SYNCHRONISED SAMPLING

Richard McKenzie and Bill Gross, Australian Bureau of Statistics
Bill Gross, Australian Bureau of Statistics PO Box 10, Belconnen, ACT, 2616, Australia
bill.gross@abs.gov.au

ABSTRACT

Synchronised Sampling is a Permanent Random Number method of selection to control sample rotation within repeated stratified surveys and to control overlap between different surveys. The sample in each stratum is specified as an interval and overlap is achieved by, for each stratum, constraining the selection interval to move within a survey range. While it does not automatically control rotation or overlap when the stratification of a survey changes, some control is achieved by careful choice of selection intervals and survey ranges. The paper describes the basic algorithms for synchronised sampling and the techniques used to set survey ranges and selection intervals when the stratification changes and assesses the effectiveness of the techniques.

Key Words: Coordination of surveys, Permanent Random Number sampling, Change of stratification

1. INTRODUCTION

Synchronised Sampling is a technique which has been used by the Australian Bureau of Statistics since 1983 to control sample rotation within surveys and overlap between surveys. It relies on assigning a permanent random number to each unit on the business register and selecting units whose random numbers lie in an interval. Rotation control is achieved by moving the interval to the right and overlap is achieved by, for each stratum, constraining the selection interval to move within a survey range.

The paper describes the basic algorithms for synchronised sampling and the techniques used to set survey ranges and selection intervals, especially after major changes in stratification. It also gives an assessment of how effective the techniques for setting survey ranges and selection intervals are.

2. DESCRIPTION OF SYNCHRONISED SAMPLING

Synchronised Sampling (Hinde and Young 1984) was developed in ABS in the early 1980's. The selection method was adapted from the JALES method (Atmer et al 1975). A more extensive description is given in Brewer et al (2000).

2.1. Basic Method

Each unit has a permanent random number from the uniform distribution on [0,1). The selections are specified as an interval in [0,1), whose start and end points lie on random numbers of units at the time of selection. The start point is in sample but the end point is not. To achieve a desired sample size, n, while allowing for births and deaths in the population, the start or end points are moved, but only to the right, to prevent units re-entering the sample.

For the first selection, a provisional start point, s_0 is set and the first n units at or to the right of s_0 are selected. They are described by the interval $[s_1,e_1)$. For the next selection, only the start and end points s_1 and e_1 need to be remembered and the number of units from the new population in the trial interval $[s_1,e_1)$ is compared with the desired sample size n. If it contains n units (i.e. there have been zero net births or deaths in the selection interval) then it remains the selection interval. If it contains more than n (i.e. there have been net births in the interval), the start point moves to the right, while if it contains less than n (i.e. there have been net deaths in the interval), the end point moves to the right. The same process applies if the sample size has changed from time 1 to time 2. For the case n=3, Fig 1 illustrates the first selection and Fig 2 shows selection after population births.

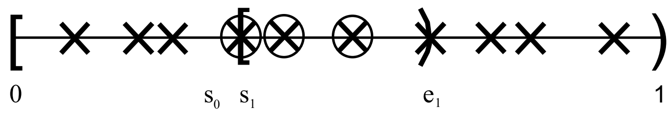

Fig 1: First selection. ($\otimes$: selected unit)

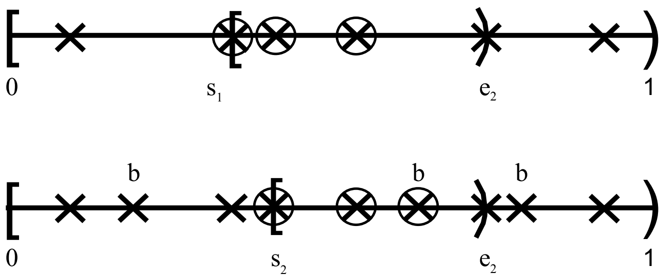

Fig 2: Selection after population births (b).

2.2. Control of Rotation

Planned rotation removes a unit after no more than R consecutive survey cycles in sample. R can be stratum specific. When a sample of size n is first selected, it is partitioned into rotation groups of sizes n_1,..,n_R by projected starts, p_r, (r=0,..,R) where $p_0=s_1$, $p_R=e_1$, so that, at the time of initial selection, there are n_r units in each interval [p_{r-1}, p_r). At the next selection, [p_1, e_1) is used as the trial interval and the start and end points are moved as for the basic method to select n units. Hence s_2 is at, or to the right of p_1. New projected start points are set by dropping p_0, shifting the numbering of the rest and taking p_R as the new end point e_2. At the (R+1)'th selection, the start point s_{R+1} is at, or to the right of $p_R=e_1$, and so all units in the first sample have been rotated out. If there is a significant number of population births, then the start point may shift over more than one projected start, and the trial interval is modified to ensure that s_{r+1} is at, or to the right of p_r. For example, if $s_2 > p_2$, then the trail interval for the third selection is [s_2, e_2) rather than [p_2, e_2). Fig 3 illustrates a simple example.

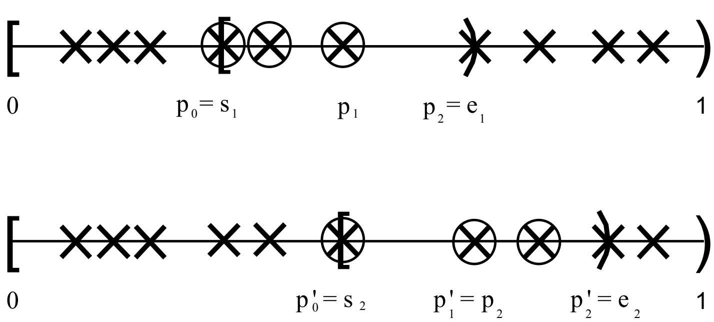

Fig 3: Rotation with pattern 2,1.

Due to planned rotation and to population changes, the selection interval moves to the right. At some time, there will not be enough units in the interval [s,1) and the additional units are selected to the right of 0, so that the selection interval becomes split into [s,1) and [0,e).

2.3 Minimising Overlap between Surveys

The aim is to prevent a unit being selected in more that one survey. Synchronised Sampling, operated correctly, can achieve this even if the surveys concerned have different stratification and different rotation rates, provided that the sum of the sampling fractions in a sub population is not too high. Protection from selection in other surveys is achieved at the expense of a decrease in the time before a unit may be reselected in the original survey. However, particularly in the 'small units' strata, where the sampling fractions are low, the reduction in the time before reselection will not appreciably increase the individual's perceived response burden.

Where there is rotation, the position of the selection interval changes from one cycle of a survey to another. To control overlap between surveys, each stratum of each survey is given a fixed interval, the survey range, and the selection interval is forced to remain within the survey range. The survey range should be large enough to hold the selection interval, plus enough unselected units to permit worthwhile rotation to occur.

In the simplest case where two surveys have the same stratification, the samples can be made disjoint by setting the survey ranges in each stratum to be disjoint. If the surveys have different stratifications then the constraints on the survey ranges are more complex. To ensure that the sample from a stratum in one survey is disjoint from the sample in another survey, the survey range for the stratum of the first survey must be disjoint from any of the survey ranges of intersecting strata of the second survey. It follows that groups of strata must be considered simultaneously - if two strata from the surveys intersect, then their survey ranges need to be considered simultaneously, and hence two strata linked by a chain of intersecting strata need to have their survey ranges considered simultaneously. To limit the length of such chains, some standard boundaries for industry and size are used. The process of allocating the position and length of the survey ranges for all the strata for all surveys is complex and is done manually with computer produced diagnostics. Hinde and Young (1984) describe how this was done when Synchronised Sampling was first used for a group of surveys. Section 3 describes the tools used to set survey ranges when a new survey is added to that group or when the stratification of an existing survey is changed.

2.4. Relationship with Stratification

While the same permanent random number is used for all surveys, at the system level, the method operates separately within each stratum in each survey. In particular, start points, survey ranges and rotation rates are stratum specific. As a consequence there is no necessary relationship between the stratifications for different surveys, although some standard boundaries are used for industry and size to make it easier to control overlap between surveys. There is also no system limitation on the number of surveys which can use Synchronised Sampling, although the total sampling fraction constrains the effectiveness of overlap control and the setting of survey ranges is more complex with a larger number of surveys.

When a unit changes stratum, Synchronised Sampling treats it as a death in the old stratum and a birth in the new stratum, so it cannot automatically take account of its selection status in the previous stratum. This is a major drawback when a survey is redesigned and the stratification is changed, since there are a very large number of stratum changes. However, it is highly desirable to maximise the common sample under the two stratifications, both to minimise the sample error on estimates of change and to minimise cost as new units in sample are more expensive to process. Synchronised Sampling can achieve this to a limited extent by appropriate choice of start points in the new stratification.

In summary, the key properties of the method described above are:

- it controls overlap between a number of surveys
- it achieves pre assigned sample sizes in each stratum
- it copes with different stratification in different surveys
- it controls rotation by guaranteeing a maximum time in sample
- the rotation rate can be stratum specific
- it allows for births and deaths on the frame

Properties described in Brewer et al (2000) are:

- it gives very nearly simple random sampling within stratum
- it can control overlap for single establishment firms between an establishment survey and a firm survey
- it does not automatically control rotation or overlap when units change stratum

3. TOOLS FOR SETTING START POINTS AND SURVEY RANGES

Start points for selection intervals need to be reset when a survey is redesigned and the stratification changes markedly, and it is desirable to set these start points in a way which maximizes the common sample under the two stratifications. Under the old stratification the sample generally consisted of one interval in each stratum, but would be two intervals if the selection interval had split because it had recently reached the end of its allocated survey range, and recommenced taking new selection from the start of the survey range. Viewed in the new stratification, the old sample consists of a larger number of "partial" intervals – one or two for each old stratum which intersects with the new stratum. A "partial' interval contains some, but not necessarily all of the points between two numbers, and a selection interval in one old stratum is a partial interval in the new stratum because only units from that old stratum population are selected. Units in the interval from a different old stratum would not, in general, be selected, and these 'dilute' the selection interval and make it partial in the new stratum. Figure 4 illustrates a simple case where two strata are combined. As the sample, of size n, for the new stratum is specified by a start point and consists of the n units at, or to the right, of the start point, the aim is to find a start point which maximizes overlap between the old and new samples.

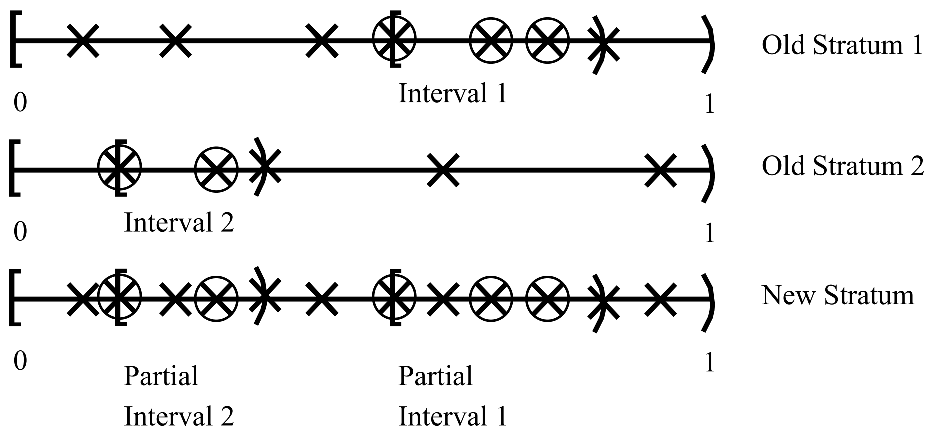

Fig 4 Selections when two strata are combined

The problem is similar when setting the survey ranges for a new or restratified survey, especially when the stratification is different from the existing surveys. Under the new stratification the survey ranges of the other surveys consist of a number of intervals or partial intervals. The aim, however, is to find an interval which minimizes overlap with the existing intervals and partial intervals. The survey ranges or selection intervals for the new strata could also be chosen to minimise overlap with the existing samples. For clarity, the rest of the description refers only to selection intervals.

In developing the tools to set start points, a key issue was how to represent and handle the partial intervals on the number line, especially when a number of different stratifications are involved. In the approach taken the realized number line is treated as a set of discrete numbers – the ones corresponding to units in the stratum - rather than as a continuum. Each unit in the new stratum is flagged, to indicate whether or not it was selected in any of the pre-existing surveys. Clearly, the unit's stratum in each pre-existing survey is used to find the appropriate selection interval and thence to set the selection flag for that survey. However, information about the old stratifications is not needed after that. In this way, any complicated relationships between the stratifications are not treated explicitly, but they are implicit in the distribution of the flags in the population.

