
7	70	0.903	70.782	70.257	66.722	70.597	69.908
8	80	0.774	79.991	80.056	80.084	80.150	80.636
9	90	0.723	89.810	89.940	90.753	90.318	90.303

Table 4.3 Realized sampling fractions (in %) for stratum movers for the different methods by survey and stratum

Table 4.4 shows the overall coverage per stratum as defined in (3). As expected, the coverage for SRS is close to 0.95 in all strata. The coverage for the other sampling methods is always larger than 0.95. The reason for this is that, due to the spread of the survey burden, units in a stratum are selected with equal frequency as much as possible. The highest coverage rates are achieved by SBS, suggesting that the best spread of survey burden is obtained by SBS. The coverage rates for the three RRNS methods are similar.

stratum	SRS	SBS	RRNSa	RRNSb	RRNSc
1	0.969	1.000	0.969	0.970	0.974
2	0.958	0.999	0.983	0.986	0.983
3	0.955	0.999	0.993	0.989	0.991
4	0.963	0.999	0.996	0.995	0.997
5	0.944	1.000	0.999	1.000	0.998
6	0.963	0.999	0.996	0.994	0.999
7	0.954	0.999	0.992	0.994	0.991
8	0.954	0.998	0.985	0.986	0.983
9	0.958	0.997	0.982	0.976	0.976

Table 4.4 Coverage statistic for the different methods by stratum

The coverage rates per stratum for births and stratum movers are respectively given in the tables 4.5 and 4.6. For the subpopulations of births and stratum movers, the coverages are slightly lower than for the overall populations. Furthermore, there are hardly any differences between the methods and the strata. Stratum 8 appears to have a lower coverage for the stratum movers for all methods including SRS. We conclude that all methods sample births and stratum movers with the expected fractions and with an accuracy similar to that of SRS.

stratum	SRS	SBS	RRNSa	RRNSb	RRNSc
1	0.954	0.958	0.962	0.942	0.966
2	0.963	0.960	0.970	0.969	0.951
3	0.959	0.964	0.969	0.960	0.962
4	0.944	0.970	0.947	0.958	0.963
5	0.971	0.969	0.961	0.971	0.958
6	0.960	0.961	0.958	0.939	0.960
7	0.966	0.960	0.954	0.953	0.960
8	0.957	0.960	0.960	0.960	0.955
9	0.962	0.971	0.966	0.959	0.970

Table 4.5 Coverage statistic for births and for the different methods by stratum

stratum	SRS	SBS	RRNSa	RRNSb	RRNSc
1	0.949	0.961	0.950	0.951	0.947
2	0.959	0.956	0.967	0.964	0.960
3	0.968	0.967	0.956	0.955	0.956
4	0.972	0.973	0.960	0.960	0.960
5	0.978	0.985	0.976	0.976	0.981
6	0.973	0.963	0.956	0.964	0.961
7	0.965	0.959	0.957	0.971	0.966
8	0.931	0.936	0.927	0.934	0.933
9	0.952	0.953	0.954	0.955	0.951

Table 4.6 Coverage statistic for stratum movers and for the different methods by stratum

4.1.2 Measures of survey burden

Table 4.7 shows the spread of the total survey burden according to (7) for all sampling methods by stratum. This measure is only computed for the stable units in the simulation, that is, units that live in a single stratum during the full simulation period. It is clear from this measure that SBS and RRNS methods distribute the survey burden quite evenly over the units, as compared to SRS. Differences between SBS and the three RRNS methods are small.

stratum	SRS	SBS	RRNSa	RRNSb	RRNSc
1	2.33	0.16	0.21	0.18	0.19
2	3.21	0.14	0.18	0.20	0.18
3	3.60	0.15	0.18	0.18	0.17
4	3.85	0.17	0.18	0.16	0.19
5	3.72	0.16	0.14	0.17	0.18
6	3.71	0.15	0.20	0.21	0.21
7	3.57	0.17	0.20	0.17	0.18
8	3.12	0.15	0.18	0.17	0.17
9	2.34	0.15	0.17	0.19	0.21

Table 4.7 Spread of total survey burden for the different sampling methods by stratum, computed for the stable units in the simulation

Table 4.8 shows the realized survey leaves in each stratum and for all considered sampling methods. The table shows both the mean, given by Equation (9), and the spread according to (10) of the duration of the survey leaves. In this table, the spread of the survey leave duration is clearly better for SBS and the RRNS methods than for SRS. Although the survey leaves have been computed only for the stable units, these units do get several survey leaves in a simulation period, so there are enough observations to obtain a sufficiently accurate estimate. There are no clear differences in the average duration of the survey leave between the sampling methods SBS, RRNSa, RRNSb and RRNSc. Only in stratum 1, the mean duration of the survey leaves seems to be slightly shorter for SRS than for the other sampling methods.

	SRS		SBS		RRNSa		RRNSb		RRNSc	
stratum	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
1	8.67	5.08	9.01	0.11	9.04	0.14	9.00	0.13	8.98	0.14
2	3.98	1.55	4.00	0.04	3.99	0.04	4.00	0.04	4.00	0.04
3	2.32	0.74	2.33	0.03	2.33	0.03	2.33	0.03	2.33	0.03
4	1.51	0.43	1.50	0.02	1.49	0.02	1.50	0.02	1.50	0.02
5	1.02	0.27	1.00	0.01	1.00	0.01	1.00	0.01	1.00	0.01
6	0.67	0.18	0.66	0.01	0.66	0.01	0.66	0.01	0.66	0.01
7	0.43	0.13	0.42	0.01	0.42	0.01	0.42	0.01	0.42	0.01
8	0.25	0.08	0.25	0.01	0.25	0.01	0.25	0.01	0.25	0.01
9	0.11	0.05	0.11	0.01	0.11	0.01	0.11	0.01	0.11	0.01

Table 4.8 Survey leave: durations of survey free periods and corresponding standard deviation for the different sampling methods by stratum, computed for the stable units in the simulation

Table 4.9 displays the percentages of units with hotspot status, averaged over the simulation runs as well as over the simulated periods (see Equation (11)). For the strata 1 to 3 with small sampling fractions, SRS clearly leads to the largest percentage of hotspots in the population. For the strata 4, 5 and 6 the percentages are actually lower for SRS. This suggests that because the samples are drawn independently by SRS some units do not become hotspots at the expense of units that are already hotpots and that therefore face an even higher survey burden. For strata 7, 8 and 9, where the sampling fractions are highest, almost all units become hotspots and there are no longer any differences between the sampling methods. There are also no clear differences in the percentages of hotspots between the methods SBS, RRNSa, RRNSb and RRNSc.

In Figure 4.1, the average number of sampled hotspots according to (12) is shown as a function of time. Since the results for the three versions of RRNS are similar we only consider RRNSa here. The variation in the average number of sampled hotspots over time is clearly greatest for SRS. This means that the spread of survey burden is less optimal for SRS. As the sampling fractions get higher, this variation over time becomes smaller. Averaged over time, the number of sampled hotspots in the strata 1-4 is at a higher level for SRS than for SBS and RRNSa, but in the strata 5

stratum	SRS	SBS	RRNSa	RRNSb	RRNSc
1	0.41	0.00	0.00	0.00	0.00
2	4.98	0.04	0.01	0.01	0.01
3	17.73	1.31	1.09	1.07	1.06
4	36.80	48.01	48.11	47.73	47.64
5	53.59	65.61	65.93	65.90	65.92
6	63.98	68.20	68.26	68.22	68.20
7	68.80	69.80	69.77	69.80	69.80
8	70.60	70.76	70.78	70.76	70.77
9	71.62	71.71	71.74	71.71	71.72

Table 4.9 Percentage of units having hotspot status by stratum, averaged over simulation runs and time

and 6 SRS is at a lower level. However, as discussed earlier, SRS results in fewer hotspots in these strata, but these hotspot units face a higher survey burden than the hotspots caused by SBS and RRNSa. For strata 7, 8 and 9 there are no longer differences between the three sampling methods. In stratum 9, we see the average number of hotspots increasing over time. This is due to the simulated population dynamics as described in Subsection 3.1.1. This stratum is constantly growing due to an influx of stratum movers. Also for the number of sampled hotspots, there are no clear differences between SRS and RRNSa.

