6. EVALUATION OF OVERLAP WITH CURRENT MCBS PSUs

6.1 Introduction

Reasons for Maximizing the Overlap

It may seem paradoxical to select a new sample and, in doing so, attempt to maximize the overlap with the old sample. However, as seen in Chapter 5, when a survey is currently running in the field and a new sample is selected, the degree of overlap between the two samples is a primary factor in the cost of fielding the new sample. Maximizing the overlap produces savings in costs from hiring and training new staff. In addition, there are quality and efficiency advantages of using experienced interviewers that are more difficult to quantify.

Methods Considered

Many methods have been developed for maximizing overlap between samples. In this evaluation, we reviewed methods developed by Causey, Cox, and Ernst (1985), Ernst (1986), Ernst and Ikeda (1995), Brick, Morganstein, and Wolter (1987), Kish and Scott (1971), and Ohlsson (1996, 1999). Each has advantages and disadvantages; none completely solves the problem.

Organization of this Chapter

This chapter first describes the nature of the overlap maximization problem and various approaches to it. Then we discuss the optimization methods we evaluated for future use in MCBS, giving an estimate of the overlap that might be reasonably be achieved in a redesigned MCBS. At the end of the chapter, we discuss the long-range implications for the overlap methods for future redesign efforts.

6.2 Overlap Maximization Problem

Components of a Survey Redesign

The survey redesign problem involves using up-to-date information to re-stratify the frame of PSUs and assign new measures of size to the PSUs. Within each new stratum, the new measures of size determine the probability of selection of each PSU. For most surveys, Westat selects two PSUs per stratum using the Durbin method. (For a description of this method, see Brewer and Hanif 1983.)

The selection of the new sample of PSUs could be done independently of the current sample. To do this we would use probabilities of selection derived from the new measure of size and implement the Durbin method. The idea of maximum overlap methods is to calculate new selection probabilities that take into account which PSUs were in the earlier sample. By using complex mathematical methods, the joint selection probabilities for pairs of PSUs in a stratum can be modified in such a way that overall selection probabilities are maintained and the overlap between the two samples is maximized.

Basic Description of the Approach

The methods developed by Ernst and his colleagues are based on fully enumerating all possible old samples and all possible new samples. Then for each possible outcome – for example, the occurrence of the i-th old sample and the j-th new sample – the probability of that outcome is manipulated so that the following conditions are met:

- The selection probabilities for all the possible new samples are preserved;
- The selection probabilities for all the possible old samples are preserved; and
- The outcomes where the old and new sample overlap are maximized.

Once these pairwise selection probabilities have been determined, the new sample is selected with probabilities that depend on which old sample was selected.

The methods used to determine the pairwise selection probabilities involve linear programming and other optimization methods. Because there are many constraints on these problems – in the most complex version, the number of constraints equals the number of possible old and new samples

– the solution of the optimum overlap requires tremendous computing resources. This is particularly true for larger strata where the number of variables in the linear programming problem can quickly become very large, making the problem intractable even with modern computing resources.

Description of Methods

The earliest approach to solving this problem was presented by Keyfitz (1951). This method was optimal for selecting one unit per stratum when units do not move from one stratum to another as a result of the survey design. However, in redesigning strata, the stratum boundaries will be shifted and it is likely, if not certain, that units will move from one stratum to another.

Perkins (1970) and Kish and Scott (1971) both discuss generalizations of the basic Keyfitz (1951) procedure. In their approaches, stratum definitions may change from one survey to another; however, these procedures still allow only the case where one unit is selected per stratum. One approach to this problem is to split the stratum at random into two parts and select one unit from each, although there is no guarantee that this solution is optimal.

Brick, Morganstein, and Wolter (1987) proposed a very simple approach that allows for more than one selection per stratum. However, their approach does not guarantee fixed sample sizes.

Causey, Cox, and Ernst (1985) presented the problem as a linear programming problem. However, the scope of the problem was well beyond computing resources of that or the current time. A modified approach given in Ernst and Ikeda (1995) reduces the number of constraints, yielding a linear programming problem that at least approaches tractability. Because of the simplifications, however, solutions found using the Ernst/Ikeda 1995 algorithm may not achieve the degree of sample overlap obtained with the algorithm designed by Causey et al. Another modified approach given by Ernst (1986) simplifies the problem even further, again at the expense of obtaining the best possible solution.