The method of choosing the new start point is also simple. Each unit in a stratum population is examined as a possible start point and given a score which measures how well the sample with that start point meets the desired overlap criteria. In order to calculate the overall score for a possible start point, a score is first calculated for each population unit, by aggregating a weight for each survey in which it is selected. The weight is positive if maximum overlap with the survey is desired and negative if minimum overlap is sought and the magnitude of the weight reflects the importance of controlling overlap with a survey, e.g. because it places a high load on respondents. For the new stratum, the desired sample size is known, and so for each possible start point, the sample which would result can be determined. The overall score for a possible start point is the sum of the unit level scores over the sample it would generate. To maximise the common sample under new and old stratifications, the point after the last occurrence of the maximum score is chosen as the start point for the selection interval. The point after the last occurrence of the maximum is chosen because using the last occurrence leads to a bias, and using the next point reduces the bias. Flack et al (2000) consider a simple case where two strata are combined, sample is selected in only one of the old strata (stratum 1) and the sampling fraction in the new stratum is much greater than in old stratum 1. In this case the bias towards selecting from stratum 1 is of order 1, if the last maximum is used, but is of order 0, if the next start point is used.

In practice the criteria for overlap may be complex – maximizing common sample as well as avoiding samples or ranges for other surveys with different levels of importance – and the weights may not fully reflect this complexity. To help assess these more complex situations, the scores for each start point are plotted against random number. In the same plot, the samples for the other surveys involved are also shown. Figure 5 shows the graph for choosing a survey range which minimizes overlap with the samples from eight other surveys. The weight for each of the surveys is –1 and a range of length 0.39 was sought. The random numbers for the samples for the eight surveys (AWE, JVO, etc) are shown as short horizontal intervals to the right of the survey acronym. The maximum score is –1, which corresponds to the interval shown and contains only the second point in the EEH sample.

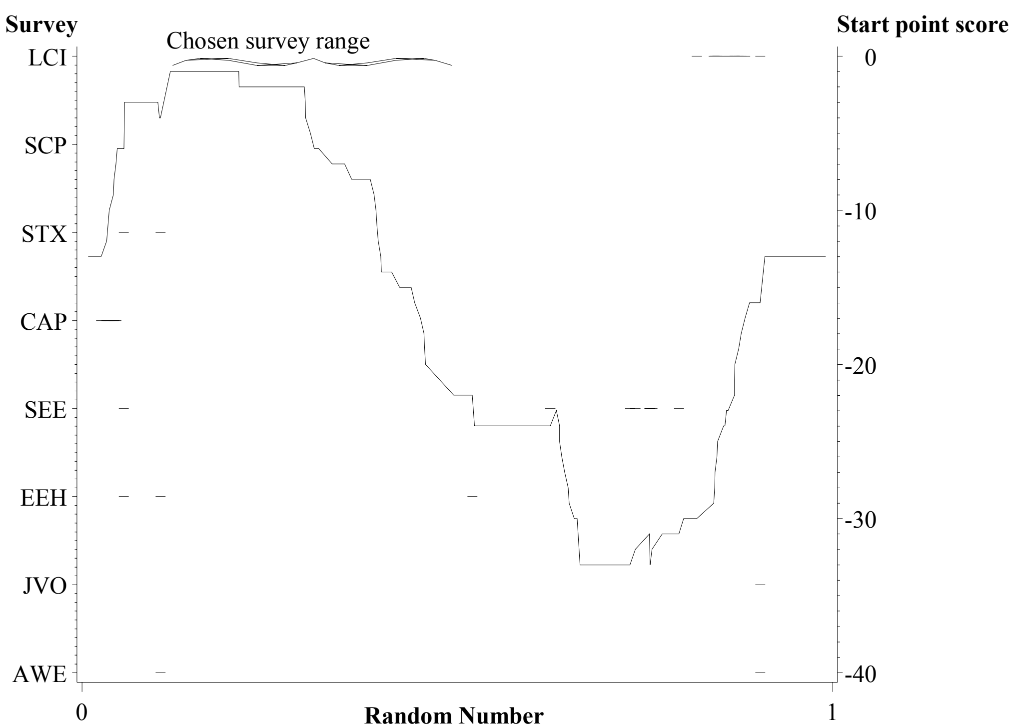

Fig 5: Graph of scores for potential start points, showing locations of existing survey samples

4. EFFECTIVENESS OF ROTATION CONTROL AT A SURVEY REDESIGN

When the stratification and allocation of a survey change some rotation is unavoidable as the probabilities of selection change. In a sub-population where the probability of selection decreases then some of the sample must be rotated out to realize the lower probability of selection. For unit i, with probabilities of selection p_{1i} under the old design and p_{2i} under the new design, the maximum probability of it being in the common sample is $p_{max\ i} = \min(p_{1i}, p_{2i})$. (Setting p_{1i} or p_{2i} to zero covers changes in scope.) The effectiveness of a method to maximise the common sample can be assessed by comparing the achieved common sample with the maximum attainable, $\Sigma_i\ p_{max\ i}$, the population total of $p_{max\ i}$.

The graphical method described in section 3 was used to set new start points for the 1999 redesign of the ABS Engineering Construction survey. The frame changed markedly from about 15,000 to about 26,000 with about 9,000 in common. The old design had a sample size of 2,500 from 129 strata, while the new design was for a sample of 3,000 from 265 strata. While the maximum attainable common sample was 1628 the start points chosen using the method gave a common sample of 501, or 30% of the maximum. As this is the best attainable using Synchronised Sampling, the example illustrates Synchronised Sampling's limited ability to maximise common sample at survey redesigns.

5. EFFECTIVENESS OF OVERLAP CONTROL

When assessing the effectiveness of a method for controlling overlap between a group of surveys it is useful to consider two aspects. The first is the expected load for a unit - defined as the sum of its probabilities of selection in the surveys in the group. The expected load is determined by the allocations for the surveys i.e. the sampling fraction chosen for each stratum.

The second is the distribution of the load (defined as the number of surveys in which a unit is selected) conditional on the expected load. It is this conditional distribution of load that the selection method affects. The aim of controlling overlap between surveys is to minimise the number of units which have a high load and the ability to achieve this is constrained by the expected load. For example, if the expected load is 1.5, then it is not possible for it to have a realised load which is only ever 0 or 1, because then the expected load on a unit would be less than or equal to 1. Hence there must be some chance of selecting the unit in more than one survey. More generally, it is easy to see that, if p_i and d_i are defined so that

p_i is an integer

$0 \leq d_i < 1$

and

expected load = p_i+d_i

then unit i must have a chance of being selected in p_i+1 surveys. A selection method which ensures that there is no chance of unit i being selected in more than p_i+1 surveys is one which results in

Pr(unit i selected in p_i+1 surveys) = d_i

Pr(unit i selected in p_i surveys) = $1-d_i$.

Such a selection method is optimal in terms of minimizing high load.

While such a method is optimal for minimizing high load, it also gives the minimum spread of load between units, which is desirable as it makes the distribution of load fair and equitable.

For a selection method, avoidable load is defined as the number of selections in multiple surveys in excess of the optimal p_i+1 for a unit with expected load p_i+d_i. More formally, if unit i with expected load p_i+d_i is selected in m_i surveys then its contribution to avoidable load is:

$$\text{avoidable load (i)} = \begin{cases} 0 & \text{if } m_i \leq p_i+1 \\ m_i - (p_i+1) & \text{if } m_i > p_i+1 \end{cases}$$

The avoidable load for the method is the population sum, Σ_i avoidable load$_i$, of each unit's contribution. We take the avoidable load as the measure of how effective a selection method is in controlling overlap between a group of surveys.

In practice, there is a hierarchy of units used in business surveys (e.g. establishment as a sub-unit of a firm) and the notions of load and expected load can be modified to allow for using more than one level in the hierarchy. Load is attributed to the highest level unit in the hierarchy and is measured by the number of sub-units selected. Expected load is the sum of the probabilities of selection of sub-units.

The effectiveness of Synchronised Sampling in controlling overlap was assessed for medium sized businesses (defined as having between 20 and 200 employees) in the twelve ABS surveys which used Synchronised Sampling in 1999. In these medium sized businesses, the sampling fractions can be quite large and so they severely test the method's effectiveness in controlling overlap. Table 1 shows avoidable load for the businesses as well as the distribution of actual load against expected load. The shaded cells represent avoidable load, i.e. counts of business which in practice are selected more times than their expected load would suggest is necessary. It shows that Synchronised Sampling does not control overlap very well for these businesses. The proportion of total load which is avoidable is reasonable at 12% of total load, but where the expected load is less than 1, the avoidable load is high, at 17%.

Table 1: Counts of Medium sized units by Expected Load and Actual Load

Expected Load	Actual Load (Number of times selected)							Avoidable Load	Total Load	% Avoidable Load
	0	1	2	3	4	5	6+			
[0,1]	25433	6941	1246	185	17	0	0	1667	10056	17
(1,2]	489	2180	1180	384	77	7	0	559	6035	9
(2,3]	17	141	457	298	106	24	5	169	2523	7
(3,4]	0	3	25	102	78	35	9	55	902	6
(4,5]	0	1	3	17	60	64	17	20	723	3
(5,6]	0	0	0	2	9	42	38	7	487	1
(6, ∞)	0	0	1	0	0	4	62	17	497	3
Total	25939	9266	2912	988	347	176	131	2494	21223	12

(Shaded cells contribute to avoidable load.)

While Table 1 shows the full range of expected load, it includes load in take all strata. However, no selection method can redistribute the load from take all strata, and so it is not useful to consider them in assessing the effectiveness of a selection method in controlling the distribution of load. Table 2 shows the avoidable load and distribution of load when take all strata are excluded, i.e. it applies to sampled strata. Units which had more than one sub-unit were also excluded from the analysis.

Table 2: Medium size simple units in sampled strata - Avoidable Load and counts by Expected Load and Actual Load

Expected Load	Actual Load							Avoidable Load	Total Load	% Avoidable Load
	0	1	2	3	4	5	6+			
[0,1]	24331	6729	1239	188	19	0	0	1672	9847	17
(1,2]	498	1151	717	251	61	5	0	388	3607	11
(2,3]	10	44	64	43	19	6	2	37	419	9
(3,4]	0	2	0	0	1	0	0	0	6	0
Total	24839	7926	2020	482	100	11	2	2097	13879	15

(Shaded cells contribute to avoidable load.)

Considering only simple units in sampled strata reduces the incidence of high expected load dramatically, and increases the percentage avoidable load slightly from 12% to 15%. As there is little change in the population of units with expected load less that 1, there is no change in the proportion of their total load which is avoidable.

ABS practice is to have survey ranges which are at least three times the sampling fraction. Hence Synchronised Sampling should control overlap very well if the expected load is less than 0.3, as it should be possible to have survey ranges which are disjoint if surveys use the same stratification. Table 3 shows detail of the avoidable load for units where the expected load is below 1.0. It indicates that, while the avoidable load is less for these low expected loads, it is still quite large. This is due to diversity in the stratification or sub-optimal setting of survey ranges.

Table 3: Medium size simple units in sampled strata - Avoidable Load for Expected Load below 1.0.

Expected Load	Avoidable Load	Total Load	% Avoidable Load
[0.0,0.1]	6	350	1.7
(0.1,0.2]	60	989	6.1
(0.2,0.3]	97	1178	8.2
(0.3,0.4]	124	1095	11.3
(0.4,0.5]	189	1328	14.2
(0.5,0.6]	228	1277	17.8
(0.6,0.7]	231	982	23.5
(0.7,0.8]	278	1062	26.2
(0.8,0.9]	265	936	28.3
(0.9,1.0]	194	650	29.8
Total	1672	9847	17.0

ACKNOWLEDGEMENTS

Tenniel Guiver and Anna Poskitt developed the programs for the method described in section 3. John Kieley, Lloyd Flack and Doug Lincoln applied them and computed the effectiveness of the methods for sections 4 and 5.

REFERENCES

Atmer, J.G., G. Thulin, and S. Bäcklund (1975).”Coordination of Samples with the JALES Technique,” *Statistik Tidskrift*, 13, pp. 443-450. (In Swedish with English summary).

Brewer, K.R.W., W.F. Gross, and G.F. Lee (2000). “PRN Sampling: The Australian Experience,” *ISI Proceedings: Invited Papers, IASS Topics, Helsinki August 10-18, 1999*, pp. 155-163.

Flack, L., W. Gross and A. Herning (2000). “Selection Bias when Maximizing Common Sample at Restratification using Synchronised Sampling,” unpublished report, Australian Bureau of Statistics

Hinde, R. and D. Young (1984) "Synchronised Sampling and Overlap Control Manual," unpublished report, Australian Bureau of Statistics.