h!tbp

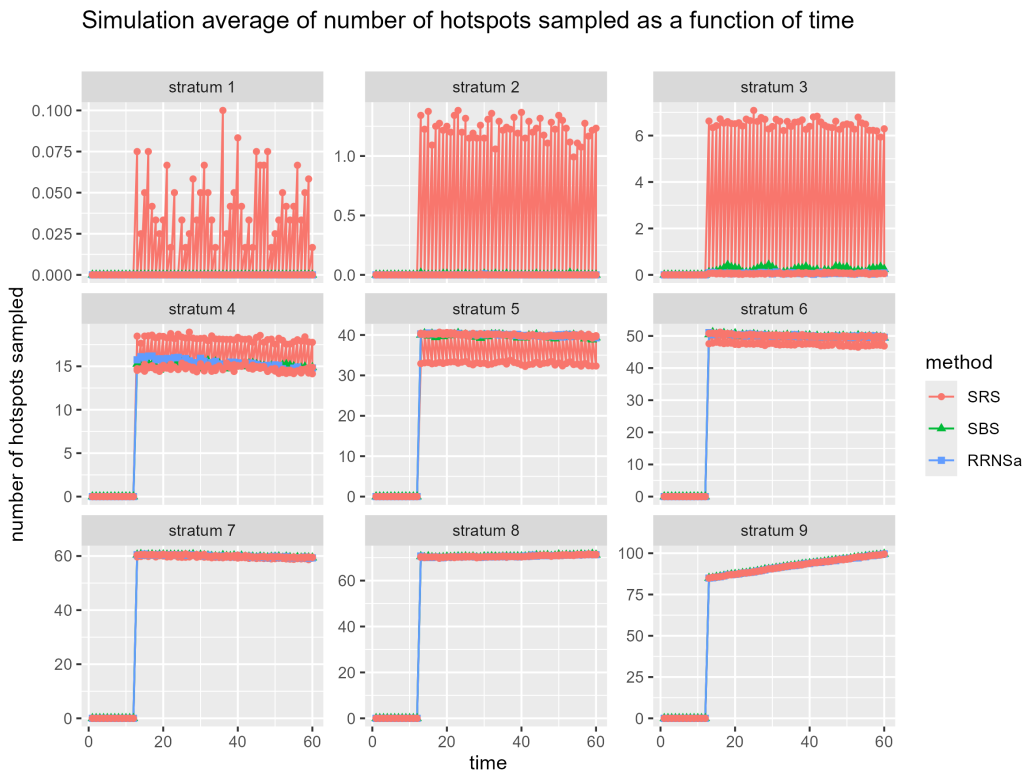

Figure 4.1 Average number of sampled hotspots over time for each stratum

4.2 Multiple (non-panel) surveys, five strata

In this section the simulation results for the setup described in Section 3.1.2 are discussed.

4.2.1 Realized sampling fractions and coverage

Table 4.10 shows the realized sampling fractions according to (4) for the different strata used in the survey designs of surveys 1 to 4. The third column contains the design fractions, as specified in Table 3.1, and the fourth column shows for reference the error margin of the realized fraction of sample inclusions under independent Bernoulli sampling with the specified design fraction, given by (5). We see that for all methods the realized overall sampling fractions are very close to the design fractions, as desired and expected.

survey	stratum	fraction (%)	error	SRS	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	5	0.016	5.000	5.000	5.001	5.000	4.999	5.000
1	2	12	0.042	11.999	11.998	12.001	12.002	12.000	12.000
1	3	25	0.074	25.004	24.997	24.997	24.996	24.994	25.004
1	4	60	0.122	59.996	60.003	59.995	59.997	60.015	60.000
1	5	100	0.000	100.000	100.000	100.000	100.000	100.000	100.000
2	1	4	0.012	4.000	4.000	4.000	4.000	4.000	4.000
2	2	18	0.054	18.003	17.999	18.003	17.996	18.002	17.996
2	3	80	0.132	79.987	80.006	80.017	79.994	79.998	79.985
3	1	7	0.016	7.000	7.001	7.000	7.000	7.000	6.999
3	2	20	0.056	20.000	20.001	20.008	20.004	20.000	19.999
3	3	75	0.142	75.011	74.998	75.013	74.992	74.999	74.989
4	1	6	0.018	5.999	6.000	6.000	6.000	6.001	6.000
4	2	9	0.038	9.001	8.999	8.998	8.997	9.003	9.000
4	3	25	0.074	24.996	24.996	25.001	25.000	24.998	25.001
4	4	75	0.108	75.008	75.002	75.007	75.010	75.004	74.994
4	5	100	0.000	100.000	100.000	100.000	100.000	100.000	100.000

Table 4.10 Realized overall sampling fractions (in %) for the different methods, by survey and stratum

Tables 4.11 and 4.12 show the realized sampling fractions for births and stratum movers, respectively. In each period, the computation is restricted to units that either were just born or changed stratum. The realized fractions again match the design fractions for all methods, strata and surveys, though with larger margins. The deviations seem to be of similar order of sizes between the methods, including SRS. This suggests that the larger deviations for all methods are due mostly to the relatively small number of births and movers in each period.

survey	stratum	fraction (%)	error	SRS	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	5	0.210	5.050	5.106	4.917	5.017	4.849	4.938
1	2	12	0.682	12.397	11.868	12.143	12.077	12.408	11.868
1	3	25	0.834	25.058	24.565	24.872	25.365	25.458	24.779
1	4	60	1.342	60.820	58.809	60.259	60.336	59.950	59.563
1	5	100	0.000	100.000	100.000	100.000	100.000	100.000	100.000
2	1	4	0.172	3.913	3.986	3.772	4.033	4.096	4.009
2	2	18	0.608	18.079	17.721	18.399	17.772	18.550	18.544
2	3	80	1.316	79.144	80.309	81.121	80.796	79.794	81.934
3	1	7	0.224	7.002	6.886	7.108	7.083	7.027	7.019
3	2	20	0.634	20.502	19.906	19.234	19.240	19.548	20.314
3	3	75	1.426	73.998	75.244	73.456	75.135	76.273	74.485
4	1	6	0.230	6.044	5.995	5.911	5.993	5.997	6.011
4	2	9	0.600	9.080	9.444	9.069	9.003	9.234	8.694
4	3	25	0.834	24.956	24.900	25.960	25.988	24.538	24.612
4	4	75	1.204	75.904	75.247	76.136	75.343	75.653	73.990
4	5	100	0.000	100.000	100.000	100.000	100.000	100.000	100.000

Table 4.11 Realized sampling fractions (in %) for births for the different methods, by survey and stratum

Table 4.13 shows overall coverage rates as defined in (3). Whereas the coverage rates for SRS are,

survey	stratum	fraction (%)	error	SRS	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	5	0.190	5.023	4.836	4.548	5.035	5.150	5.014
1	2	12	0.228	12.019	12.280	12.069	12.127	11.919	11.877
1	3	25	0.502	24.983	25.442	24.930	25.114	24.816	24.584
1	4	60	0.752	60.540	61.751	61.434	60.475	60.969	60.357
1	5	100	0.000	100.000	100.000	100.000	100.000	100.000	100.000
2	1	4	0.372	3.894	3.579	3.831	3.741	4.451	3.903
2	2	18	0.472	17.736	18.139	18.573	18.162	18.030	17.818
2	3	80	1.002	80.261	82.697	82.870	79.978	79.617	81.518
3	1	7	0.484	6.754	7.230	7.572	7.410	6.888	6.942
3	2	20	0.492	20.118	19.636	19.937	19.805	20.556	20.345
3	3	75	1.086	74.697	78.564	74.305	74.933	75.373	74.100
4	1	6	0.206	6.030	5.633	5.933	5.935	5.992	6.028
4	2	9	0.200	9.150	8.936	9.049	9.004	9.064	8.961
4	3	25	0.502	25.144	24.598	25.650	24.946	24.581	25.365
4	4	75	0.664	75.193	76.745	75.492	75.498	75.392	74.881
4	5	100	0.000	100.000	100.000	100.000	100.000	100.000	100.000

Table 4.12 Realized sampling fractions (in %) for stratum movers for the different methods, by survey and stratum

as expected, close to 0.95, the coverage rates of all other methods are always larger than 0.95 and mostly close to 1. This is due to the negative coordination of these methods, forcing smaller variation in the number of times units are sampled over the simulation time span. Method SBS has coverage rates closest to 1, which is not surprising as this method spreads survey burden optimally over time for each survey separately. The other coordinating methods also spread survey burden over time, but across the 4 surveys. Of the RRNS methods RRNSb on average has slightly lower coverage rates, presumably because it injects some noise through regeneration of all random numbers after each sample draw.