Ohlsson (1996 and 1999) has developed a very simple method for maximizing overlap between samples. His method uses random numbers that are permanently assigned to each PSU. His method was originally developed for samples of one PSU per stratum, but later was extended to two PSUs per stratum. Ohlsson's method has the advantage that it is very simple to use and can be used over and over again on the same sampling frame, regardless of changes in stratification. In his later paper, Ohlsson compares a number of methods and finds that his method compares favorably others, performing about as well as the Ernst (1986) algorithm. Ohlsson's method is best used prospectively. The key element of the procedure (the permanent random numbers) can be retrospectively assigned for samples of size one, but not for larger samples. The MCBS was drawn with samples of two per stratum.

In an unpublished manuscript provided to Westat, Ernst (1999) provides an exhaustive discussion of approaches to the overlap problem. These methods are summarized in Table 6-1, which is adapted from Ernst's manuscript. As one can see from that table, many of these methods allow only selections of samples of size $n = 1$. While this can be dealt with in a number of ways (such as splitting the stratum into two parts), there are often assumptions about the way the earlier sample was selected which cause further difficulties. For example, Kish and Scott's method assumes that the earlier sample also had $n = 1$ per stratum. There are further difficulties when the earlier sample was selected using an overlap maximizing procedure.

Table 6-1. Summary of overlap-maximizing procedures

Procedure	Sample size	Different strata?	Optimal?	Uses linear programming	Surveys
Keyfitz (1951)	$n = 1$	No	Yes	No	2
Perkins (1970)	$n = 1$	Yes	No	No	2
Kish and Scott (1971)	$n = 1$	Yes	No	No	2
Sunter (1989)	Small	No	No	No	2
Ohlsson (1999)	Small	Yes	No	No	≥ 2
Causey et al (1985)	Small	Yes	Yes	TP	2
Ernst (1986)	Small	Yes	No	LP	2
Ernst and Ikeda (1995)	Small	Yes	No	TP	2
Pollock (1984)	Large	Yes	Yes	No	2
Ernst (1995)	Large	Yes	No	No	2
Mitra and Pathak (1984)	$n = 1$	No	≤ 3 surveys	No	≥ 2
Perry et al (1993)	Large	Yes	No	IP	≥ 2
Ernst (1996)	$n = 1$	Yes	Yes	TP	2
Ernst (1998)	Large	No	Yes	TP	2
Ernst (1999)	Small	No	Yes	LP	≥ 2

Source: Adapted from L.R. Ernst (1999), The maximization and minimization of sample overlap problems: a half century of results. (Unpublished)

6.3 Methods Used to Construct Test Strata

In order to evaluate the different methods of overlap maximization, we constructed a set of strata using a measure of size based on number of Medicare beneficiaries in each PSU. This section describes the methods used to construct these strata.

Table 6-2 lists the PSUs that are identified as "certainty" PSUs with respect to the counts of Medicare beneficiaries (HCFA MOS) when using a certainty cutoff of 1.0. These are the 20 PSUs with the largest MOS. Each of these 20 certainty PSUs constitutes its own stratum.

Table 6-2. Listing of certainty PSUs (certainty cut-off of 1.0)

STRATUM	PSU	1999 Medicare beneficiaries
A410	001	969,552
A210	001	926,476
A340	001	791,892
A120	001	741,136
A113	001	639,477
A112	001	625,827
A220	001	603,234
A420	001	511,120
A130	001	498,325
A320	001	438,050
B330	005	422,889
A140	001	393,084
A310	001	375,816
A230	001	365,081
B460	005	359,886
B460	001	354,963
A111	001	351,843
A360	001	349,677
B420	020	335,362
A350	001	327,959

The remaining 1,358 PSUs were identified as noncertainty PSUs. They were stratified according to various demographic and socioeconomic characteristics. We considered the following rules during the process of construction of strata:

- Strata should be roughly of equal size with respect to HCFA MOS;
- Strata should not cross Census Regions; and
- Strata should not contain both metropolitan and nonmetropolitan PSUs.

Table 6-3 shows distribution of PSUs within the major frame divisions (Region by Metro status).