ISSUES IN COORDINATED SAMPLING AT STATISTICS CANADA

Don Royce, Statistics Canada
R.H. Coats Building, 11th Floor, Ottawa, Ontario, K1A 0T6, Canada
roycdon@statcan.ca

ABSTRACT

Until the mid-1990s, the environment for business surveys at Statistics Canada was not conducive to the use of coordinated sampling. With the arrival of the Project to Improve Provincial Economic Statistics in 1997, however, the situation began to change. One of these changes has been the development of a new Unified Enterprise Survey, which integrates the frames and the sample designs of many of our annual business surveys. This and related developments have raised many opportunities for Statistics Canada concerning the future use of coordinated sampling. Some related initiatives to reduce response burden are also described.

Key Words: Business Register, Generalized Software, Response Burden, Random Numbers

1. INTRODUCTION

As described by Ohlsson (1995), coordinated sampling refers to a wide variety of techniques, often based on permanent random numbers (PRNs), used to coordinate the units selected in samples for a program of business surveys of a statistical agency. The reasons for wanting to coordinate samples are many. In some cases, reducing response burden is the objective – we wish to avoid having a business fall into too many surveys within a short time frame – and in such cases, *negative coordination* of samples is the goal. In other cases, we may wish to maximize the overlap between samples, or to have *positive coordination*. This may be because the sample for a survey is being redesigned and we wish to minimize the disruption to the estimates and the costs of transition to the new sample. It may also be because we wish to use data from one survey in the editing or analysis of the data from another survey, or because we wish to reduce the costs of collecting data from a completely independent sample. In still other cases, we may want elements of both positive and negative coordination; for example we may wish to maintain a high degree of sample overlap between occasions of a monthly survey, while still ensuring that units remain in sample for a limited period of time.

Examples of all of these reasons and more can be found at Statistics Canada. Why, then, has Statistics Canada not adopted a systematic approach to coordinated sampling, as have statistical agencies in Sweden, Australia, France and New Zealand? As later sections of this paper will show, various methodologies, sometimes quite complex and sophisticated, have been developed to deal with each of the situations described above. However these methods have generally been developed in isolation from each other, applying to only one or a few surveys at a time. In this paper, I explore some of the reasons for the absence of a more systematic approach, and point to some indications that I believe will enable Statistics Canada to take a more unified approach to the issue of coordinated sampling in the future.

Much of the change in environment is due to a new project known as the Project to Improve Provincial Economic Statistics (PIPES), which is dealt with in some length in Section 3 of the paper. Prior to describing PIPES and its impact, however, I describe the situation with business surveys at Statistics Canada before PIPES in Section 2, including a description of some of the methods used for sample rotation and overlap. Section 4 then looks at the future, describing some of the developments planned for the next few years in areas that should facilitate a more systematic approach, such as the use of the Business Register (BR) and the use of the Generalized Sampling System (GSAM). Finally, Section 5 describes some additional measures Statistics Canada is considering related to the response burden issue.

2. BUSINESS SURVEYS AT STATISTICS CANADA BEFORE 1997

During the mid-1980s, business surveys at Statistics Canada were characterized by their isolation from each other. Responsibility for various industry sectors was divided among a number of divisions, each of which maintained its own frame, sampling methodology, collection operations, processing systems and dissemination channels. There was of course some standardization of concepts and classifications (e.g., the 1980 Standard Industrial Classification, the System of National Accounts), but other than this, divisions were generally free to design their programs to meet the individual needs of their clients. Unlike many countries, there was no periodic economic census. There were annual surveys of most industries, and sub-annual surveys for industries (e.g., manufacturing, distributive trades) where current information was required. There were also several cross-industry surveys on specific topics, such as financial statistics, employment, and capital investment, as well as numerous special surveys on various types of economic activity or on specific commodities.

In 1985, Statistics Canada received funding for a three-year redesign of its business surveys program. The Business Survey Redesign Project (Colledge, 1987) had, as a main objective, the development of certain common elements of a survey-taking infrastructure. Chief among these was a new Business Register, a central register of all businesses in Canada, which was to be used for selecting samples for all business surveys. Other infrastructure elements included common software tools for sampling, collection, editing, imputation and estimation, and a common Tax Estimates Program for sampling and processing income tax records from Revenue Canada (now the Canada Customs and Revenue Agency).

By the time the project formally ended, only one major survey – the Monthly Wholesale and Retail Trade Survey - was fully "hooked" to the new Business Register. The groundwork had been laid, however, and during the years that followed (1992 to 1996), a few surveys were hooked up each year. Unfortunately, due to lack of funding, this incremental and slow-paced approach again led to different flavors of hookup, different sampling methodologies, and different software being used for each survey.

Many of the major programs, among them agriculture, manufacturing, and financial statistics continued to keep their own frames, each one maintained from different sources. In the case of the Business Register, the main source was payroll deduction accounts from Revenue Canada. The Agriculture Division maintained its own Farm Register, based on the Census of Agriculture held every five years in conjunction with the Census of Population. In the case of the manufacturing program, the main source was annual income tax records, sometimes supplemented with subject matter knowledge gained from the media. Other surveys made extensive use of industry association lists. In regulated industries, Statistics Canada often entered into partnerships with the regulating body, using their lists as the frame. In the case of the Mining industry, Statistics Canada delegated the collection to another federal department, Natural Resources Canada, which collected the data using its own frame, questionnaire, and processing systems.

Within statistical programs, however, there was usually some attempt to coordinate approaches. This extended to the coordination of different samples within these programs. Several examples of the sample coordination approaches used within various programs follow. The examples are roughly in chronological order.

2.1 Rotating Panel Design – the Monthly Wholesale and Retail Trade Survey

The Monthly Wholesale and Retail Trade Survey (MWRTS) was designed in the late 1980s with sample coordination in mind. In order to control response burden, the sample uses a rotating panel design, where the number of panels depends on the sampling fraction in the stratum and the time-in and time-out constraints. Sampling units are initially assigned to panels in a systematic fashion that balances the panels to within one unit. Rotation is achieved by dropping one panel each month and adding a new one. Births are added each month, following the systematic ordering so as to maintain the balance among panels as much as possible. Since the sample size is variable, a ratio estimate is used. As well, deaths can cause the panels to become unbalanced after some time, leading to a need for re-balancing. Procedures for doing this while maximizing sample overlap were developed, as were procedures for re-stratifying the sample while maximizing overlap between the old and new samples. These methods use the Kish and Scott (1971) method adapted for a rotating panel design (Hidiroglou et al 1991). The rotating panel approach ensures full control over the time-in and time-out constraints, along with relatively good stability of the achieved sample size. One disadvantage of the method is that it does not lend itself to coordination with other surveys. As well, the methods can become quite complex, even involving the use of "dummy" panels in some cases.

The rotating panel methodology was implemented in software known as the Generalized Sampling System (GSS), which was intended to become the basis of sample designs for other surveys. However, developments on a different sampling system, known as GSAM, were also underway and eventually overtook work on GSS. Today, the MWRTS is the only survey still using GSS.

2.2 Use of Bernoulli Sampling in the Tax Estimates Program

The Tax Estimates Program that was originally designed under the Business Survey Redesign Project provides another example of the use of coordinated sampling techniques at Statistics Canada. The sample was chosen from tax records at Revenue Canada as tax forms were captured and processed. Since there was no complete frame available in advance of the sampling operation, stratified Bernoulli sampling was used to select the sample. That is, each tax record was assigned a random number in the interval [0,1], and if the record's random number fell within the sampling interval $[0, f_h)$ then the unit was selected.

Since there is no respondent burden involved with using tax records, maximum overlap of samples from consecutive years is desirable for the purpose of measuring year-to-year changes. Sunter (1986) describes a useful technique for maximizing the overlap in such situations. The technique involves the use of a hashing algorithm, which transforms the unique identifying numbers on the administrative files to pseudorandom numbers used to implement the Bernoulli sampling. Since the identification numbers remain the same from year to year, the pseudorandom number generated by the hashing algorithm is also the same, thus maximizing the overlap.

2.3 Surveys Based on GSAM

At nearly the same time as the development of GSS, some other surveys in Statistics Canada were taking a different approach to the need for sample rotation while keeping sample sizes relatively stable and allowing re-stratification on occasion. The method adopted by these surveys was a variant of collocated sampling first described by Brewer et al. (1972). The method was adopted as the official method of sample selection for a new system, also called the Generalized Sampling System but with the acronym GSAM, and was implemented in a number of sub-annual surveys, including the Monthly Restaurants, Caterers and Taverns Survey, the Quarterly Motor Carrier Freight Survey and the Monthly Survey of Manufacturing. The method is as follows:

Selection numbers ν_i are calculated for all units within each stratum h.

$$\nu_i = \frac{\text{rank}(u_i) - \varepsilon_h}{N_h} \text{ for unit i in stratum h, where } u_i, \varepsilon_h \text{ are random values from U(0,1),}$$

$\text{rank}(u_i)$ is the order of u_i among $u_1, u_2, \ldots, u_{N_h}$

and N_h is the number of units in stratum h.

Note The value of ε_h provides a random start in the assignment of equally spaced selection numbers on the [0, 1] interval.

The sample is then selected within each stratum according to pre-defined sampling rates f_h (these rates are either computed within or outside GSAM), by selecting (initially) all units whose selection numbers fall in the interval $[0, f_h)$. This interval is gradually shifted to allow the rotation of units. As with the rotating panel approach, the rotation scheme can take into account specific time-in and time-out constraints. Births are added each month by assigning them their own selection numbers in the [0,1] interval, independently of the selection numbers for units already in the stratum.

Over time, of course, the sample size can become more variable due to both deaths and the clustering of the selection numbers for births. At this point the sample can be re-balanced by recalculating the selection numbers (but keeping the u_i unchanged). Sample overlap can be maximized for either re-balancing or re-stratification by judicious choice of the selection interval. For example, a re-stratification that requires a higher sampling rate can be achieved by increasing the upper boundary of the sampling interval. More complex operations are described in Hidiroglou, Carpenter and Estevao (1996) and Statistics Canada (1996).

2.4 "Piggyback" Surveys

Sections 2.1 to 2.3 gave examples of sample rotation and/or positive coordination within specific surveys. However, during the mid-1990s a number of programs also adopted approaches that maximized sample overlap between related surveys. In the case of wholesale and retail trade, a so-called "Period 13" approach was developed that saw the Annual Wholesale and Retail Trade Survey use the sample from the Monthly Wholesale and Retail Trade Survey. The approach was called "Period 13" because the annual survey's questionnaire was mailed shortly after December's monthly questionnaire (Period 12). In addition, a Quarterly Retail Commodity Survey used a sub-sample of the retail component of MWRTS as its sample (Bérard et al. 1999; Brodeur et al. 1997). A similar approach was used in the case of motor carrier transport, where the Annual Survey became a "Quarter 5" supplement to the Quarterly Motor Carrier Freight survey. A number of ad hoc surveys also piggybacked on existing surveys.

This approach provided many advantages. The costs of sample design were negligible, collection costs were much lower since the respondent was already in the monthly survey, and data from the monthly survey could be used in editing and imputing the piggyback surveys. In the case of the Annual Wholesale and Retail Trade Survey, the data from the monthly survey were also used to calendarize data from the annual survey for respondents who provided fiscal year instead of calendar year data. The high degree of sample overlap meant that estimates from the monthly and annual surveys were reasonably consistent.

A disadvantage of the approach, however, was that respondents selected in one survey had to respond not only to the monthly survey, but also to the annual survey and possibly a commodity survey as well. However the advantages of maximizing the overlap of samples for related surveys were felt to outweigh the disadvantage of a more highly concentrated response burden.

2.5 Agriculture Surveys

In the case of Statistics Canada's agriculture statistics program, the program has long maintained its own Farm Register. The Farm Register undergoes a major update every five years using the quinquennial Census of Agriculture, and is used for all agriculture surveys. Denis et al (1999) describes a method used to coordinate the samples for a series of crop surveys. These surveys are conducted six times per year on various crop-related topics; however the sample size on each occasion is different, and various degrees of overlap, ranging from none to partial to complete, are desired for these various samples. In addition, annual rotation of the samples is desired. The solution used is to divide the population (excluding large take-all units) into a number of equal-size replicates within each sampling stratum, where the number of replicates depends on the desired overlap pattern. A master sample is then selected within each stratum-replicate combination. Sub-samples are chosen by selecting the appropriate replicates to satisfy the required sample sizes and overlap patterns for the six survey occasions.