survey	stratum	SRS	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	0.954	0.999	0.996	0.995	0.989	0.996
1	2	0.962	1.000	0.994	0.991	0.967	0.996
1	3	0.953	1.000	0.997	0.998	0.998	0.998
1	4	0.956	1.000	0.999	1.000	1.000	0.999
1	5	1.000	1.000	1.000	1.000	1.000	1.000
2	1	0.949	0.998	0.994	0.991	0.979	0.996
2	2	0.956	0.998	0.998	0.996	1.000	0.999
2	3	0.954	0.998	0.997	0.997	0.993	0.995
3	1	0.952	0.999	0.997	0.996	0.993	0.999
3	2	0.950	0.999	0.999	0.999	1.000	0.999
3	3	0.959	1.000	0.999	0.998	0.999	0.998
4	1	0.957	0.999	0.998	1.000	0.994	0.997
4	2	0.959	0.999	0.986	0.991	0.972	0.995
4	3	0.961	0.999	0.999	1.000	0.999	0.997
4	4	0.954	1.000	1.000	0.999	0.999	1.000
4	5	1.000	1.000	1.000	1.000	1.000	1.000

Table 4.13 Coverage statistic for the different methods by survey and stratum

Tables 4.14 and 4.15 show the coverage rates for births and movers, respectively. These coverage rates are based on much fewer units and therefore fluctuate more, but still do so around the 0.95 value, or slightly above for the coordinating methods. So all methods sample births and movers with the expected fractions and with accuracy similar to that of SRS. Stratum 3 of survey 2 seems to have somewhat lower coverage for stratum movers. This may be due to the fact that this is the smallest stratum with sampling fraction less than 1, so the reported coverage rates are more uncertain.

survey	stratum	SRS	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	0.961	0.944	0.966	0.966	0.965	0.968
1	2	0.961	0.962	0.972	0.954	0.956	0.969
1	3	0.959	0.955	0.956	0.940	0.966	0.964
1	4	0.964	0.977	0.959	0.964	0.956	0.956
1	5	1.000	1.000	1.000	1.000	1.000	1.000
2	1	0.978	0.981	0.976	0.966	0.950	0.966
2	2	0.960	0.952	0.957	0.955	0.960	0.961
2	3	0.948	0.956	0.964	0.959	0.935	0.967
3	1	0.950	0.947	0.949	0.963	0.949	0.954
3	2	0.961	0.968	0.965	0.968	0.958	0.961
3	3	0.959	0.966	0.962	0.974	0.971	0.956
4	1	0.937	0.955	0.963	0.960	0.952	0.953
4	2	0.972	0.961	0.958	0.958	0.943	0.962
4	3	0.979	0.956	0.948	0.947	0.968	0.948
4	4	0.972	0.971	0.974	0.971	0.966	0.969
4	5	1.000	1.000	1.000	1.000	1.000	1.000

Table 4.14 Coverage statistic for births, for the different methods by survey and stratum

survey	stratum	SRS	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	0.959	0.961	0.972	0.958	0.959	0.958
1	2	0.959	0.959	0.959	0.952	0.953	0.953
1	3	0.965	0.955	0.971	0.959	0.969	0.960
1	4	0.958	0.975	0.964	0.973	0.964	0.963
1	5	1.000	1.000	1.000	1.000	1.000	1.000
2	1	0.957	0.962	0.958	0.954	0.954	0.962
2	2	0.960	0.963	0.943	0.960	0.963	0.959
2	3	0.903	0.923	0.932	0.910	0.885	0.911
3	1	0.959	0.954	0.952	0.960	0.955	0.955
3	2	0.967	0.962	0.951	0.963	0.961	0.963
3	3	0.969	0.971	0.964	0.982	0.973	0.961
4	1	0.948	0.971	0.959	0.966	0.955	0.950
4	2	0.951	0.960	0.956	0.954	0.959	0.957
4	3	0.962	0.959	0.959	0.960	0.964	0.959
4	4	0.971	0.962	0.967	0.970	0.966	0.964
4	5	1.000	1.000	1.000	1.000	1.000	1.000

Table 4.15 Coverage statistic for stratum movers, for the different methods by survey and stratum

4.2.2 Measures of survey burden

Table 4.16 shows the simulation results for a measure of survey burden inequality. The measure is computed from the standard deviations of the total number of times each unit is sampled within the simulation time span, computed for each stratum[1)]. These standard deviations are then averaged over simulation runs, see (7). For convenience these measures have been computed based only on units not affected by population dynamics, i.e. stable units. From the table it is clear that all coordinating methods drastically reduce this inequality as compared to SRS. The best spread of survey burden is achieved by SBS1 and the RRNS methods, due to the fact that these methods spread the burden not only over time but across the different surveys as well. By contrast, SBS spreads survey burden over time only, independently for each survey.

1) Here and in several other tables, simulation results are presented by the most detailed of the two stratification variables used, i.e. Strat1 of Table 3.1

stratum	SRS	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	4.58	0.26	0.15	0.15	0.24	0.17
2	5.52	0.40	0.26	0.23	0.44	0.25
3	8.05	0.50	0.31	0.27	0.48	0.31
4	8.74	0.60	0.37	0.36	0.62	0.41
5	6.04	0.54	0.42	0.46	0.40	0.45

Table 4.16 Standard deviation of the total number of times sampled over all surveys by detailed stratum, for the stable units in the simulation

To further illustrate how much the negative coordination over surveys of the SBS1 and RRNS methods reduces the survey burden inequality, Table 4.17 shows the standard deviations of the total number of times sampled for the first survey only. Compared to Table 4.16 the inequality measures for the SBS1 and RRNS methods are much larger for a single survey, so there must be strong negative correlations between number of times sampled for each of the four surveys.

stratum	SRS	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	2.15	0.15	1.10	1.13	1.67	1.11
2	3.37	0.25	1.50	1.50	2.84	1.53
3	4.20	0.27	1.18	1.58	1.88	1.21
4	5.02	0.31	1.67	1.93	2.33	1.74
5	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.17 Standard deviation of the total number of times sampled for survey 1 only by detailed stratum, for the stable units in the simulation

A measure for peaks in survey burden is the number of times that units are drawn simultaneously in two or more surveys, 'stacking' of surveys in short. These numbers are computed by stratum for the different sampling methods, and are displayed in Figure 4.2. Stacking is most pronounced for SRS and SBS, as these methods do not spread survey burden across the different surveys. The other methods restrict the amount of stacking as much as possible. In particular, these methods prevent stacking from happening altogether in the first three strata with the lowest sampling fractions, and in stratum 4 no units are drawn in more than 2 surveys simultaneously. Note that stacking cannot be completely prevented in strata where the sum of the sampling fractions of the participating surveys exceeds 1.

Figure 4.3 shows the distributions of survey leave durations, i.e. the time spans between two selections for any of the surveys, for stable units. Note that the y-axis in Figure 4.3 uses a square root scale to more clearly show that SRS yields occasionally very long survey leave durations in stratum 1. For SBS the maximum observed survey leave period is 14 (in stratum 1) and for the other methods the survey leave of stable units concentrates solely on the visible peaks shown in the figure. The figure shows that by spreading survey burden over time, method SBS already achieves a fairer distribution of survey leave durations than simple random sampling. Moreover, the other 4 methods that also spread the burden across surveys are clearly superior in this regard. For example, all units living in stratum 1 have survey leaves of 3 or 4 periods, and neither smaller nor larger values occur. These methods provide the most narrow, and most fair, distribution of survey leave durations, the particular survey leave durations being determined by the sum over surveys of sampling fractions and hardly by chance.

As before, a hotspot is defined as a unit that in the past 12 periods has been selected at least 5 times in total for all surveys. Table 4.18 lists the percentages of units that become hotspot averaged over the simulation runs as well as over the simulated periods. For stratum 1 with small sampling fractions over 11% of units are hotspot on average at any moment if samples are drawn using simple random sampling. SBS with each of the 4 surveys in a separate group reduces this figure to less than half a percent. A further improvement is achieved by SBS1 where all surveys