Table 6-3. Number of strata within major frame divisions

Region	MSA	Number of PSUs	HCFA Medicare Beneficiaries Counts	Number of Strata	Average Stratum Size
1	No	65	996,707	2	498,354
1	Yes	53	3,793,144	6	632,191
2	No	376	2,967,523	5	593,505
2	Yes	89	4,285,893	7	612,270
3	No	449	4,498,922	7	642,703
3	Yes	138	6,412,399	10	641,240
4	No	137	1,312,586	2	656,293
4	Yes	51	3,311,184	5	662,237
Total		1,358	27,578,358	44	

Note: This stratification procedure was done only to create examples for testing the overlap procedures. The sampling frame does not include Alaska or the Puerto Rico PSUs.

The strata were defined within each Region by metropolitan status class, according to the distribution of percent minority (Black or Hispanic), and average per capita income in the following way:

- Within region and MSA status, we formed size classes based on Medicare beneficiary counts;
- Within each size class, we sorted PSUs by per capita income, then formed strata of roughly equal size, making additional breaks by percentage minority where appropriate; and
- In non-MSAs, we formed strata based on per capita income alone.

Using this procedure, we created 44 strata for noncertainty PSUs, with each stratum having an approximately equal aggregate measure of size.

Note that this method and these stratification variables are not necessarily the ones that would be used in a full MCBS redesign. These methods were used simply to construct a reasonable stratification system that could be used to test various overlap procedures.

6.4 Results of the Evaluation

6.4.1 Certainty PSUs and Special Cases

If a certainty cutoff of 1.0 is used, there are 20 certainty PSUs based on 1999 Medicare beneficiary counts. All these PSUs are in the current MCBS sample. Table 3-5 compares the overlap between certainty PSUs with a 75 percent cutoff for Medicare beneficiary counts with the current MCBS certainties. For this cut-off, there is almost 80 percent overlap; with a 50 percent cutoff, all but two of the current certainties would be included in a redesigned sample.

Among the 44 noncertainty test strata, 4 have no PSUs that were in the old sample. On the other hand, there were three strata where *all* of the PSUs were in the old sample. Assuming two PSUs selected per stratum, the overlap in these seven strata will be six (of a possible 14), regardless of the selection method. That leaves 37 strata, which are the primary focus of the remainder of this chapter.

6.4.2 Results for Ernst's 1986 Algorithm

Overlap and Feasibility

Table 6-4 summarizes the results of running the 1986 Ernst algorithm on the 44 noncertainty strata. The table shows the size of each stratum, target sample size, number of PSUs from the current MCBS sample, the maximum that could be selected in the new sample, and two expected overlap calculations.

The sizes of the strata range from 3 to 81 PSUs. As noted earlier, the computational difficulty increases with the number of PSUs. We attempted to run Ernst's 1986 algorithm on all 37 strata, with successful outcomes in 26 strata. In the 10 largest strata, the number of PSUs was too large for the SAS linear programming procedure.

The expected overlap calculation is first computed assuming independent sampling in the redesigned sample. The sum of the expected unconditional overlap over all strata is 19.8, indicating that approximately 20 PSUs would be expected to overlap with no attempt at controlling the degree of overlap.

Table 6-4. Expected overlap with Ernst's 1986 algorithm

Redesign stratum	Size of stratum	PSUs in MCBS	Maximum overlap	Expected overlap: Unconditional	Expected overlap: Conditional: Ernst 1986	Expected overlap: Conditional: Ernst/Ikeda
14	2	2	2	2.00	2.00	
15	3	2	2	1.41	2.00	
28	3	1	1	0.98	1.00	
40	3	3	2	2.00	2.00	
3	4	3	2	1.67	2.00	
5	4	1	1	0.40	1.00	
16	4	4	2	2.00	2.00	
4	5	1	1	0.42	0.66	1.00
29	5	3	2	1.25	2.00	
34	5	1	1	0.37	0.56	
41	6	2	2	0.84	1.25	2.00
43	8	2	2	0.73	1.40	
7	9	0	0	0.00	0.00	
30	9	3	2	0.78	1.20	
35	10	1	1	0.17	1.00	1.00
6	13	4	2	0.79	1.13	
17	16	3	2	0.54	1.30	
33	16	3	2	0.35	1.89	
36	16	2	2	0.37	1.08	
44	16	1	1	0.14	1.00	1.00
18	17	0	0	0.00	0.00	
31	17	1	1	0.09	0.14	
8	18	4	2	0.73	2.00	
42	18	1	1	0.19	0.26	
19	19	2	2	0.43	1.14	
37	22	2	2	0.14	1.05	2.00
1	28	1	1	0.04	0.90	
20	28	1	1	0.07	0.16	
32	35	1	1	0.09	0.14	
2	37	1	1	0.04	0.92	
21	46	3	2	0.11	1.03	
24	56	1	1	0.03	0.97	
22	58	2	2	0.11		
38	63	1	1	0.02		
23	64	0	0	0.00		
9	69	3	2	0.10		
10	69	1	1	0.01		
25	69	2	2	0.10		
39	74	3	2	0.08		
12	76	0	0	0.00		
27	76	1	1	0.06		
26	80	3	2	0.08		
11	81	3	2	0.10		
13	81	1	1	0.02		
			63	19.8	35.2	