Grondin (1997) also describes the method used for minimizing overlap between the crop surveys and a livestock survey. Many farms fall into the target populations for both surveys. The collocated sampling method described in section 2.2 was used for selecting both samples. Selection intervals had to be chosen carefully to minimize overlap between the two surveys throughout the intercensal period, taking into account sample rotation. Various scenarios were simulated, using various selection intervals and sample rotation hypotheses. The option chosen uses the interval $[0, f_c)$ for the crop survey and the other end of the interval for the livestock survey, i.e., $(1-f_l, 1]$. Rotation for the crop surveys is achieved by shifting the selection interval to the right, while for the livestock survey the interval shifts to the left. Overlap is minimized for the first year and is fairly well controlled for subsequent years. This option ensures minimal overlap over five years.

2.6 Survey of Employment, Payrolls and Hours (SEPH)

SEPH is a monthly survey of employers whose objective is to produce statistics on employment, payrolls and hours worked. The survey makes extensive use of administrative data, with a sample of about 200,000-payroll deduction accounts sampled using Bernoulli sampling each month. In addition, there is a small sample of about 10,000 establishments, for whom more detailed data are collected. The data from the two sources are combined in a modelling framework to produce detailed estimates for the population. The establishment component uses the Business Register as the frame, and GSAM is used to select and rotate the establishment sample on a monthly basis.

For non-take-all units, the normal time in sample is 12 months, although in certain strata with high sampling fractions a slower rotation may be necessary in order to ensure that, once a unit is rotated out, it stays out of sample for at least twelve months. Special measures are taken to ensure that even if a unit changes stratum, it is not brought back into sample until it has been out at least twelve months. In the case of multi-establishment enterprises, when one or more of the establishments belonging to the same enterprise are selected, one establishment is randomly chosen to control rotation for the entire enterprise. No establishments within the enterprise can rotate into or out of sample until the controlling establishment rotates out. Once this establishment does rotate out, the whole enterprise and all of its establishments must remain out of sample for twelve months. The purpose of this procedure is to maintain the stability of reporting arrangements for the enterprise during the time that it is in sample.

It is recognized that some of these procedures do introduce a bias into the survey sample. In the context in which the data are used, however, these biases are felt to be acceptable in relation to the reduction in response burden achieved.

3. THE IMPACT OF PIPES

The Project to Improve Provincial Economic Statistics (PIPES)) is a direct result of an agreement between the Government of Canada and three provincial governments to harmonize their federal and provincial sales taxes. Under the accord, which took effect on April 1, 1997, the federal government collects the Harmonized Sales Tax (HST), and the pooled revenues are allocated using the Provincial Economic Accounts of Statistics Canada. As part of the agreement, Statistics Canada received permanent funding to make the improvements needed to support the HST allocation formula. It should be noted that because these provinces trade with other provinces, it is necessary to improve the quality of economic data for all provinces, and not just for the three provinces who are part of the HST agreement.

The major focus of these improvements has been the development of a new integrated program, known as the Unified Enterprise Survey (UES), that is eventually replacing the existing program of annual business surveys. This unified approach is designed to achieve data quality improvements in four areas: (i) better consistency of the methods used across surveys, (ii) better coherence of the data collected from different levels of the business, (iii) better coverage of industries, and (iv) better depth of information, in the sense of more content detail and estimates for more detailed domains. A more complete description of PIPES and the UES is given in Royce (1998).

In designing the UES, a number of basic principles were adopted. Of interest in the current context were the adoption of the Business Register as the sole source of the frame for the UES, the use of tax data to minimize response burden, and the adoption of an "enterprise-centric" approach, whereby all of the data for an enterprise and its associated establishments would be analyzed together. Many of these ideas were not new, but unlike the Business Survey Redesign Project, Statistics Canada now has the long-term funding to implement these principles.

In the case of the UES sample, the frame is drawn once per year from the Business Register. It is used for selecting a single sample covering all the industries in the UES. The enterprise and the establishment are the two statistical units for which data are compiled. This standardization of statistical units, frames and sample designs has largely eliminated the problems of overlapping surveys directed to the same unit, at least for those units in the UES. The manufacturing sector, which previously maintained its own frame, switched to using the UES frame in 1998, although it is maintaining its old sample design until the survey undergoes a more complete redesign for the reference year 2000. The mining sector also used the UES frame for the first time in 1999, although the survey is still conducted by Natural Resources Canada.

A key feature of the UES sample design is the use of adaptive cluster sampling, sometimes called network sampling (Simard and Hidiroglou 1999). In the UES, it was deemed essential to survey enterprises and their establishments as a whole. However enterprises which have establishments in more than one province or industry cannot be stratified by these variables. One solution could have been to make all such "complex" enterprises take-all units; however analysis of the frame showed that there were many relatively small complex enterprises for whom this approach would be an unwarranted burden, as they would be in the sample every year. The solution was to define the initial sampling units as clusters of establishments within an enterprise and within the same province and industry. These clusters could then be stratified by province and industry. If a cluster belonging to an enterprise was selected, then the enterprise and all other establishments in the enterprise were also selected in the sample. This is an example of positive coordination of samples across different levels, as described in Ohlsson (1998).

Another of the major achievements of PIPES has been the integration of the sampling frame previously used by the Quarterly Financial Survey (QFS) into the Business Register. The reconciliation and integration of these two frames took a considerable amount of work, but the end result has been a much-improved Business Register. The change of frame also provided the opportunity for a complete redesign of the Quarterly Financial Survey, using the Business Register as its frame. As part of the redesign, a portion of the new sample was coordinated with the sample for the Unified Enterprise Survey. All complex enterprises (those with establishments in more than one province and/or more than one industry) above a certain size threshold were included with certainty in both the QFS and UES samples. In the case of the UES, all of the establishments belonging to these pre-specified take-all complex enterprises were also selected. In this way, the enterprise-level financial data from the QFS can be combined with the establishment-level data from the UES to carry out so-called "coherence analysis." While there is no such coordination of QFS and UES samples for complex enterprises below the size threshold, these smaller enterprises account for only a very small portion of the economic activity of all complex enterprises.

As well as the Quarterly Financial Survey, the sample for the Capital Expenditures Survey (CAPEX) was redesigned to be better coordinated with the UES. Although CAPEX already used the BR, its sampling frame was modified so that it used exactly the same frame extracted for the UES. In this case, however, positive sample coordination is required not only for the large complex units, but for as much of the sample as possible, so that the results from CAPEX can be combined with those of the UES. Both surveys (UES and CAPEX) use collocated sampling, although the UES uses the cluster as its sampling unit while CAPEX uses the establishment. In order to achieve a high overlap, the random numbers used to calculate the selection numbers were generated in two stages. The first nine digits are generated by running the enterprise identification number through a hashing algorithm. The second nine digits are generated in a similar fashion using the cluster (for UES) or the establishment (for CAPEX). This resulted in a very high degree of overlap, with over 99% of the CAPEX sample also falling into the UES.

Finally, the 1997 UES sample design for the simple businesses was coordinated with the design of a sample of tax data in a two-phase approach. The Business Register was the frame for both samples. Tax data were obtained for the first phase, and a smaller sample was chosen for which survey data were also collected. Data from the first phase tax sample were used as auxiliary data at the estimation stage to improve the estimates from the survey data. However, because of the timetable for tax data processing, the first phase sample was selected after the second phase sample. Again, this was accomplished by the use of random numbers to ensure that the two samples overlapped. For 1998 and future years, the Canada Customs and Revenue Agency has introduced a new processing system for corporate tax returns. Under this system, Statistics Canada receives, in electronic form, detailed financial data for the entire universe of corporate tax filers.

This movement to the Business Register, the standardization of sample designs for industries in the UES, and the coordination of the UES sample with related surveys such as the QFS and CAPEX has moved Statistics Canada far along the road to taking a more systematic approach to the coordination of samples for its business surveys. In the next section I look forward and discuss some issues that need to be addressed in the coming years.

4. FUTURE ISSUES IN THE USE OF COORDINATED SAMPLING

By reference year 2000, the UES will cover the majority of annual business surveys, including mining, construction, manufacturing, distributive trades, and large portions of the services sector. The use of common concepts, frame and sample designs across industries goes a great distance towards the objective of a more coordinated and coherent approach to business surveys. Integration of the samples of the UES, the QFS, CAPEX and tax data has led to better integration of data on Canadian businesses. However there are still a number of developments needed before Statistics Canada can fully achieve the desired level of coordination of the samples for all of its business surveys.

4.1 Types of Sample Coordination to be Achieved

As mentioned in the introduction, sample coordination can have many different objectives. At this point in our development, there appear to be three main types of sample coordination that would be desirable to achieve within the next few years. They are listed in perceived increasing order of difficulty.

4.1.1 Rotation of Annual Samples

The annual samples selected for the three editions of the UES to date (1997 through 1999) have not incorporated sample rotation. Because the sample design has evolved as additional surveys come into the UES and as the Business Register has improved, the approach so far has been to select independent samples each year. However, the UES questionnaire is a relatively heavy burden on respondents due to the detailed data needed for PIPES. It is hoped to introduce annual rotation of the sample beginning in reference year 2001 or 2002, both to control this burden and to provide better estimates of year-to-year changes. The collocated sampling method using PRNs will form the basis for this rotation. Introduction of rotation into the UES will also imply the rotation of the sample for the Capital Expenditures Survey, whose sample is closely integrated with that of the UES itself.

4.1.2 Positive Coordination of Annual and Sub-Annual Surveys

The UES is limited to annual surveys, but in some cases there is a wish to have positive coordination of the samples of the annual survey and the sub-annual survey for the same industry. For example, the annual sum of the estimates of sales from the Monthly Retail Trade Survey has been used in an important transfer payment program known as equalization. Users naturally compare the estimates derived from the monthly survey to similar estimates from the annual survey. Due to the sensitive use of these data, deviations between the monthly and annual survey estimates are undesirable, so it would be beneficial to maximize the overlap between the two surveys. Such coordination had been achieved under the Period 13 approach described in section 2.4, but now that the Annual Retail Trade Survey is being conducted under the UES framework this is no longer the case. And as mentioned in Section 2.1, coordination of the current monthly survey sample with other surveys is not easy to achieve.

The MWRTS sample will be redesigned in the next three years, likely using GSAM, and coordination of its sample with the corresponding UES annual sample may be one aspect of the new design. The methodology may become quite complex, as it will have to take account of desired rotation patterns in both the monthly and the annual survey. One solution would be to rotate the new MWRTS only on an annual basis rather than on a monthly basis as is done now, but this might introduce unacceptable breaks in the monthly series. Another solution could be to let the rotation pattern for the monthly survey drive the rotation pattern for the annual survey. More study is needed before specific methods can be implemented.

4.1.3 Negative Coordination of Surveys Targeting the Same Population

As well as its regular program of annual and sub-annual surveys, Statistics Canada conducts a large number of ad hoc surveys, on topics ranging from Year 2000 preparedness to the use of technology in specific industries. Such surveys often target the same populations as the regular program of surveys. For surveys such as these, there are often few advantages to the positive coordination of samples, as the content of the ad hoc survey is often unrelated to that of the regular survey. For such cases, it would be desirable to implement some form of negative coordination of samples in order to minimize the burden on respondents. Doing so will be challenging, because by their nature the sample design requirements for such ad hoc surveys are difficult to anticipate in advance. In addition, not all such surveys use the Business Register as the frame. For those that do, however, it may be possible to use methods such as those described for the crop and livestock surveys to minimize the overlap.

Negative coordination of samples may also be desirable for some of our ongoing surveys that cover several industries. One such survey is the Workplace and Employees Survey (WES), an annual survey that collects employment-related data from both employers and employees in an integrated fashion. To date there has been no study of the feasibility of negative coordination between the sample for WES and those of other surveys. Such coordination would be complicated since WES is a longitudinal survey.

4.2 The Infrastructure for Coordinated Sampling

While much progress has been achieved during the past three years, further progress on coordinated sampling will require developments of some key elements of an infrastructure. The three most important elements of this infrastructure are the Business Register, GSAM, and, perhaps most important of all, a management process to determine exactly what kind of sample coordination is desired. Within this framework, we can then proceed to use tools such as those described in Ohlsson (1995) to achieve our goals.

4.2.1 Standardization on the Business Register

Thanks to PIPES, the Business Register is now firmly established as the frame for most major business surveys at Statistics Canada. While there will always remain a need for specialized frames for some industries, the use of the BR for most surveys now makes it possible to conceive of a more comprehensive strategy for coordinating samples. As a first step in this direction, the BR has already developed a Contact Archival and Tracking System (CATS) that records all contacts made with a particular business from all surveys using the BR as the frame. This allows us to monitor the response burden on any given business. CATS will be an important tool in monitoring how well our future attempts to coordinate samples are working. A second step should be the assignment of Permanent Random Numbers to all units on the Business Register, rather than the current situation where each program generates its own set of PRNs. The assignment of PRNs should take into account the need to coordinate samples across the various levels of the BR. Finally, further hookup of additional surveys to the BR, which generally is accompanied by a redesign of the sample, will allow us to consider the possibility of sample coordination at the time the sample is redesigned.