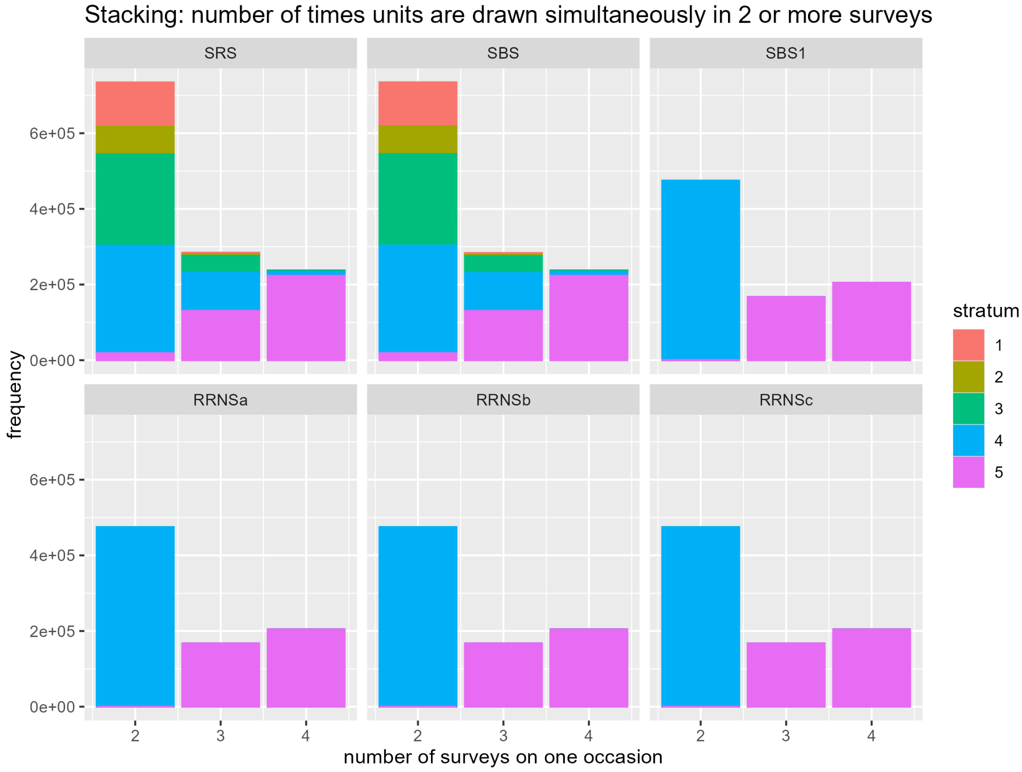

Figure 4.2 Stacking: simultaneous inclusion of units in multiple surveys, by stratum, for each of the 6 sampling methods considered

belong to the same group. All RRNS methods perform on par with SBS1. Large reductions of numbers of hotspots are realized by SBS and especially SBS1 and RRNS methods in stratum 2 as well. For stratum 3 and higher the percentage is slightly lower for SRS than for the other methods. This indicates that due to the uncoordinated SRS sampling some units do not become hotspot at the expense of units that are already hotspot and are burdened even more.

stratum	SRS	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	11.10	0.48	0.15	0.15	0.15	0.15
2	28.68	16.91	3.99	3.98	4.00	3.98
3	85.92	88.06	88.47	88.44	88.43	88.46
4	94.82	95.13	95.25	95.26	95.26	95.26
5	96.69	96.70	96.71	96.71	96.71	96.71

Table 4.18 Percentage of units having hotspot status by (detailed) stratum, averaged over simulation runs and time

Finally, Figure 4.4 shows the number of hotspots sampled over time for each survey and detailed strata 1, 2 and 3. Graphs for strata 4 and 5 have been left out as they show similar patterns as stratum 3. Also, graphs for methods RRNSb and RRNSc have been left out as they are similar to the graphs for RRNSa. In stratum 1 hardly any hotspot is ever sampled by the coordinating methods, where under SRS on average about 5 hotspots per period (after a short simulation start-up period in which the number of hotspots has not yet reached its stable value) get sampled by each survey. In stratum 2 the situation is similar, except that due to larger sampling fractions hotspots are sometimes drawn by all methods, though much more often by SRS, and more often by SBS compared to the other coordinating methods. In stratum 3 the larger sampling fractions are such that practically all units become hotspots after the start-up phase. The increasing trends are due to the simulated population dynamics, which cause a small nett incoming flow of stratum movers.

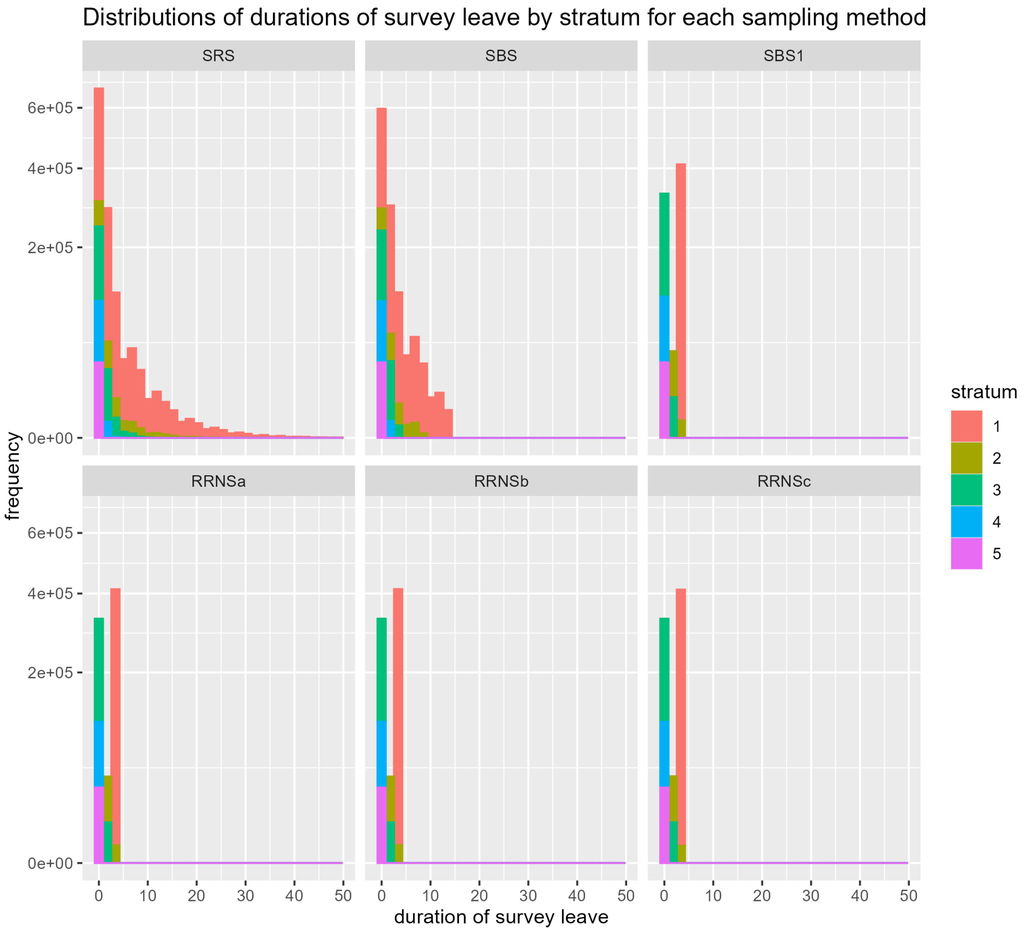

Figure 4.3 Survey leave: distribution of durations of survey free periods, by stratum, for stable units, for each of the 6 sampling methods considered

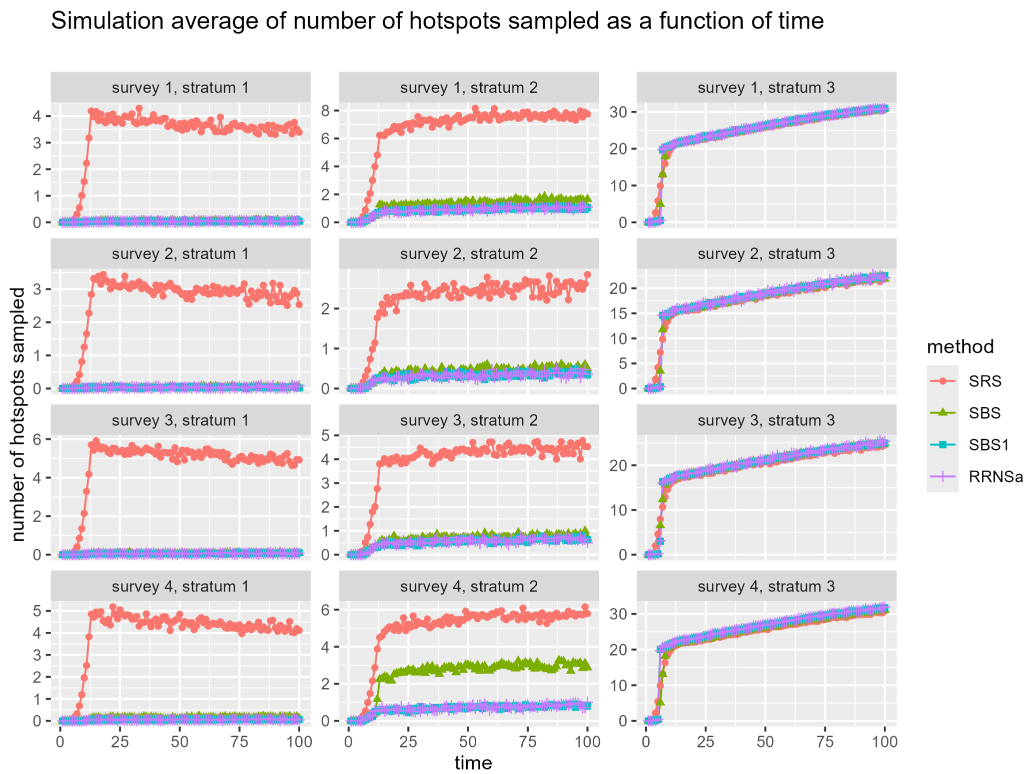

Figure 4.4 Average number of hotspots sampled over time for each survey and detailed strata 1-3

4.3 Multiple surveys including a rotating panel survey, five strata

In this section the simulation results for the setup described in Section 3.1.3 are discussed. Here three surveys are considered of which the first is a rotating panel survey with rotation fraction 10%.