The second overlap calculation is based on the overlap maximization procedure. In all cases, the optimized, conditional expected overlap is at least as large as the expected unconditional overlap. This value is missing for the 11 "large" strata were the overlap procedure could not be used. The sum of the expected conditional overlap for the other strata is 35.2, indicating an expected overlap of about 35 PSUs, which should be compared with the 19 expected overlapping PSUs in 26 strata where the Ernst procedure was run.

Thus there is an expected overlap of 35.2 PSUs among the 31 strata where either the outcome is known or the Ernst algorithm can be applied. In addition, even an independent sampling of PSUs from the 11 other strata would yield an expected overlap of 0.67 PSUs. Thus, there would be a total expected overlap of about 36 PSUs. Since 88 PSUs would be selected from the 44 noncertainty strata, this yields an estimated 41 percent overlap for the noncertainty strata. Overall, there would be an expected overlap of about 56 PSUs, assuming 20 certainty PSUs.

This degree of improvement compares favorably with results reported by Ernst (1986). He reports a 59 percent overlap using the optimization algorithm versus 39 percent with unconditional sampling for the Current Population Survey (CPS) and 81 percent versus 59 percent for the National Crime Survey (NCS). In both cases, the proportion of overlapping PSUs was higher, but the degree of improvement is higher for our example.

There remains the question of whether the expected overlap for the 11 very large strata could be improved. This topic is discussed in the next section.

Performance of the 1986 Algorithm

The programming of the 1986 algorithm is somewhat complex (though not nearly as difficult as the 1995 algorithm discussed below). In this section, we discuss methods used to evaluate the reliability of the algorithm and difficulties encountered in using the algorithm.

We have thoroughly tested the programming and performance of the 1986 algorithm using elementary examples; we have also examined of the SAS OR output for more complex computations done on the test strata. We have worked simple examples by hand and compared the results with the SAS output. We have also used different software for the linear programming portion of the software and obtained the same results. For more complex computations, we have checked to see that the constraints are satisfied and checked the conditional probabilities obtained from the algorithm.

As indicated above, a number of problems were encountered for larger strata. The primary problem seems to be difficulties in meeting the extraordinary number of equality constraints required by this algorithm, often numbering in the thousands. One approach to this problem has been to scale up the constraint equations; for example, we found that multiplying the constraints by 10,000 and then rounding to the nearest whole number allowed us to run the program on larger strata. Another approach to this problem might be to replace some equality constraints with inequality constraints.

Improvements in Performance for the 1986 Algorithm

In running the Ernst algorithm on the largest strata, we encountered difficulties in solving the linear program. As discussed earlier, linear programs are difficult to solve when the number of constraints and variables becomes large. For these largest strata, the number of constraints and variables greatly taxes the computing resources that we have brought to bear on the problem.

It appears that some of the difficulty stems from round-off error within the SAS procedure. We have developed an approach for solving this problem and are continuing to test this on larger strata.

Another approach would be to run the procedure on a larger computer. In testing, we have used a relatively powerful PC. We have not used the mainframe due to the unavailability of the SAS OR software for our mainframe. Also, other (and possibly more efficient) linear programming algorithms are available.

Moreover, by the time that an MCBS redesign would actually be carried out, computing power available even on PCs will be greater, possibly allowing us to run the 1986 Ernst procedure for larger strata even on a small computer.