4.2.2 Standardization and Extension of GSAM

During the past few years, GSAM has been adopted as the sampling system for most new or redesigned business surveys at Statistics Canada. As mentioned in Section 2.2, GSAM permits a variety of sample design and sample maintenance activities, including sample rotation, sample re-balancing, and re-stratification while maximizing sample overlap. Due to the complexity and size of the UES, GSAM itself was not adopted for the UES, but the UES sampling system does incorporate the same basic collocated sampling methods as GSAM. Efforts are being made to ensure that the methods available in GSAM are incorporated into the UES when it does move to a rotating design.

However, considerably more work would be needed to permit GSAM to coordinate different surveys at the same time. While GSAM could be used to handle methods such as those described for the crop and livestock surveys in section 2.5, it is not designed to do so automatically. During the coming year we plan to study systems in place in other statistical agencies in order to determine what modifications would be necessary for GSAM to handle multiple surveys.

4.2.3 Deciding What We Want to Do

As should be evident, there can be many competing objectives when it comes to sample coordination. In some cases we want to maximize the overlap among samples, either to make the results more consistent or to reduce collection and processing costs. In other cases we wish to minimize the overlap in order to spread out the respondent burden. Where only one survey, or a closely related set of surveys, is involved, the approach is relatively straightforward. However the picture is not so clear when it comes to coordinating samples for different surveys that are managed by different parts of the agency. Before putting any kind of coordinated sampling strategy in place, a management process needs to be established to bring these conflicting demands together and to make choices about how samples should be coordinated. The establishment of a senior-level Task Force on response burden in late 1999 has led to some preliminary discussions on this subject. As well, a small working group has been established to look at the more technical issues of coordination of samples from different surveys to see what is feasible. However there needs to be much more work to bring the various objectives together and to make decisions before any kind of comprehensive sample coordination system can be designed. Arriving at such decisions may in fact be the biggest obstacle to implementing a more comprehensive approach.

5. SOME RELATED RESPONSE BURDEN INITIATIVES

As has been noted, one of the motivations for sample coordination is to reduce the burden on respondents, or at least distribute it more evenly. The arrival of PIPES has meant an increase in this response burden. At the same time, there is steadily increasing pressure for Statistics Canada to reduce the burden on respondents. While the agency has made great strides in this area over the past twenty years, it has recently been studying the possibility of taking additional steps towards the reduction of response burden.

One of these steps is to completely eliminate the burden on very small respondents. This could been done through establishing a set of “exclusion thresholds” based on size, below which a business is excluded from surveying. The exclusion thresholds would vary by province and industry. Adjustments would be made for the excluded units, using whatever auxiliary information is available, such as tax data. Using these thresholds, nearly half of the businesses on the Business Register could be excluded from surveying, even though these businesses account for less than 2% of gross business income.

A second measure is to require that medium-sized businesses, i.e., those above the exclusion thresholds but still relatively small, should be sampled at a rate of no more than 2%. The rule applies at the level of the survey, not for each sampling stratum, since in some cases strata are too small to permit such low sampling fractions. While somewhat arbitrary, this rule would limit the chance that a unit will fall into more than one survey at any given time.

Finally, medium-sized businesses would be rotated out of survey samples after at most two years in sample. For a monthly survey, this would mean that one-twenty-fourth of such businesses would rotate each month, giving a good basis for measuring trends. For annual surveys, half of these businesses would rotate each year. Of course, the overall rotation rate for the survey would be considerably lower.

At the present time, these measures are general principles rather than fully implemented procedures. In many cases these measures cannot be implemented until the survey is redesigned, and even then there will be a need for exceptions. A senior-level committee has been set up to examine the feasibility of these measures for each of our survey programs. However just the process of developing and discussing these principles has served to raise the awareness of survey managers of the importance of response burden and of the need to consider it in the design of the sample.

6. CONCLUSION

This paper has provided a description of the use of coordinated sampling at Statistics Canada, as well as some indication of the issues that must be addressed if the agency is to take a more comprehensive approach to the subject in the next few years. To a large extent, efforts at sample coordination have been rather isolated, limited to one or a small group of related surveys. However the advent of PIPES and the UES has given us the opportunity to better coordinate and standardize our methods. The move to a common Business Register, the development of GSAM, and the techniques of sample coordination using Permanent Random Numbers have now given us many of the tools we need to move forward in a more integrated fashion. The biggest challenge will undoubtedly be in deciding what to do, because there are many conflicting objectives when it comes to sample coordination. However the experience of the UES has given us valuable lessons on how to work together to solve common problems. The fact that we rely so heavily on our respondents also makes the development of a more comprehensive approach a necessity.

7. REFERENCES

Bérard, H., Brodeur, M. and St-Pierre, M. (1999) “The Retail Commodity Survey.” *Proceedings of the Survey Methods Section, Statistical Society of Canada,* pp. 141-146.

Brewer, K. R. W., Early, L.J. and Joyce, S.F. (1972) “Selecting Several Samples from a Single Population.” *Australian Journal of Statistics,* Vol. 14, pp. 231-239.

Brodeur, M., Jocelyn, W. and Trépanier, J. (1997) "A Solution to the Design and Implementation of a Fast-track Survey: Two-Phase Sampling." *Proceedings of Symposium 97: New Directions in Surveys and Censuses,* Statistics Canada, pp. 31-35.

Colledge, M.J. (1987) "The Business Survey Redesign Project: Implementation of a New Strategy at Statistics Canada." *Proceedings of the Third Annual Research Conference.* Bureau of the Census, pp. 550-576.

Denis, J., Hamel, N., Tardif, C., Grondin, C. and Pelletier, É. (1999) *1996 Redesign of Agricultural Commodity Surveys: - Panel of crop surveys;-Semi-annual livestock survey;-Semi-annual survey of the Atlantic provinces.* Methodology Branch Working Paper BSMD99-002-E/F, Statistics Canada.

Grondin, C. (1997) "Contrôle du chevauchement entre l'enquête sur les cultures et l'enquête sur le bétail." Internal report, Statistics Canada.

Hidiroglou, M.A., Choudhry, H. and Lavallée, P. (1991) "A Sampling and Estimation Methodology for Sub-Annual Business Surveys." *Survey Methodology,* Vol. 17, No. 2, pp. 195-210.

Hidiroglou, M.A., and Srinath, K.P. (1993) "Problems Associated with Designing Subannual Business Surveys." *Journal of Business & Economic Statistics,* Vol. 11, No. 4, pp. 397-405.

Hidiroglou, M.A., Carpenter, R. and Estevao, V. (1996) "Methods Used in Sampling and Estimation at Statistics Canada." *Proceedings of the Section on Survey Research Methods, American Statistical Association,* pp. 73-80.

Kish, L. and Scott, A. (1971) "Retaining Units after Changing Strata and Probabilities." *Journal of the American Statistical Association,* 66, pp. 461-470.

Ohlsson, E. (1995) "Coordination of Samples Using Permanent Random Numbers." *Business Survey Methods,* John Wiley and Sons, Inc. pp. 153-169.

Ohlsson, E. (1998) "Spreading Respondent Burden and Updating Samples by Using Permanent Random Numbers." *Methodological Issues in Official Statistics,* Statistics Sweden.

Outrata, E. and Chinnappa, N. (1989) "General Survey Functions Design at Statistics Canada." *Bulletin of the ISI, Proceedings of the 47th Session,* pp. 219-238.

Royce, D.A. (1998) "Project to Improve Provincial Economic Statistics." *Proceedings of the Joint IASS/IAOS Conference, Statistics for Economic and Social Development.* Instituto Nacional de Estadistica, Geografia e Informatica.

Simard, M. and Hidiroglou, M.A. (1999) "Estimation for Annual Business Surveys Based On Two-Phase Network Sampling." *Proceedings of the Survey Methods Section, Statistical Society of Canada,* pp. 11-19.

Statistics Canada (1996) "Generalized Sampling System: User Guide." Internal report, Statistics Canada.

Sunter, A. (1986) "Implicit longitudinal sampling from administrative files: A useful technique." *Journal of Official Statistics,* 2, pp. 161-168.

COORDINATION OF PPS SAMPLES OVER TIME

Esbjörn Ohlsson, Stockholm University
Mathematical Statistics, Stockholm University, S–106 91 Stockholm, Sweden
esbj@matematik.su.se

ABSTRACT

Probability proportional to size (PPS) sampling finds application in business surveys both for the first stage of a multi-stage design, and for direct sampling of businesses from a list frame. In both cases there is a need to update samples from time to time, while retaining as many units as possible from the old sample. In the second type of application there is also a need for negative coordination of surveys to get an even distribution of response burden. Further requirements on the sampling procedures are simplicity in application and in estimation of variance. We present various permanent random number techniques that meet these requirements and compare them to a few other methods, and present a simulation study on expected overlap.

Key Words: Positive Coordination, Negative Coordination, Overlap Control, Permanent Random Numbers

1. THE PROBLEM

Applications of PPS sampling in business surveys can be split into two main categories: (a) sampling from area frames with a multi-stage design and (b) sampling ultimate units directly from a list frame. The classical situation for (a) is a multi-stage sample where primary sampling units are drawn with probability proportional to some measure of the unit's size (PPS). Typically, due to extensive stratification, just a few units are selected in each stratum.

As an example we consider a master sample of stretches of road, maintained by the Swedish National Road Administration. Even though the first stage sample size is 84, extensive stratification results in stratum sample sizes n=1, throughout. In each sampled unit, investments are made, e.g. on equipment to measure traffic flow, and this, of course, is a strong argument for retaining the sample over the years. Another argument is that the main interest in samples from this frame is in estimates of change over time, for which retaining the same units improves precision.

The sizes of the units are based on a rotating census that gives estimates of traffic mileage per year for one third of the units each year. Since these sizes changes substantially over the years, there is a regular need to update the samples, if not every third year so at least each decade. Else, there would finally be a great loss in efficiency in the estimates from surveys using the master sample. Furthermore, at irregular intervals there are changes in the classification of roads that underlies the stratification: e.g., roads may change from county to national or even European highway status. Also, there are some new roads (births) and roads that are no longer in the population (deaths).

We conclude that in this survey, as in most repetitive surveys, there is a need for updating the sample to account for new sizes, stratum/classification changes plus births and deaths, while retaining as many units as possible from the old sample.

In this example, the within strata sample sizes were n=1, as is the case with several major surveys such as the US Consumer Expenditure Survey and the US Current Population Survey. This might explain why this case has received considerable attention in the literature see e.g., Keyfitz (1951), Kish and Scott (1971) and Causey, Cox and Ernst (1985). Several newer references can be found in the overview by Ernst (1999). In the present paper we will put a special emphasis on the case n=1, but we will also consider PPS samples of other.

The second type of application of PPS sampling is the case where we have a list frame with ultimate sampling units, often in the form of a business register. The most common design in this case seems to be a stratified simple random sample (STSI). Sample coordination with this kind of design was discussed extensively at the 1993 ICES meeting, see Ohlsson (1995) and Srinath and Carpenter (1995). It is also a topic at the ICES II meeting, see McKenzie and Gross (1999) and Royce (1999). In these cases, there is both stratification by industry and by size. The PPS alternative means that the stratification by industry is kept while stratification by size is replaced by PPS sampling. One advantage of this kind of design is that the size measure is more extensively exploited.

One example of a PPS design of this second kind is the sample of outlets for the Swedish CPI. In an investigation of the sample of retail traders for the Swedish Consumer Price Index, Ålenius (1990) considered the sample of n=41 department stores out of N=259 units in that industrial stratum, with prices of two different commodities as target variables. With a standard size stratification, using four strata, STSI gave twice the variance of PPS sampling. Even when the STSI design was refined to use 41 strata and a ratio estimator was used (treating size as an auxiliary variable), STSI gave around 25 percent larger variance than PPS. For the NASS Crops Survey, Bailey & Kott (1997) found PPS sampling to be more efficient than STSI for most crops.

The conclusion is that there are business surveys for which a PPS design is preferable to STSI from an efficiency point of view. Furthermore, if the PPS sampling procedure is simple to implement, a PPS design is simpler to administrate, with much fewer strata to construct, allocate and maintain. Even if we do not claim PPS to be generally preferable to STSI, it should be considered a strong candidate for business survey designs. A necessary condition for this is that the PPS procedure involved is simple to use in practice. Simple and efficient PPS procedures are the main topic of this paper.