4.3.1 Realized sampling fractions and coverage

The realized sampling fractions under this setup are again very close to the design fractions as shown in Table 4.19 for all units by stratum and for each survey. For births and stratum movers (Tables 4.20 and 4.21) the realized fractions also fluctuate around the design fractions although the deviations are larger due to the much smaller number of units that are born or have moved from another stratum at any one period.

survey	stratum	fraction (%)	error	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	5	0.016	5.000	5.000	4.999	4.999	5.001
1	2	12	0.042	11.998	11.999	12.000	11.999	11.999
1	3	25	0.074	25.003	25.001	24.999	25.003	24.998
1	4	60	0.122	60.007	59.993	60.000	59.998	59.995
1	5	100	0.000	100.000	100.000	100.000	100.000	100.000
2	1	4	0.012	4.000	3.999	4.000	3.999	4.000
2	2	18	0.054	17.999	18.002	17.998	17.999	18.002
2	3	80	0.132	80.003	79.994	80.024	80.002	79.994
3	1	6	0.018	6.000	6.000	6.000	6.000	6.001
3	2	9	0.038	9.002	8.999	9.000	8.997	9.002
3	3	25	0.074	25.003	25.000	24.999	25.002	25.005
3	4	75	0.108	75.001	75.002	74.997	74.998	74.998
3	5	100	0.000	100.000	100.000	100.000	100.000	100.000

Table 4.19 Realized sampling fractions overall for the different methods by survey and stratum

survey	stratum	fraction (%)	error	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	5	0.210	5.160	5.099	4.882	5.181	5.078
1	2	12	0.678	13.109	13.011	12.576	12.685	11.782
1	3	25	0.832	25.718	26.253	25.681	25.930	24.896
1	4	60	1.362	61.457	61.689	61.650	61.901	61.341
1	5	100	0.000	100.000	100.000	100.000	100.000	100.000
2	1	4	0.172	4.012	3.998	3.927	3.908	3.998
2	2	18	0.608	17.453	17.459	18.390	17.340	18.121
2	3	80	1.332	79.528	79.196	79.528	80.777	80.888
3	1	6	0.230	5.994	5.889	6.047	5.910	6.071
3	2	9	0.596	8.899	8.812	8.834	8.649	8.605
3	3	25	0.832	25.293	25.136	24.241	24.887	25.948
3	4	75	1.204	73.764	74.247	75.174	76.681	74.807
3	5	100	0.000	100.000	100.000	100.000	100.000	100.000

Table 4.20 Realized sampling fractions for births for the different methods by survey and stratum

The coverage rates in Table 4.22 are mostly fine for the non-panel surveys 2 and 3, but generally well below 0.95 for panel survey 1. This is probably due to the strong dependence between sample draws of the panel survey. The small rotation fraction of 10% induces strong positive dependence over subsequent draws. Note that the coverage rates under method SBS are higher than for the other methods. Only SBS treats the surveys independently, so the coverage rates for surveys 2 and 3 are as high as in Table 4.13, i.e. very close to 1. Surprisingly, for births and stratum movers the coverage rates are reasonably close to 0.95 for all surveys including the panel

survey	stratum	fraction (%)	error	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	5	0.190	4.850	4.916	5.138	5.472	5.074
1	2	12	0.228	12.315	12.160	12.371	12.492	12.210
1	3	25	0.502	26.030	25.893	25.709	25.769	25.608
1	4	60	0.752	62.463	62.876	61.431	61.738	61.679
1	5	100	0.000	100.000	100.000	100.000	100.000	100.000
2	1	4	0.372	3.433	4.033	3.728	3.934	4.149
2	2	18	0.472	17.787	18.036	17.840	17.606	18.311
2	3	80	1.002	83.370	83.636	80.956	80.188	80.658
3	1	6	0.206	5.740	6.224	6.060	5.938	6.106
3	2	9	0.200	8.931	9.233	8.906	8.874	8.883
3	3	25	0.502	24.827	25.384	24.834	24.794	25.434
3	4	75	0.664	76.990	73.411	74.708	75.274	74.643
3	5	100	0.000	100.000	100.000	100.000	100.000	100.000

Table 4.21 Realized sampling fractions for stratum movers for the different methods by survey and stratum

survey, see Tables 4.23 and 4.24. The positive coordination by the panel survey is less strong for these units, as on average their panel membership is shorter due to the way dynamics is handled in the sampling algorithms (Section 2.2). So apparently the positive panel coordination and negative coordination by SBS, SBS1 and the RRNS methods roughly cancel each other out so that coverage rates are near their expected values under independence. Further analysis or simulation would be necessary to check this hypothesis.

survey	stratum	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	0.561	0.556	0.499	0.499	0.474
1	2	0.638	0.596	0.566	0.545	0.578
1	3	0.783	0.620	0.584	0.599	0.590
1	4	0.964	0.791	0.798	0.802	0.858
1	5	1.000	1.000	1.000	1.000	1.000
2	1	0.999	0.916	0.913	0.921	0.875
2	2	0.999	0.915	0.967	0.969	0.981
2	3	0.998	0.995	0.995	0.996	0.996
3	1	0.998	0.929	0.918	0.950	0.890
3	2	0.999	0.853	0.983	0.987	0.986
3	3	0.999	0.885	0.920	0.928	0.933
3	4	1.000	0.970	0.989	0.994	0.998
3	5	1.000	1.000	1.000	1.000	1.000

Table 4.22 Coverage statistic for the different methods by survey and stratum

survey	stratum	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	0.945	0.954	0.962	0.959	0.953
1	2	0.950	0.950	0.951	0.961	0.953
1	3	0.944	0.945	0.965	0.971	0.982
1	4	0.961	0.976	0.968	0.959	0.971
1	5	1.000	1.000	1.000	1.000	1.000
2	1	0.967	0.972	0.972	0.969	0.973
2	2	0.962	0.958	0.958	0.958	0.966
2	3	0.945	0.951	0.936	0.964	0.945
3	1	0.961	0.962	0.932	0.962	0.951
3	2	0.966	0.956	0.959	0.966	0.972
3	3	0.957	0.971	0.947	0.966	0.963
3	4	0.953	0.981	0.958	0.985	0.968
3	5	1.000	1.000	1.000	1.000	1.000

Table 4.23 Coverage statistic for births, for the different methods by survey and stratum

survey	stratum	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	1	0.949	0.941	0.941	0.947	0.949
1	2	0.957	0.946	0.949	0.944	0.954
1	3	0.968	0.945	0.958	0.956	0.963
1	4	0.970	0.973	0.969	0.969	0.967
1	5	1.000	1.000	1.000	1.000	1.000
2	1	0.958	0.949	0.956	0.956	0.951
2	2	0.946	0.958	0.948	0.948	0.950
2	3	0.911	0.916	0.926	0.932	0.920
3	1	0.959	0.965	0.952	0.952	0.962
3	2	0.960	0.970	0.962	0.957	0.954
3	3	0.959	0.965	0.968	0.966	0.966
3	4	0.969	0.971	0.971	0.967	0.973
3	5	1.000	1.000	1.000	1.000	1.000

Table 4.24 Coverage statistic for stratum movers, for the different methods by survey and stratum

4.3.2 Measures of survey burden

Table 4.25 shows the standard deviation of the total number of times stable units are drawn in a sample from any of the 3 surveys. Clearly the SBS methods achieve lower values and therefore a better spread of survey burden compared to the RRNS methods, according to this measure. As expected, method SBS1, which coordinates all surveys in a single group, is superior to SBS, which coordinates only over time for each survey separately.

stratum	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	4.86	2.30	5.89	5.66	6.04
2	2.99	2.31	6.51	6.67	6.73
3	3.22	0.94	3.17	3.15	3.10
4	1.32	0.40	5.16	4.68	4.04
5	0.38	0.34	0.38	0.42	0.43

Table 4.25 Standard deviation of the total number of times sampled over all surveys by detailed stratum, for the stable units in the simulation

The RRNS methods seem to work less well compared to SBS for distributing the survey burden over surveys where at least one of the surveys is a panel survey. The reason is that unlike SBS the RRNS methods do not store a built-up survey burden. They are therefore not able to account for the duration of panel membership. For SBS the survey burden value is increased for each period a unit remains in a panel, and so after a lengthy stay in a panel such a unit will probably not be selected for another (or the same) survey soon. RRNS methods do not have a long-term memory in this sense so that units leaving a panel may be selected for the next survey much sooner.