Finally, the overlap might be improved somewhat by a more careful construction of strata. In Table 6-4, notice that 4 strata have no MCBS PSUs, while 17 have only 1 MCBS PSU. This construction reduces the chances for overlap with the MCBS PSU sample. Through these various devices, we expect that the expected overlap in the 11 large strata could be increased to about 10 PSUs.

The strata for a redesigned MCBS could be constructed to maximize overlap with the strata used in 1980, making adjustments as necessary for new measures of size. As seen in Chapter 3, the new measures of size (with some exceptions) are approximately proportionate to the 1980 population counts. Thus, it is possible that some strata could remain unchanged, or nearly unchanged. This would allow for a greater unconditional expected overlap and also increase the expected conditional overlap.

Results for Ernst-Ikeda Algorithm (1995)

This algorithm is extraordinarily difficult to program.* We have completed a program for running it and have tested it on several of the redesigned strata. The results of these tests are shown in Table 6-4. We have been able to run this procedure on strata with as many as 22 PSUs. We have attempted to run this procedure on strata with 58 PSUs; after 48 hours the procedure was still not completed. When the procedure has been run, it has been successful in achieving a greater overlap than the 1986 Ernst procedure.

Results for Keyfitz Methods

We attempted to use two procedures that are closely related to the original Keyfitz procedure. One, a method devised by Brick et al (1987), worked relatively well in terms of achieving overlap. However, as noted earlier, this method does not control the sample size, with the result that sample sizes varied between 1 and 3; the larger the number of current PSUs, the greater the sample selected using this method.

We also tried using a method developed by Kish and Scott (1971). Since this method is intended for one selection per stratum, one could divide each stratum at random into two "pseudo-strata" of approximately equal size. Unfortunately, this method is not intended for use when more than one PSU per stratum was selected in the original design. We attempted to generalize this method so that it could incorporate more than two original selections; however, we have not yet been successful in this attempt.

---

* Westat is indebted to Dr. Moshe Feder of the University of Southampton for his assistance with developing the software to run this procedure.

6.5 Overlap Maximization and Future Redesign Efforts

Introduction

In many cases, overlap control methods are used with only the current survey in mind. This approach is acceptable for suveys that are unlikely to be repeated. However, in developing overlap methods for on-going panel surveys such as MCBS, future redesign efforts should be considered as well as the current one.

Independence between Strata

One primary concern about re-use of overlap control methods is the extent to which the method require and affect independent sampling between strata. The Ernst (1986) method does not require independent sampling, so that it can be used over and over again. On the other hand, the Ernst and Ikeda (1995) method does require independent sampling between strata; moreover, since the use of this method destroys independence, it can be used at most once.

Besides its ease of application, the Ohlsson (1999) method has the advantage that it neither requires nor affects independence between strata. Thus the Ohlsson method can be used repeatedly.

Conditioning on Earlier Samples

Another concern is that most of the methods developed for controlling the overlap operate by conditioning on the sample selected in the earlier survey. These methods are unbiased when *all* earlier samples are considered, but they are strongly influenced by the earlier sample. Since a major point of the redesign would be to revise outdated measures of size, this continuing influence of the earlier sample – despite possible changes in the measure of size – is troubling.

Because of its prospective nature, Ohlsson's method appears to be unique in *not* conditioning on the earlier sample. Morever, if the measure of size of a PSU in the earlier sample changes dramatically, its chance of being retained in the new sample would decrease. In most methods, the fact that it was in the earlier sample would outweigh the change in MOS.

6.6 Implementing the Ohlsson Overlap Method

Because of the advantages of the Ohlsson method that were described in Section 6.5, this section will briefly discuss some of the issues involved in implementing it.

As noted earlier, the Ohlsson method is primarily intended for prospective use. If the earlier sample was selected with one per stratum, then retrospective permanent random numbers can be generated. However, the method for generating these numbers does not extend easily to samples of two per stratum or more.

There are two possible solutions. First, one could select a new sample independently, accepting whatever overlap occurs. Our analysis of the sampling frame indicates that we could expect an overlap of about 40 PSUs overall, versus the 65 that could be expected with the Ernst (1986) method.

Another approach would be to randomly split the current sample in each stratum into two parts. Using retrospectively assigned permanent random numbers, the new design could maximize the overlap with one of the random subsamples. Using this approach, we could expect an overlap of about 53 PSUs, again comparing this to the 65 that could be expected with the Ernst method.