2. SPECIFICATION FOR PPS COORDINATION

There are an immense number of PPS procedures available in the literature, see Brewer & Hanif (1983). Most of these are not proper for sample coordination, though. In fact, we believe that lack of simple, efficient PPS procedures that can produce coordinated samples may be a reason for the extensive use of STSI instead of PPS sampling. We now specify in more detail our requirements on a simple and efficient PPS sampling procedure with capability of handling sample coordination.

The (stratum) population is $U=1,2,...,N$. In the frame (which is supposed to be a list even in case a) there is a non-negative auxiliary variable $p_1, p_2, ..., p_N$. In applications, p_i is usually a measure of the size of unit i. We assume that the p_i:s have been normed so that $\sum p_i = 1$, within each stratum.

Särndal, Swensson and Wretman (1992, p. 90) give a list on desirable properties of a PPS procedure. Applied to the situation with two samples that are to be coordinated, their first two properties give us the following three requirements:

(i) Relative simplicity in application.
(ii) For the first sample $\pi_i = \Pr(i \in s) = np_i$, $i \in U$.
(iii) For the second sample $\pi'_i = \Pr(i \in s') = n'p'_i$, $i \in U'$.

Here $\pi_i = \Pr(i \in s)$ denotes actual inclusion probability of unit i in the first sample s. All quantities relating to the second sample s' will be equipped with a prime, as in p'_i.

Särndal et al. add three conditions that enable variance estimation with the Sen-Yates-Grundy estimator. In a later article, Särndal (1996) argues for the use of procedures that allow simple, single-sum variance estimation. We agree with this point of view and get

(iv) Availability of a variance estimator, preferably expressed as a single sum. (Not relevant for $n=1$.)

Finally, we add three conditions that are particular for the problem of overlap control. Note that the expected number of units in common to two samples is

$$\sum_i \Pr(i \in s, i \in s') \qquad (1)$$

(v) Possibility of positive sample coordination of two or more samples with different size measures p, different strata and different n, preferably with maximization of the expected sample overlap in (1).

(vi) Possibility of negative sample coordination, of two or more samples with different size measures p, different strata and different n, preferably with minimization of the expected sample overlap in (1).

(vii) On each occasion, all strata are sampled independent of each other. For the second sample this means that $\Pr(i \in s', j \in s') = \Pr(i \in s')\Pr(j \in s')$ whenever i and j are in *different* strata.

As noted by Ernst (1999), condition (vii) is not satisfied by most overlap procedures that allow for different stratification. This condition is important for obtaining unbiased variance estimates and, above all, for the possibility to apply the sample overlap procedure repeatedly.

Särndal et al. (1992, p. 90) remark that it is not easy to devise a (fixed-size) procedure having the properties desirable for PPS sampling, even at a single occasion. It is thus futile to hope for a procedure for sampling on two occasions, with overlap control, that fulfils all our requirements (i)-(vii). Instead, we have to look for procedures that are reasonably good at (i)-(vii). We start out with a procedure that is not fixed-size, i.e. it gives a random sample size.

3. POISSON SAMPLING AND THE IDEA OF PRN

In Poisson sampling, each unit is given an independent, uniformly distributed random number X_i on the interval (0,1). Unit i is included in the sample s if $X_i \le np_i$. Poisson sampling can be used for sample coordination by saving the X_i as *permanent random numbers* (PRN). This idea is due to Brewer, Early and Joyce (1972) and means that when a second sample is to be drawn, we use the same random numbers as in the first, but we update the sizes p, the stratification, and the sample size n. Note that the quantities p_i' and n' relates to the stratum where i is located in the second design, which may or may not be the same as in the first design. (For simplicity in notation, we refrain from using stratum sub-indexes.)

A virtue of Poisson sampling is that it is very simple to apply, i.e. requirement (i) met. Further, it is readily seen that this procedure is strictly PPS so that (ii) and (iii) are fulfilled. A single-sum variance estimator (iv) is available, see Brewer and Hanif (1983, p. 83). The probability of including unit i in two samples drawn with PRN Poisson sampling is obviously

$$\Pr(i \in s, i \in s') = \min(np_i, n'p_i') \qquad (2)$$

This is of course the largest possible probability, yielding a strict maximum i (1). Negative coordination can be achieved by shifting the PRN an amount c to the right before the selection of the second sample, giving new random numbers $X_i^* = X_i + c$. If m samples are to be negatively coordinated, the choice $c = 1/m$ should give a small sample overlap. In particular, if the target inclusion probabilities np_i are less than $1/m$ for all units i in all m designs, the expected overlap is 0. In the case of m=2, an alternative is to use antithetic random numbers, $X_i^* = 1 - X_i$ which gives minimum expected sample overlap for any target probabilities.

We conclude that both (v) and (vi) are satisfied. Finally, all units are sampled independently and in particular (vii) is met.

It may appear as if we have found the optimal procedure for PPS sample coordination. However, Poisson sampling has the drawback of giving a random sample size (with expected value n). This has two implications: The first is that there is a risk for n=0 in some stratum. Since the random sample size is approximately Poisson distributed, the probability of this to happen when we have H strata, all with sample size n, is $1-\left(1-e^{-n}\right)^H$. In order for this quantity to be negligible we must avoid small n, where the magnitude of "small" of course depends on H. The conclusion is that Poisson sampling can not be used in all situations, and in particular not in those with n=1 or 2. Even when the probability of some zero sample size is negligible, the randomness in sample size may seriously disturb the intended sample allocation over strata, with a loss in efficiency of the estimates.

The second, less serious, drawback of the random sample size is that it should be clear that inference should be made conditional on the sample size actually obtained. Conditional on the actual sample size, the probabilities of inclusion are no longer exactly PPS and they are in fact very hard to compute, see Aires (1999).

4. FIXED SIZE PROCEDURES

We now look at fixed-size alternatives to Poisson sampling, starting with the case *n*=1. In this section we restrict the attention to PRN procedures. Collocated sampling, by Brewer et al (1972), cannot be used to coordinate samples with different stratification, and will therefore not be considered here.

4.1 The case n=1

A sampling design is a probability distribution on all possible sets of samples. With this definition, PPS sampling with fixed size *n*=1 is unique. The traditional sampling procedure for realizing a PPS size one sample uses just one random number. For the application of PRN sample coordination, we need a procedure that uses individual random numbers for all units. Such a procedure is *Exponential sampling*, presented in Ohlsson (1996). Starting with a set of PRN, $\{X_i; i=1,2,...N\}$, we compute the transformed random numbers $\xi_i = -\log(1-X_i)/p_i$, which are exponentially distributed with mean $1/p_i$. The unit with the smallest ξ_i is selected for the sample. By a well-known result from probability theory, the probability of selecting unit *i* is p_i, as required.

Coordination of samples is achieved by using PRN as described for Poisson sampling above. Exponential sampling does not reach the optimal expected overlap in (v) and (vi), but is not too far away, see the numerical examples below. A formula for expected overlap for positive coordination was given in Ohlsson (1996).

Since this procedure is very simple to implement, (i) is fulfilled along with (ii)-(iii); (iv) is not relevant since *n*=1. We just mentioned that (v)-(vi) are fulfilled. Finally, it is not hard to see that (vii) is satisfied. We conclude that Exponential sampling is a strong candidate for coordinated PPS sampling when *n*=1.

4.2 The case n>1

Unlike the *n*=1 case, PPS sampling with *n*>1 can be done in several different ways. Most of these can not be used in connection with PRN, since this requires procedures that use individual random numbers for the units. A natural idea is to extend Exponential sampling to *n*>1, by selecting the units with the *n* smallest transformed random numbers. Unfortunately, this yields so called successive sampling (Cochran, 1977, Section 9A.8) which is not strictly PPS. The actual inclusion probabilities may be quite far from the target values. Cochran (1977) presents several techniques for handling this problem, of which we consider Brewer's method for *n*=2. In our context, this method can be applied by drawing the first unit as in Exponential sampling, but with transformed random numbers

$$\xi_i = -\frac{\log(1-X_i)}{p_i(1-p_i)/(1-2p_i)} \qquad (3)$$

After removing the unit drawn in the first round, a second unit is drawn with the transformed random numbers of Exponential sampling $\xi_i = -\log(1-X_i)/p_i$. Cochran (1977, Section 9A.8) shows that these two steps yield a size *n*=2 sample with the required inclusion probabilities, i.e. meets (ii) and (iii). The procedure is relatively simple (i), and has the same properties as Exponential sampling as regards (v)-(vii). It also allows unbiased variance estimation with the Sen-Yates-Grundy estimator. This is not a single-sum estimator, though, so (iv) is not completely met.

The extension of Brewer's method to *n*>2 by Sampford (1967) is rather complicated and will not be considered here.

Ohlsson (1990 and 1998) gave an alteration of Poisson sampling called Sequential Poisson sampling (SPS). This procedure uses the transformed random numbers $\xi_i = X_i / p_i$ and selects the *n* units with the smallest such numbers.

The sample will be close to a Poisson sample, since the latter selects the units with $\xi_i \le n$. Not surprising, the properties are approximately the same as those of Poisson sampling, but the size is fixed. The procedure is very simple (i), admits simple-sum variance estimation (iv) and yields independent strata (vii). It is approximately PPS, in the meaning of (ii) and (iii), a fact which is motivated by asymptotics and simulation in Ohlsson (1998). Even though sample coordination is not optimal, in terms of maximum and minimum expected overlap in (v) and (vi), respectively, we can expect the overlap not to be too far from the optimum of Poisson sampling. The case of positive coordination is investigated in a simulation study in the next chapter.

Rosén (1997) gave an alteration of SPS, called *Pareto sampling* (PAS), with transformed random numbers of "odds ratio type"

$$\xi_i = \frac{X_i/(1-X_i)}{np_i/(1-np_i)} \qquad (4)$$

The properties are similar to those of SPS, but PAS is somewhat closer to the target inclusion probabilities in (ii) and (iii). The closeness to the optimum in (v) is investigated in the simulation study below.

5. COMPARISON WITH OTHER PROCEDURES

5.1 The case n=1

In the literature, there are several (non-PRN) procedures for positive coordination (maximizing overlap) of two PPS size n=1 samples. The pioneering procedure by Keyfitz (1952) assumes the same stratification for both samples. Conditional on the first sample, Keyfitz' method focuses on the second sample. Suppose unit i was selected in the first sample. Keyfitz' method retains this unit in sample if it is increasing, i. e. $p_i \le p'_i$. Else, the unit is retained with probability p'_i / p_i. If the unit is rejected, one of the increasing units are selected with probability proportional to the increments $(p'_i - p_i)$. It is easily verified that this simple procedure is strictly PPS and is optimal in terms of maximizing expected sample overlap, when we have the same stratification for both samples.

Keyfitz' procedure can be extended to the case with (somewhat) changing strata. We will describe this procedure for a single new stratum. First identify an old stratum which will be considered as the predecessor of our new stratum. Units coming from other old strata, *immigrants*, are treated as births, i.e., they are assigned the value $p_i = 0$. Any first sample selection among immigrants is ignored. Then the original Keyfitz algorithm is applied to the new stratum, but the first two steps are only applied to an eventual initial selection that is not an immigrant.

Kish and Scott (1971) note that this procedure can be far from optimal in terms of expected overlap unless we have very small differences in the stratification of the two samples. They provide three methods (beside the extended Keyfitz procedure) for the case with arbitrary stratification of the two samples. We shall consider only Method II, which is claimed by Kish and Scott to give the largest overlap of the three, without being very complicated. The procedure is an elaborated extension of Keyfitz' method. For a description, we refer to the original article.

The procedure has the disadvantage of distorting the independence between the strata, i.e., it does not fulfil requirement (vii). As already noted, this implies that the procedure can not be applied repeatedly to the same survey. Like Keyfitz' procedure, Kish and Scott only concerns the second sample, so that (ii) is trivial. In their section 6.2, Kish and Scott (1971) prove that the procedure fulfils (iii). The proof relies on the independence of the initial strata, though. A consequence of the dependence of the new strata is therefore that the procedure is not strictly valid for repeated use, i.e., (iii) is only valid the first time the method is applied. The expected overlap is quite close to optimum, see Section 6.1.

5.2 The case n>1

Causey, Cox and Ernst (1985) suggested a procedure which maximizes (or minimizes) the expected overlap, subject to the constraints of having the required target probabilities in both samples, i.e. conditions (ii) and (iii). The problem is solved by linear programming methods.

By design, this procedure fulfils our requirement (ii)-(iii) and is optimal in terms of our choice of (v) or (vi). Ernst and Ikeda (1995) note two difficulties with the procedure which can make it unusable in practice. One is that (vii) is not fulfilled, with the same difficulty as for Kish and Scott to apply the procedure repeatedly for the same survey, especially if n>1. The second difficulty is that the transportation problem may be too large to solve in practice. In any case, (i) is not fulfilled.