The standard deviations in Table 4.26 concerning only the panel survey (survey 1) are much larger than the combined ones in Table 4.25 except for method SBS, consistent with the fact that SBS1 and the RRNS methods apply negative coordination over all surveys, whereas SBS only applies negative coordination over time for each survey.

stratum	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	4.85	5.23	7.45	7.13	7.63
2	2.98	6.47	8.77	8.93	9.06
3	3.23	9.64	10.51	10.46	10.30
4	1.27	8.02	7.87	7.08	6.16
5	0.00	0.00	0.00	0.00	0.00

Table 4.26 Standard deviation of the total number of times sampled for survey 1 (the panel) only by detailed stratum, for the stable units in the simulation

Figure 4.5 shows the stacking measure, the number of times units are drawn simultaneously in 2 or more surveys. Stacking of two or more surveys happens more often under method SBS, which does not coordinate across surveys. The other 4 methods restrict the amount of stacking as much as possible. In particular, as in the case of non-panel surveys only, these methods prevent stacking altogether in the first three strata with the lowest sampling fractions, and in stratum 4 no units are drawn in more than 2 surveys simultaneously. So regarding the stacking measure the RRNS methods are as effective as SBS1 also if the surveys coordinated over include panels.

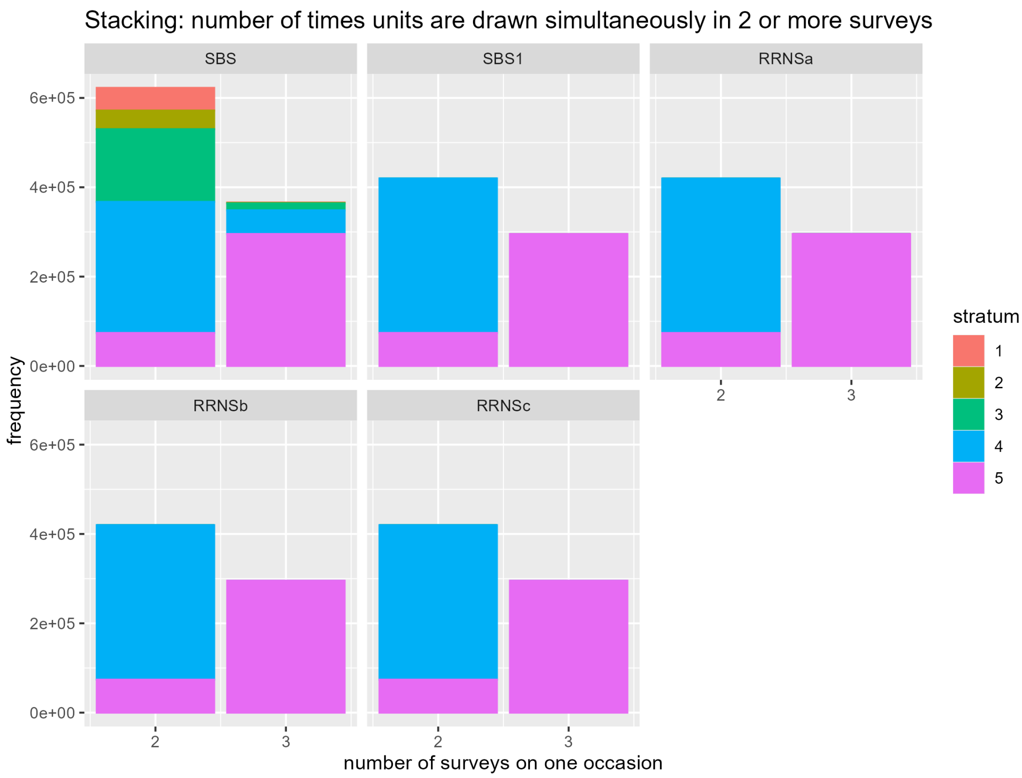

Figure 4.5 Stacking: simultaneous inclusion of units in multiple surveys, by stratum, for each of the 5 sampling methods considered

Figure 4.6 shows the distributions of survey leave durations, i.e. the time spans between two selections for any of the surveys, for stable units. Note that the y-axis in Figure 4.6 uses a square root scale to more clearly show the longer survey leave periods of method SBS1 mostly in stratum 1 but also in stratum 2. These correspond to units that after having been in the panel (on average approximately for 10 periods due to the 10% rotation fraction) enjoy a long survey free period due to SBS1's use of built-up survey burden and negative coordination across different surveys. The RRNS methods yield the narrowest distributions of survey leave duration. In particular, for stratum 1 there is not only a peak at 0 (units remaining in the panel) but also a sharp peak around 8, the maximum survey leave being 9 periods. So from the perspective of equality of survey leave durations the RRNS methods perform well.

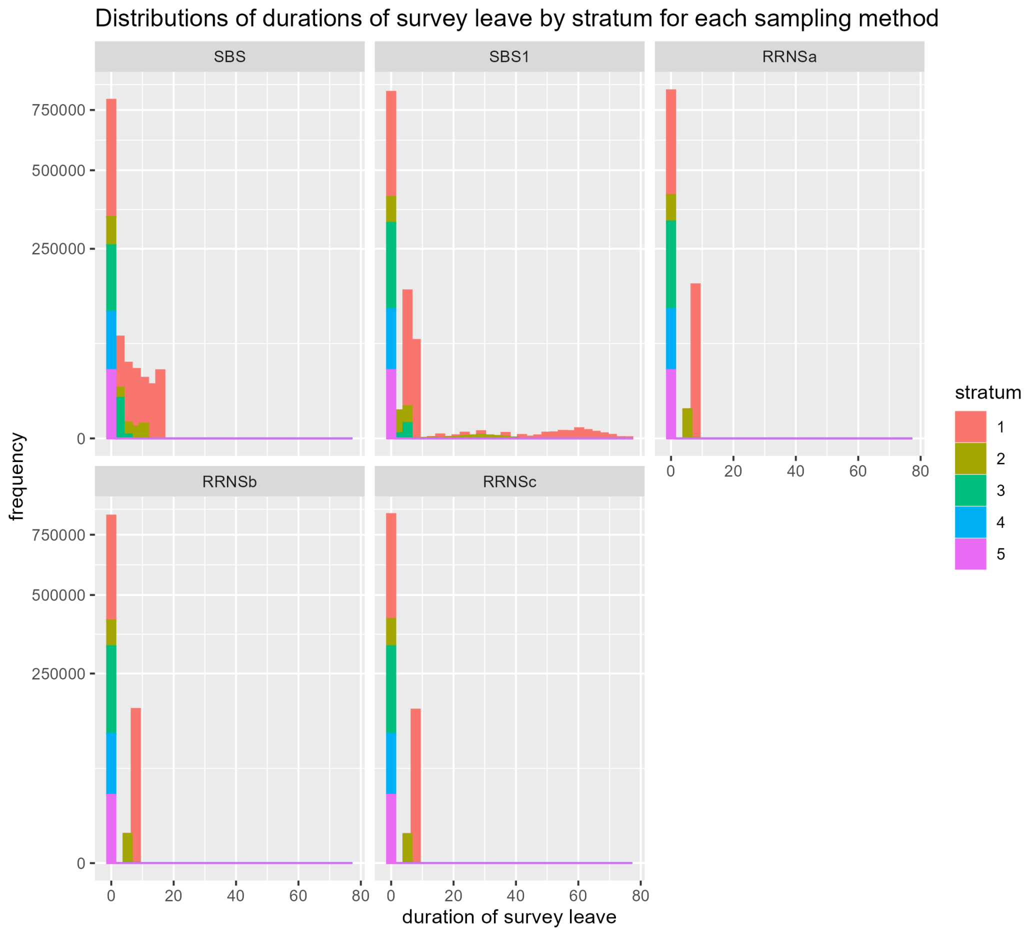

Figure 4.6 Survey leave: distribution of durations of survey free periods, by stratum, for stable units, for each of the 6 sampling methods considered

Figure 4.7 shows that the distribution of durations of panel membership is very similar between the different methods. The highest peak is around 10 as can be expected for a rotation fraction of 10%. The figure also shows a wide tail to the left, which arises due to population dynamics: for example, births flow into the panel according to their assigned random numbers and are therefore comparable to any existing panel member regarding the length of (remaining) stay in the panel. We have also checked that all methods yield roughly the same size of panel overlaps between subsequent periods, as they should because of the use of the same rotation fraction as well as the same simulated population dynamics.