Ernst (1999) reviews several alterations of the Causey, Cox and Ernst procedure, non of which fulfills all our requirements.

Sunter (1989) presents an interesting procedure that is applicable for maximizing overlap of two samples with any sample size n. It is a generalization of Keyfitz' procedure. Like the latter, Sunter's procedure was not primarily designed for handling stratum changes and can be expected to give an overlap far from maximum when we have large stratum differences between the two samples.

6. NUMERICAL EXAMPLES

Below we report results from two numerical studies, one for n=1 and one for n=2 to 4.

6.1 The case n=1

The first study concerns expected overlap of different procedures for maximizing overlap in the case n=1. It is based on the so called MU284 population of 284 Swedish municipalities presented in Appendix B of Särndal et al. (1992). See http://statlib@lib.stat.cmu.edu/datasets/mu284.

Draw probabilities are either equal or proportional to the number of inhabitants, for the first sample we use figures from 1975 (P75), and for the second sample those of 1985 (P85). First sample strata are either defined by the regional REG variable, giving 8 strata of sizes between 15 and 56, or by the CL variable, with 50 small strata with sizes between 5 and 8.

The expected overlap was computed exactly (no simulation) for PRN Exponential sampling (EXP), the extended Keyfitz (KEY), the Kish and Scott (K&S), and the Cox-Causey-Ernst (CCE) procedure.

As benchmarks we use (a) the case without any overlap control, with both samples drawn independently (IND) and (b) the non-achievable upper limit (2), added over the strata. For further details on the study, see Ohlsson (1996).

Table 0. MU284 population. Expected overlap in eight regional (REG) strata. Unequal probabilities. 50% of the units move from each stratum to an adjacent stratum.

Stratum	N_h	*IND*	*EXP*	*KEY*	*K&S*	*CCE*	*Limit*
1	36	0.040	0.524	0.431	0.592	0.592	0.718
2	37	0.183	0.772	0.588	0.818	0.874	0.997
3	35	0.084	0.689	0.554	0.793	0.793	0.942
4	35	0.042	0.614	0.488	0.703	0.703	0.845
5	48	0.030	0.578	0.446	0.676	0.676	0.764
6	49	0.105	0.700	0.581	0.814	0.814	0.916
7	21	0.089	0.677	0.609	0.763	0.769	0.948
8	23	0.109	0.645	0.552	0.716	0.741	0.928
Sum	*284*	*0.682*	*5.200*	*4.249*	*5.875*	*5.964*	*7.058*
Percent		*9*	*65*	*53*	*73*	*75*	*88*

For the remaining set-ups, we present no stratum details but just the sum in percent of the number of sampled units.

Table 1. n=1. MU284 population. Expected overlap in percent of total sample size. No changes in strata.

Strata	*Prob*	*IND*	*EXP*	*KEY*	*K&S*	*CCE*	*Limit*
Medium (REG)	Unequal	9	96	97	97	97	97
	Equal	1	100	100	100	100	100
Small (CL)	Unequal	30	97	98	98	98	98

Table 2. n=1. MU284 population. Expected overlap in percent of total sample size. One unit changes stratum.

Strata	*Prob*	*IND*	*EXP*	*KEY*	*K&S*	*CCE*	*Limit*
Medium (REG)	Unequal	9	92	94	94	95	96
	Equal	3	95	97	97	97	99
Small (CL)	Unequal	29	78	80	80	84	87

Table 3. n=1. MU284 population. Expected overlap in percent of total sample size. 50% of units change stratum.

Strata	*Prob*	*IND*	*EXP*	*KEY*	*K&S*	*CCE*	*Limit*
Medium (REG)	Unequal	9	65	53	73	75	88
	Equal	3	66	52	75	75	89
Small (CL)	Unequal	26	64	48	69	72	81

Table 4. n=1. MU284 population. Expected overlap in percent of total sample size. One third of the units move to the next stratum, one sixth to the following.

Strata	*Prob*	*IND*	*EXP*	*KEY*	*K&S*	*CCE*	*Limit*
Medium (REG)	Unequal	7	59	46	70	-	83
	Equal	3	61	42	69	-	88
Small (CL)	Unequal	-	-	-	-	-	-

Note: CCE (Causey-Cox-Ernst procedure) intractable in two strata. Unequal case not treated.

6.2 The case n>1

For n=2, 3, 4 we have conducted a simulation study of expected overlap for sequential Poisson sampling (SEP), and Pareto sampling (PAR). For n=2, we also consider exponential sampling (EXP), in the form using Brewer's recalculated draw probabilities. The benchmarks IND and Limit are as in the preceding section.

We use data from a master frame of the Swedish National Road Administration. The statistical units are stretches of road and the size measure is derived from traffic mileage per year. The units are stratified according to region and type of road, altogether 28 strata with 2523 units. We let 10% of the units change strata between the two sampling occasions. A more detailed report of the study is given in Ohlsson (1999). The data are available at the address http://www.matematik.su.se/~esbj/roads.dat. The simulations were run with 14000 iterations.

Table 5. n=2. Road population. Expected overlap by stratum.

Stratum no.	N_h	*IND*	*EXP*	*SEP*	*PAR*	*Limit*
11	44	0.12	1.77	1.77	1.77	1.97
12	41	0.13	1.66	1.66	1.66	1.78
13	58	0.12	1.80	1.79	1.79	1.93
14	292	0.03	1.75	1.74	1.74	1.93
21	35	0.14	1.78	1.78	1.78	1.94
22	80	0.07	1.73	1.72	1.72	1.91
23	82	0.08	1.72	1.72	1.72	1.91
24	304	0.03	1.73	1.73	1.73	1.91
31	12	0.49	1.77	1.77	1.78	1.93
32	6	0.63	1.61	1.60	1.61	1.77
33	34	0.19	1.56	1.54	1.55	1.79
34	61	0.11	1.77	1.76	1.77	1.96
41	39	0.16	1.72	1.72	1.73	1.95
42	79	0.07	1.70	1.70	1.70	1.88
43	73	0.08	1.69	1.69	1.69	1.91
44	318	0.02	1.73	1.73	1.73	1.94
51	30	0.16	1.70	1.69	1.70	1.95
52	56	0.08	1.70	1.70	1.70	1.89
53	44	0.12	1.66	1.66	1.66	1.91
54	170	0.04	1.67	1.66	1.66	1.92
61	37	0.16	1.73	1.73	1.73	1.90
62	68	0.08	1.75	1.75	1.75	1.96
63	53	0.10	1.66	1.65	1.66	1.92
64	301	0.02	1.65	1.64	1.64	1.86
71	21	0.25	1.66	1.66	1.66	1.86
72	34	0.15	1.65	1.64	1.65	1.94
73	29	0.17	1.63	1.62	1.63	1.93
74	122	0.05	1.71	1.71	1.71	1.92
Sum	*2523*	*3.9*	*47.6*	*47.5*	*47.6*	*53.4*

For n=3 and 4 we only give the sum over strata.

Table 6. Road population. Expected overlap, aggregated over strata. 10% of units change stratum.

Sample size	*IND*	*EXP*	*SEP*	*PAR*	*Limit*	*n*
n=2	3.9	47.6	47.5	47.6	53.4	56
n=3	8.7	-	72.8	72.9	80.1	84
n=4	15.5	-	98.6	98.7	106.7	112

6.3 Inclusion probabilities

The PRN techniques for n>1 are only approximately (asymptotically) PPS. Numerical studies indicate that the approximation is very good in many situations. Ohlsson (1990, 1998) reports a simulation study on CPI data, where SEP is very close to being unbiased. Rosén (1998) studies exact actual inclusion probabilities (AIP) for SEP and PAR in an artificial, but nevertheless interesting, situation, viz. when all units have the same inclusion probability except for one odd unit. Even with sample sizes as small as n=2, the relative error in the AIP for PAR is never larger than 2% and in most cases it is much smaller. SEP is a bit further from the target probability. The AIPs were also computed in the simulation study mentioned in Section 6.2. The results are reported in Ohlsson (1999). Here again PAR is quite close to the unbiased case, with SEP performing a little bit less good.

7. CONCLUSIONS

We first consider the problem of *maximizing* overlap, which is the concern of the numerical studies. Using any of the procedures under consideration gives a great increase in (expected) sample overlap, as opposed to drawing independent samples (IND). For *n*=1, the Kish and Scott procedure is quite close to the optimal (in this respect) Cox, Causey and Ernst procedure. When there are great differences between the strata of the two samples, Keyfitz' procedure is rather far from the maximum expected overlap. Exponential sampling is a bit away from the optimum, but much less so than Keyfitz. Since the K&S and CCE both suffer from the problem of dependent strata, i.e. violate (vii), and since the former does not fulfill (vii) and the latter violates (i), we consider Exponential sampling a good compromise in the search for a procedure that fulfills (i)-(iii) and (v)-(vii) in as far as possible. Of the mentioned procedures, Exponential sampling and CCE are the only ones that can be used for *minimizing* overlap (vi).

Turning to the case *n*>1, we concluded in Section 5.2 that CCE can be unusable in practice and that Sunter's procedure can be expected to give a low sample overlap when there are great differences in stratification between the two samples. This leaves us with the PRN procedures, which are all equal in sample overlap in our simulation studies. Since PAR is a bit closer to the right inclusion probabilities it is generally preferable to SEP. EXP is an alternative for *n*=2, being strictly unbiased, but it suffers from a more complicated variance estimation procedure.

In any case, the various PRN techniques investigated here are simple and efficient for simultaneous negative and positive sample coordination of any number of surveys, with any draw probabilities and any stratification. When *n*>1, they all allow for variance estimation, and all but Exponential have an associated single-sum variance estimator. In summary, the PRN techniques fulfill (i)-(vii) even though they do not strictly maximize/minimize the expected sample overlap.

The overall conclusion is that PRN procedures, n.b. Exponential sampling for *n*=1 and maybe *n*=2, and Pareto sampling for *n>1*, are competitive as procedures for controlling sample overlap.

8. REFERENCES

Aires, N. (1999). "Comparisons Between Conditional Poisson Sampling and Pareto πps Sampling Designs," Contributed paper, Bulletin of the International Statistical Institute, 52nd Session.

Ålenius, M. (1990). "Storleksstratifiering eller pps-urval? En jämförande studie för KPI-data," *F-METOD NR 25, Statistics Sweden*. (In Swedish.)

Bailey, J. T. and Kott, P.S. (1997). "An Application of Multiple List Frame Sampling for Multi-Purpose Surveys," *ASA Proceedings of the Section on Survey Research Methods*.

Brewer, K.R.W., Early, L.J. and Joyce, S.F. (1972). "Selecting several samples from a single population," *Australian Journal of Statistics*, 14, 231-239.

Brewer, K.R.W., and Hanif, M. (1983). *Sampling With Unequal Probabilities*. Springer, New York.

Causey, B. D., Cox, L. H. and Ernst, L. R. (1985), "Application of Transportation Theory to Statistical Problems," *Journal of the American Statistical Association*, 80, 903-909.

Cochran, W. G. (1977). *Sampling Techniques 3d ed*., New York: John Wiley.

Ernst, L. R. (1999), "The Maximization and Minimization of Sample Overlap Problems: A Half Century of Results," Invited Paper, Bulletin of the International Statistical Institute, 52nd Session.

Ernst, L. R. and Ikeda, M. M. (1995), "A Reduced-Size Transportation Algorithm for Maximizing the Overlap Between Surveys," *Survey Methodology*, 21, 147-157.

Keyfitz, N. (1951), “Sampling with Probabilities Proportional to Size,” *Journal of the American Statistical Association*, 46, 105-109.

Kish, L. and Scott, A. (1971), “Retaining Units After Changing Strata and Probabilities,” *Journal of the American Statistical Association*, 66, 461-470.

McKenzie, R. and Gross, B. (1999), Synchronized Sampling at the Australian Bureau of Statistics. In this volume.

Ohlsson, E. (1990), *Sequential Poisson Sampling from a Business Register and its Application to the Swedish Consumer Price Index,* R & D Report 1990:6, Statistics Sweden.

Ohlsson, E. (1995), “Coordination of Samples Using Permanent Random Numbers,” In *Business Survey Methods*, New York: Wiley, 153-169.

Ohlsson, E. (1996), *Methods for PPS Size One Sample Coordination*, Research Report No. 194, Institute of Actuarial Mathematics and Mathematical Statistics, Stockholm University.

Ohlsson, E. (1998), “Sequential Poisson Sampling,” *Journal of Official Statistics*, 14, 149-162.

Ohlsson, E. (1999), *Methods for PPS Size One Sample Coordination*, Research Report No. 210, Institute of Actuarial Mathematics and Mathematical Statistics, Stockholm University.

Rosén, B. (1997b), “On Sampling with Probability Proportional to Size,” *Journal of Statistical Planning and Inference*, 62, 159-191.

Rosén, B. (1998), On Inclusion Probabilities for Order Sampling, R & D Report 1998:2, Statistics Sweden. To appear in Journal of Statistical Planning and Inference?

Royce, D. (1999), “Issues in Co-ordinated Sampling at Statistics Canada,” In this volume.

Sampford, M. R., (1967), “On Sampling Without Replacement with Unequal Probabilities of Selection,” *Biometrika*, 54, 499-513.

Särndal, C. E., (1996), “Efficient Estimators With Simple Variance in Unequal Probability Sampling,” *Journal of the American Statistical Association*, 91, 1289-1300.

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

Srinath, K.P. and Carpenter, R.M. (1995). “Sampling Methods for Repeated Business Surveys,” In *Business Survey Methods*, New York: Wiley, 153-169.

Sunter, A. B (1989), Updating Size Measures in a PPSWOR Design, Survey Methodology, 15, 253-260.

DISCUSSION OF SESSION 31: COORDINATING SAMPLING BETWEEN AND WITHIN SURVEYS

Lawrence R. Ernst, U.S. Bureau of Labor Statistics
BLS, 2 Massachusetts Ave., N.E., Room 3160, Washington, DC 20212-0001, U.S.A.
ernst_l@bls.gov

1. CATEGORIZATION OF THE THREE PAPERS

Although there are three papers in this session, I view them as fitting into two categories of procedures for controlling sample overlap. The major focus of Ohlsson's paper is on an overlap procedure that he developed, exponential sampling, which is in a class of procedures most commonly used in selecting PSUs for household surveys, although some establishment surveys also use PSUs. Perhaps the key characteristic of such procedures, which originated with Keyfitz (1951), is that the sample size per stratum, *n*, is always small, typically either 1 or 2, and we consequently designate these procedures as S procedures. The following are common characteristics of S procedures. For a given stratum, *n* is predetermined, although in some designs *n* may vary by stratum. Sample units are selected pps. If $n > 1$, the joint selection probabilities within a stratum are predetermined. Most S procedures, including Ohlsson's procedure, although not all, allow for different stratifications in the designs being overlapped. These procedures have been developed for overlap maximization and/or minimization, but not partial rotation. S procedures generally do not use PRNs, with Ohlsson's procedure the only exception that I am aware of among procedures that strictly preserve the desired selection probabilities. (We restrict the discussion to such procedures.) Finally, S procedures typically overlap only two samples at a time, with again Ohlsson's procedure an exception.

The McKenzie and Gross (M&G) and Royce papers, in contrast, discuss overlap procedures typically used for selecting establishments from a stratified list frame, with a key characteristic that these procedures must be capable of overlapping samples for which *n* is large. Consequently, we designate these procedures as L procedures. In L procedures, *n* for a stratum may be either predetermined, as in synchronized sampling of the M&G paper, or variable, as in all the Statistics Canada (STC) procedures described in the Royce paper. The selection of the sample units is commonly, as in both of these papers, although not exclusively, with equal probability. Joint selection probabilities are generally neither predetermined, nor calculated. Typically L procedures do not attempt to control overlap for units changing strata. This is because, in addition to the extra complexity of attempting to do so, the most common stratification change is the occasional establishment changing size class. Both of these papers are exceptions because they do discuss procedures that control overlap with restratification. In fact, this is a key issue in the M&G paper. L procedures are used for overlap maximization, minimization, and partial overlap, and all three applications are discussed in both of these papers. L procedures commonly use PRNs, with the procedure used by the Canadian Monthly Wholesale and Retail Trade Survey the only exception in these two papers. Some L procedures are applicable to the overlap of more than two surveys at a time.

2. OHLSSON PAPER

I consider the procedure described in Ohlsson's paper to be highly innovative and an extremely important contribution to controlling overlap. Before discussing other details of the features of this procedure, let me mention one important feature that is not mentioned in this paper, but is described in Ohlsson (1996). Suppose an initial sample has been chosen, that is one not overlapped with a previous sample, without using Ohlsson's procedure. Although it might appear to be too late then to overlap a subsequent sample with this initial sample using his procedure, it is shown in Ohlsson (1996) that for $n = 1$, PRNs can be assigned retrospectively, conditioned on the initial sample, and the Ohlsson procedure then applied to subsequent samples with all the properties of the procedure remaining unchanged. It is not presently known whether this result can be extended to $n > 1$.

Ohlsson's procedure, like all overlap procedures, has both advantages and disadvantages. The short list of disadvantages will be mentioned first. It is the only S overlap procedure that I am aware of that in the case when $n = 1$ and two samples with identical stratification are overlapped does not yield the optimal overlap. All other S procedures reduce to Keyfitz's (1951) procedure in that case. In particular, Keyfitz's procedure always retains in the new sample any sample unit in the initial design that has a selection probability at least as large in the new design as in the initial design. However, with Ohlsson's procedure, such a unit can be replaced in the new sample by a unit with a selection probability that has increased by a larger percentage in the new design. In addition, there are limitations on the use of this procedure. If $n > 1$ and Ohlsson's procedure was not used in selecting the initial sample, then it cannot be used to select subsequent samples since the procedure for retrospectively assigning PRNs mentioned above has only been developed for the case $n = 1$. In addition, regardless of *n*, if another overlap procedure had previously been used that destroyed the independence of sampling from stratum to stratum, then Ohlsson's procedure cannot be used.

Most of the advantages of Ohlsson's procedure particularly apply when the surveys being overlapped have different stratifications, so we will confine our discussion to this case, which is the more common case in practice when *n* is small. Ohlsson's procedure is quite simple to implement. This is particularly noteworthy when $n > 1$, since most alternatives procedures, including Causey, et al. (1985), Ernst (1986), and Ernst and Ikeda (1995), employ linear programming algorithms, which are generally not simple to implement. Also, which is the key point of the procedure, it preserves the independence of sampling from stratum to stratum. Overlap procedures that require the same stratifications in the surveys overlapped also automatically satisfy this independence. Besides these procedures, the only other procedures that I am aware of that preserve this independence do not predetermine the sample size and hence are not S procedures. In addition to the advantages in variance estimation, Ohlsson's procedure can be applied repeatedly as a result of this independence property. Among alternative procedures, some, such as Kish and Scott (1971), Causey, et al. (1985), and Ernst and Ikeda (1995), can only be used once, since they assume this independence; while others, such as Perkins (1970) and Ernst (1986), which can be used repeatedly, tend to yield an overlap that is further from optimal. To illustrate the resulting problems, consider the selection of the PSUs for the U.S. Census Bureau's Survey of Income and Program Participation. This survey is redesigned every 10 years. The PSUs were selected independently in the 1980s redesign. For the 1990s selection, the sample was overlapped with the 1980s sample using the Ernst and Ikeda (1995) procedure. For the 2000s redesign, since the Ernst and Ikeda procedure cannot be used again, the Ernst (1986) procedure will be used, which should not produce as large an overlap. It might have been better to have used Ohlsson's procedure throughout. It would be interesting to empirically compare the overlap produced by Ohlsson's procedure to that produced by the Ernst (1986) procedure, which produces the best overlap among those procedures that can be used repeatedly because they do not assume independence from stratum to stratum. I suspect that the Ernst procedure would be superior when the surveys overlapped have similar stratifications, since the Ohlsson procedure is not optimal when the stratifications are identical. However, when the stratifications are very different, Ohlsson's procedure may be superior, since it is known that the Ernst procedure does not generally produce an overlap that is close to optimal under this condition.

3. MCKENZIE AND GROSS PAPER

I have always been impressed with the general approach in synchronized sampling of moving the endpoints to the right while keeping the sampling size fixed, which prevents units that leave the sample at one time period as a result of more births than deaths from reentering the next time period as a result of more deaths than births.

The focus in this paper, however, is not on synchronized sampling in general, but on maximizing or minimizing overlap with another stratification of the same survey or with one or more other surveys. As the authors note, the overlap attained using their procedure is far from optimal because of the "partial intervals" problem. The only way that I am aware of to avoid this problem without a transformation of the PRNs is to have the new sample consist of several partial intervals with only units in each partial interval from the corresponding old stratum included, if possible, rather than a single interval with all units included. I suspect that this solution would cause too many operational problems to be seriously considered.

The fact that there is a "bias" with their procedure when the starting point of the selection interval is the last point that produces the maximum score is illustrated by the following simple example. Suppose a stratum in the new design consists of two complete old strata, A and B, with $n = 1$ for the new stratum and both old strata, and with A consisting of more units than B. Then the probability is .5 that the last point that produces the maximum score is the sample point in stratum A and the probability is .5 that it is the sample point in stratum B. Consequently, the selection of the sample unit in the new stratum would not be with equal probability, but instead each unit that was in stratum B would have a higher selection probability than each unit that was in stratum A. I agree, as mentioned in M&G, that the use of the point after last occurrence of maximum score rather than the point of the last occurrence as the starting point of the selection interval tends to reduce the size of the misallocation, although I believe this whole issue is complex and needs further study. Furthermore, for very small *n*, particularly for $n = 1$, this choice of starting point may reduce the effectiveness of the overlap. To illustrate, consider the above example with the point after the last occurrence of the maximum score as the starting point. Then the sample unit in the new stratum would only have been a sample unit in the old design if the stratum A sample unit and stratum B sample unit are the first and last points the new stratum, an event with probability lower than the probability of overlap when the sample unit in the new stratum is selected independently of the selection of the sample units in the old strata.

The avoidable load measure used in the empirical study is interesting. However, it is a measure of high load, while the maximum score used in the overlap procedure attempts to minimize a somewhat different measure. For example, selecting two units, one with a high load and one with a low load may result in a lower score but a higher contribution to the avoidable load than selecting two units with medium loads. I believe it also would be worthwhile to compare the avoidable load obtained in the empirical study using their overlap procedure to the avoidable load obtained selecting the samples for these 12 surveys independently.

4. ROYCE PAPER

Although several approaches for sample coordination are discussed in this paper, STC appears to be standardizing for within survey sample coordination around GSAM, which uses collocated random numbers (CRNs) that have at least two advantages over standard PRNs due to the equal spacing of the selection numbers. First, CRNs help reduce the variability of the sample size when using a fixed selection interval, although, particularly because of births and deaths, they do not completely eliminate it. In addition, an attempt to minimize overlap among surveys by assigning each survey different starting points an appropriate distance apart cannot fail with CRNs, as it can with PRNs if most of the PRNs are clustered close to each other. CRNs also have disadvantages in comparison with PRNs. CRNs cannot be assigned on a flow basis, which is why they could not be used in the Tax Estimates Program. Also coordinating surveys with different stratifications or when restratifying can be more complicated with CRNs, as noted by Ohlsson (1995), since the CRN for a unit is a function of the number of units in the stratum.

One of the important features of GSAM is a procedure for maximizing overlap when restratifying a survey. GSAM accomplishes this while avoiding the partial intervals problem discussed in the M&G paper and thus generally produces a larger overlap. This is done by assigning new selection numbers to the units in a new stratum in a manner that clusters together as much as possible the units that were in sample under the old stratification. Using this approach for coordinating among different surveys might lead to complications, however, because the same unit would have different selection numbers for different surveys.

An alternative to the procedure described for minimizing the overlap between the crops and livestock surveys would be to choose appropriate starting points for the selection intervals for the two surveys and move both intervals to the right. For example, if the sample for one survey was three times larger than the sample for the other in a stratum, then the selection interval for the larger sample could start at 0 and for the smaller sample at 0.75. This alternative should result in a longer time period before the samples could overlap. However, I understand that this alternative was among those considered by STC, but that it produced a larger overlap than the procedure adopted.

The procedure described for reducing respondent burden in SEPH for multi-establishment enterprises results in biased estimates. In particular, I believe there may be a large underallocation of establishments in large enterprises because of this procedure.

The network sampling approach used in UES is quite interesting. It complicates estimation and variance estimation, however, as discussed in Simard and Hidiroglou (1999).

5. REFERENCES NOT LISTED IN SESSION PAPERS

Ernst, L. R. (1986), "Maximizing the Overlap Between Surveys When Information Is Incomplete*," European Journal of Operational Research*, **27**, pp. 192-200.

Perkins, W. M. (1970), "1970 CPS Redesign: Proposed Method for Deriving Sample PSU Selection Probabilities Within 1970 NSR Strata," memorandum to Joseph Waksberg, Washington, DC: U.S. Bureau of the Census.

Any opinions expressed in this discussion are those of the author and do not constitute policy of BLS.