As before, a hotspot is defined as a unit that in the past 12 periods has been selected at least 5 times in total for all surveys. For the panel survey (survey 1), we only count the inflow into the panel as selection event contributing to the hotspot status. Table 4.27 shows the percentage of units having hotspot status under the different methods. The differences between methods are not very large, although SBS has a somewhat larger fraction of hotspots in stratum 3 and SBS1 has a slightly larger fraction in stratum 2. This is in line with Figure 4.8 showing the average number

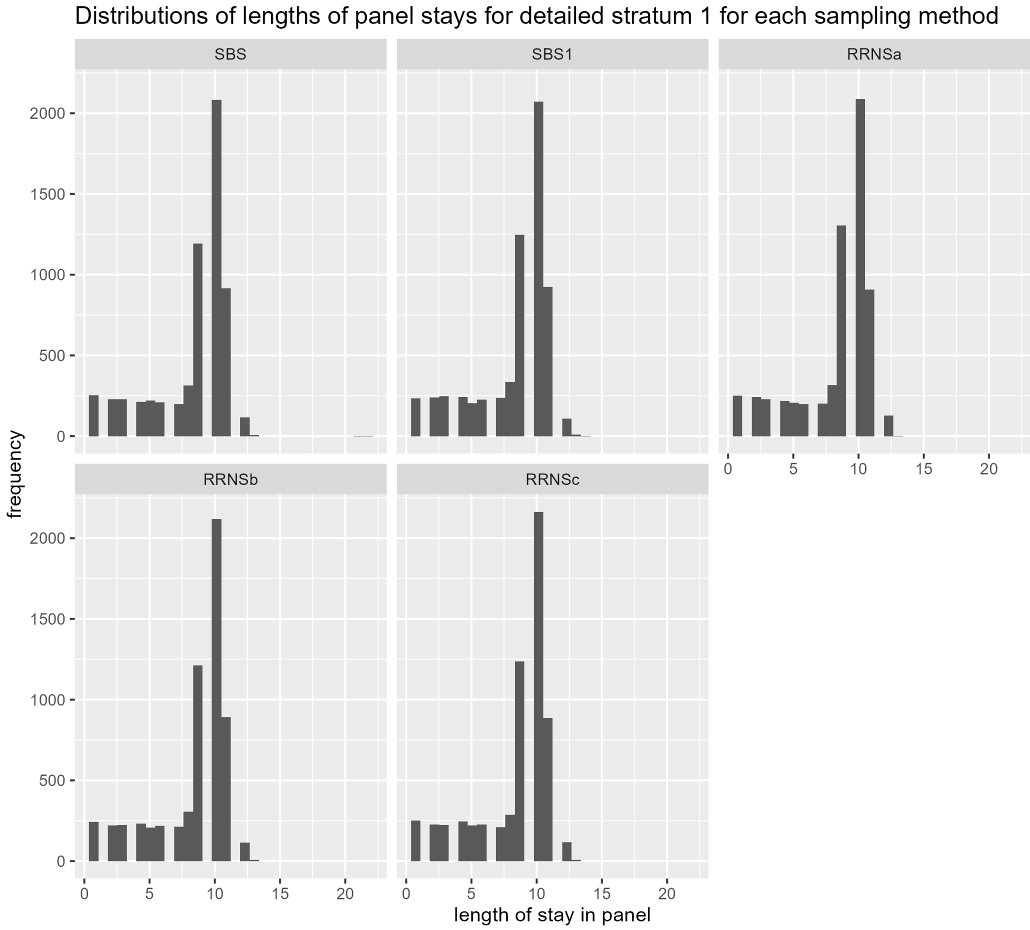

Figure 4.7 Duration of panel membership: distribution of lengths of panel stays, for the stable units in stratum 1, for each of the 5 sampling methods considered

of hotspots *sampled* as a function of time. The differences of numbers of hotspots sampled in strata 4 and 5 are quite small between all methods and therefore not shown in Figure 4.8. RRNSb and RRNSc results are very similar to those of RRNSa and therefore omitted as well.

stratum	SBS	SBS1	RRNSa	RRNSb	RRNSc
1	0.02	0.05	0.04	0.04	0.04
2	1.32	1.75	1.55	1.55	1.57
3	73.38	52.25	53.94	54.24	54.06
4	92.54	90.09	91.15	91.07	91.08
5	96.03	95.97	96.00	96.01	95.99

Table 4.27 Percentage of units having hotspot status by (detailed) stratum, averaged over simulation runs and time

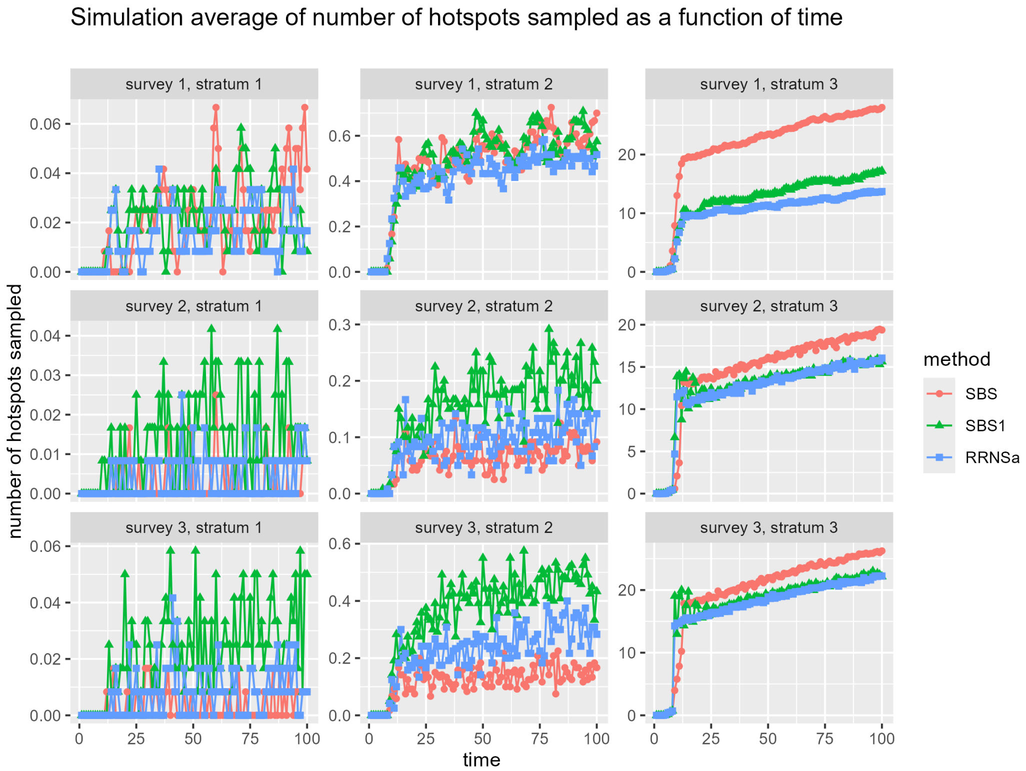

Figure 4.8 Average number of hotspots sampled over time for each survey and detailed strata 1-3

5 Discussion

In this report we investigate a new simple method to evenly distribute the burden caused by a set of sample surveys over sampling units. The method, RRNS for rotating random number sampling, uses random numbers (RNs) that rotate over the units after each survey sample draw. The method was introduced in Boonstra et al. (2024) where it was found to be a potentially valuable alternative to SBS (Survey Burden Sampling), the method currently used to draw many business surveys at Statistics Netherlands (Smeets and Boonstra 2018). A limitation of SBS is that in order to spread the total survey burden caused by multiple surveys over the sampling units, the stratifications of these surveys must be aligned. In practice this turns out to be difficult to achieve as surveys can be very different regarding their inferential targets and survey managers are often hesitant to changes in stratification. If one would have to account for each survey's specific targets simultaneously one might end up with a very detailed stratification. Such detailed stratifications would quickly become impractical due to small population sizes, small sample sizes, rounding issues, large numbers of stratum movers, and other reasons.

The RRNS method does not impose any restrictions on stratifications, so it can spread the burden caused by multiple surveys evenly over the sampling units without any harmonization of stratifications. This is done by using the same RN variable for all surveys and by appropriately rotating the random numbers over units after each sample draw. Unlike SBS, the RRNS method does not keep a record of built-up survey burden, so in that sense it is also a simpler method. However, this also means that it cannot differentiate between the amount of burden caused by each survey. So in case of a group of surveys with aligned stratifications and substantial differences between the surveys' burdens, SBS is the more appropriate coordination method.

In Boonstra et al. (2024) it was found by way of a simulation study that the RRNS method works properly in the sense that it respects the inclusion probabilities of the sampling design, at least for a population that does not change over time. It was also found to be able to reduce the number of hotspots drawn in a sample, a specific measure of (peak) survey burden currently used at Statistics Netherlands. In the current report we examine the RRNS method in more detail. In particular, we extend the simulation studies to incorporate population dynamics, i.e. births, deaths and stratum movers, generalize the method to deal with (rotating) panel surveys and assess the realized samples based on a wider set of survey burden measures.

The simulation results show that RRNS works well under population dynamics. All units, including births and stratum movers are drawn with the correct design fractions. In the simulations with one or several cross-sectional surveys, coverage rates of approximate 95% intervals for the estimated sampling fractions are close to and mostly above 95% for births and stratum movers. Overall coverage rates for all units are often very close to 1 due to the negative coordination over time. In the simulation study that includes a rotating panel survey, all sampling procedures including the RRNS methods sample from births, stratum movers and other units according to the correct design fractions, even though the naive (independence-based) intervals have coverage rates well below 95%. This happens for all sampling methods and is due to the positive coordination of the (slowly) rotating panel survey.

Three versions of the RRNS method have been tested. Under method RRNSa the random numbers never change, only rotate over the units. Methods RRNSb and RRNSc also regenerate the random numbers but assign them in the same order as RRNSa to the units. Method RRNSb regenerates the random numbers from the uniform distribution whereas RRNSc only uses a single draw from the uniform distribution to assign equidistant ’random numbers’ to all units after each draw. The differences between the three RRNS methods are small for most simulation measures considered. However, method RRNSb yields a somewhat larger standard deviation over sampling units of the total number of times sampled under a set of non-panel surveys. This is probably due to the fact that RRNSb injects more noise to the sampling procedure by its way of regenerating the random numbers. This is a (small) disadvantage of RRNSb. Also, in the simulation of a single survey the realized sampling fractions for method RRNSa appeared to deviate slightly from the design fractions in some strata. Therefore, RRNSc with equidistant regeneration of random numbers is currently the preferred version of RRNS methods.

In the simulation studies we have compared the new RRNS methods to the SBS method. In the simulation study of a single survey the results for both methods are very similar. In the simulation study of four non-panel surveys two versions of the SBS method have been considered: the version in which coordination is over time only (in the results section denoted by SBS) and the version in which coordination is over time and over the four surveys simultaneously (denoted by SBS1, as all surveys belong to a single SBS group). The results show that the coordination achieved by RRNS is very similar to that of SBS1, both of which spread the survey burden much better than the SBS method that coordinates each survey separately over time. In the simulation study the four surveys use two different stratifications. But as these stratifications are nested, the most detailed of these two stratifications could be used as the single base stratification for the SBS1 method. In practice, however, stratifications of business surveys can be very different and there may be many obstacles to harmonizing them, which makes it hard to apply SBS1. The RRNS method, however, can be applied without any restrictions on the different stratifications, which makes spreading the survey burden over time and across surveys much easier in practice.

In another simulation setup, we compared how the methods spread the sample burden over three repeated surveys, one of them being a rotating panel with a rotation fraction of 10%. For this situation, which mixes positive coordination by the panel over time and negative coordination by the survey burden methods over both time and surveys, the results show a disadvantage of RRNS compared to SBS. Whereas SBS accounts for the duration that units remain in the panel by keeping track of the built-up survey burden, RRNS sees only the current random number variable and so cannot account for the duration of panel membership. As a result the standard deviation of the number of times units are sampled over the simulation time span is much higher for RRNS than it is for SBS and SBS1. However, not all measures are worse for RRNS: regarding the stacking and hotspot measures RRNS is similar to SBS1, and the distribution of survey leave is actually more favourable under RRNS. So in the case of panel surveys it is a trade-off between RRNS's flexibility of stratification and its worse (according to some criteria) spread of burden that determines whether it is a good alternative to SBS.

In conclusion, the new RRNS method is a valuable method to spread the survey burden over time and across a set of sample surveys. It is easier to apply in practice, because unlike SBS it imposes no restrictions on stratifications. This latter restriction of SBS is the reason that its current use at Statistics Netherlands mostly spreads the burden for each survey over time but not across surveys. It is expected that by using the RRNS method for drawing samples for the different (non-panel) business surveys of Statistics Netherlands the spread of survey burden will further improve. For example, the number of times a business will be in two or more surveys within the same period is expected to decrease to a minimum. The number of hotspots will probably also decrease somewhat, although large reductions of the hotspot measure can only be achieved by decreasing sampling fractions, thereby lowering the overall burden. That said, the SBS method still has the advantage of being able to account for differences in burden caused by different surveys. And for panel surveys SBS accounts for the duration of panel membership as well.

The new RRNS method has been implemented in R package SBS used at Statistics Netherlands, which now allows choosing between SRS, SBS and RRNS methods. One possible future extension will be the generalization of RRNS to PPS samples, i.e. samples with unequal inclusion probabilities within strata, proportional to some size measure. For SBS a preliminary study (Smeets and Boonstra 2019) has already shown that this seems possible. Further research is also needed to see whether the RRNS method can be improved for the situation where both negative and positive coordination must be applied, i.e. in the case of (possibly multiple) panel surveys.

Acknowledgments

The views expressed in this paper are those of the authors and do not necessarily reflect the policies of Statistics Netherlands. The authors are grateful to Rob Willems, Leanne Houben en Remco Paulissen for their comments.

References

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

Boonstra, H. J. and M. J. E. Smeets (2015). *Enquêtedrukspreiding voor panels and niet-panels - Proof of Concept. Internal CBS report*. PPM-2015-06-05-HBTA-MSET.

– (2018). *Enquêtedrukspreiding voor stratumwisselaars. Internal CBS report*. PPM-2018-02-27-HBTA-MSET.

Boonstra, H. J., M. J. E. Smeets, and J. Klingwort (2024). *Enquêtedrukspreiding voor hotspots. Internal CBS report*.

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

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

Erikson, J., D. Giesen, L. Houben, and P. Saraiva (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*. Hoboken: John Wiley & Sons, pp. 193–223. DOI: 10.1002/9781119672333.ch10.

Ernst, I. (1999). “The maximization and minimization of sample overlap problems: a half century of results.” In: *Proceedings of the International Statistical Institue* 52. Helsinki, Finland, pp. 168–182.

Giesen, D., M. Vella, C. F. Brady, P. Brown, D. Ravindra, and A. Vaasen-Otten (2018). “Response Burden Management for Establishment Surveys at Four National Statistical Institutes”. In: *Journal of Official Statistics* 34.2, pp. 397–418. DOI: 10.2478/jos-2018-0018.

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

Ippel, L., G. Gommans, R. Paulissen, L. Burgerjon, R. Kitzen, J. Klingwort, M. Smeets, L. Houben, and G. Snijkers (2023). *Hotspots project. Evaluatie en geleerde lessen. Internal CBS memo*.

Klingwort, J. and M. J. E. Smeets (2023). *Hotspot Analyses Using the Research & Development and Community Innovation Survey. Internal CBS report*.

Matei, A., P. Smith, M. Smeets, and J. Klingwort (2023). “Targetted double control of burden in multiple surveys”. In: *Survey Methodology* 49.2, pp. 363–384.

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

Smeets, M. and H. J. Boonstra (2019). *Enquêtedrukspreiding bij PPS-steekproeven*.

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

Smeets, M. and J. Klingwort (2023). “Sampling methods to better spread the response burden for small businesses”. In: *CBS Discussion paper*. Den Haag: Centraal Bureau voor de Statistiek, pp. 1–30.

Strategische commissie betere regelgeving bedrijven (2020). *Onevenredige enquêtebelasting van specifieke ondernemers in het MKB*.

Appendices

A Overview of sampling methods used

	SRS	SBS		RRNS		
feature		SBS	SBS1	RRNSa	RRNSb	RRNSc
coordination over time	no	yes	yes	yes	yes	yes
coordination over surveys	no	no	yes	yes	yes	yes
common stratification required	no	no	yes	no	no	no
account for built-up survey burden	no	yes	yes	no	no	no
type of random numbers	-	permanent	permanent	rotating, not regenerated	rotating, regenerated independently	rotating, regenerated equidistantly

Table A.1 Overview of the sampling methods compared in the simulation study

Colophon

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress
Statistics Netherlands, Grafimedia

Design
Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contact form: www.cbs.nl/information