

DISCUSSION

Joseph Waksberg, Westat, Inc.

A. General Comments

1. During the major part of my professional life that I spent at the Census Bureau, I took it for granted that this type of research would accompany large-scale statistical projects. However, in the years that I've been gone I've come to realize how unusual it is for an organization to be willing to devote the extensive money and staff time required in such long range planning. The Census Bureau should be commended for the thoroughness with which it approaches such problems. Of course, the scale of activities to which this research is to be applied is almost inconceivable to anyone in the survey business outside the Government, and is what justifies this kind of effort. The fact that a group of fairly minor improvements can produce increases in efficiency that are worth tens of millions of dollars is indicative of the size of the surveys involved. The thoroughness does, however, create problems in timing. I was surprised to see that not all the research was completed and many decisions still remain to be made. The papers did not give a timetable for the introduction of the revised samples. It seems quite late to be still considering features of the new design. When will the new samples be introduced and why are they so late?

2. Gary Shapiro's paper used some language that can cause problems. There are a number of references in the papers to how much money will be saved by the redesign efforts. It seems to me this is dangerous language to use in the current political climate. If the idea that you're saving \$34,000,000 gets circulated, it won't take long before someone is around to take it away from you. I'd suggest you change your language to something that is also probably more technically correct; for example, that the improvements you're introducing would require that much additional money if they were to be accomplished only by increasing the sample size.

3. It would be useful to find some way of getting the information reported in these papers to students taking courses in sampling theory. Optimizing sample designs is typically treated as a mathematical problem in sampling theory textbooks. These papers are a good illustration of the fact that the mathematical problems are the easy ones. The difficult ones are issues like: which statistics should one use as the basis of optimization, or what kind of compromises should one make among different optima arising from a variety of items; how to choose the best balance between reducing variance and potential bias; what is an appropriate cost function, and how is this affected by the administrative structure of the data-collection operations; what are the parameter values for the cost

function, and are they likely to change as the sample design varies; etc.

The Census Bureau could make a useful contribution to training methods for statisticians by putting together a short manual on the practical issues involved in developing what is naively referred to as an optimum sample design.

4. The papers have been rather narrowly focused on the Census Bureau's redesign efforts but in the course of the research, a great deal of information must have been put together that could be helpful to other survey statisticians. I'm thinking particularly of cost models, unit costs, and components of variance. Statisticians who develop sample designs are frequently concerned with reasonable cost functions, how to get data on components of variances, etc. A little information on these subjects was presented here, but not in a form readily adaptable to other surveys. It would be helpful if the Census Bureau could organize the data that must have been developed for the analyses discussed here and make them easily available. The written versions of the papers delivered here do contain bibliographies but they are mostly of internal Census memoranda that are inconvenient to get hold of and seem to be generally addressed to very specific Census problems.

5. The redesign plans described today assume that essentially the same data collection procedures will be used as in the past, that is most of the surveys will rely on face-to-face interviews and the CPS will use a mixture of face-to-face and telephone interviews with possibly some increase in the use of telephones.

The Census Bureau has been experimenting with telephone surveys, including CATI methods. Private organizations have, by now, converted many of their surveys to telephone with random digit dialing as the method of sampling. The Census Bureau will probably eventually move in that direction also, although possibly with more of a mixture of telephone and face-to-face interviews than is used by others. This is likely to require major revisions in sample design. Quite different cost and variance relations exist for telephone operations than the ones described in today's papers. I suggest that it's not too early for the Bureau to start considering the changes that will be necessary at such time, and possibly begin to take them into account in the current plans.

B. Specific Comments

1. PSU Definition and Stratification

It was pointed out that the increasing needs for labor force data for individual states made it advisable to redefine the PSU's and explore new methods of stratification. Of course the same requirements are not present for the other household surveys conducted by the Census Bureau, and for that matter for the various CPS supplements. New PSU definitions and stratification will sharply reduce the overlap between current and future samples and thus result in a major turnover of interviewers, affecting both cost and quality. It is conceivable that the most efficient system would be two separate samples including PSU definition; one for surveys requiring state data and the other for national statistics. The second set would presumably cover such surveys as the Annual Housing Survey, the National Crime Survey, etc. The design for the second set could take into account the major constraints used in the past; that is, to maximize overlap and to have a common design for a number of surveys so that essentially the same interviewers could be used. Has the Bureau considered this possibility? In particular, I would think that a sharp break with the past would create serious problems with both the Annual Housing Survey and the National Crime Survey, the first of which is virtually a longitudinal study and the second a rotating panel with a long period of retention for sample units.

I have one other question regarding the PSU definitions. Gary Shapiro mentioned that there will probably be a general reduction in size of PSU's and this would probably increase the variance slightly but result in a substantial reduction in travel costs. I assume that this is based on data on unit costs and variance that were not included in the papers. It would be helpful to make this available since the question of PSU size is one other statisticians face. Incidentally, reducing PSU size will exhaust PSU's more frequently increasing turnover rates for interviewers. I assume this been taken into account in the cost analysis?

2. Area vs. List Segments

If there is going to be an increased use of area samples, the Bureau again will be faced with dealing with the poor quality of many of the maps used to establish area segments. This was an important consideration in Census moving toward list samples. I think the mapping material has improved over the past 20 years but there are still many problems in working with them, such as block numbers on census tapes not appearing on the maps and the lack of any indexing system. I hope Gary Shapiro and the others involved in the redesign efforts exert pressure to clean up the maps, and to get them indexed better.

They will not only increase their own efficiency but will fulfill a useful service to everyone else using Census data for sampling and related purposes.

It would also be helpful to get more information on any experimentation or other forms of research on methods of assigning measures of size to segments in rural or other nonblocked areas. Gary mentioned alternate types of procedures for allocating measures of size to segments. Except for the Census Bureau, almost everyone relies on area segments for household surveys. We are always faced with decisions on whether to make rather arbitrary allocations for measures of size, whether to chunk a large segment, or whether it is better to have the entire large segment in the sample and list it. If the Census Bureau has any cost data and the effects on segment variability of alternate approaches, it would be nice if they made them available to other survey organizations.

3. Clustering and Sorting of ED's

The description of the new way of clustering and sorting ED's in order to increase the efficiency of systematic sampling within PSU's was quite interesting. However, I was rather surprised at the decision to use ED's as the units to be treated. In our work, we have tended to use block groups where they existed and ED's in the rest of the U.S. We were under the impression that the Census STF1A tape which contains block group instead of ED data in the blocked areas was the most convenient tape to work with. Also, I assumed that block maps were more easily available than ED maps, and they were in better shape. Are we wrong in that? What is the advantage of using ED's as units?

In fact, for our last few area samples, we used the block summaries in the STF1B tapes and selected individual blocks, or groups of neighboring blocks, directly from the Census tapes. Of course, the incentives for using blocks instead of larger areas such as BG's or ED's are much greater if area samples are used than for list samples because of the significant reduction in listing costs. With a list sample the cost of selecting a Census address is the same if one takes it from a sample ED or a sample block, and it is probably simpler to take the set of samples used for sample rotation from a larger cluster like an ED or BG than from separate sample blocks. However, I would think the effectiveness of stratification would be improved if blocks were used as the sorting units.

4. Sample Selection of Permit Places

I can visualize some problems with the revised system planned for the selection of permit places. As I understand the procedure, each year a new sample of permit places will be selected with PPS, with the measure of size being the number of permits issued during the year. A Census employee will then visit the Permit Office and list the permits issued during that year.

This would seem to create a serious lag between the time a permit is issued and the time the unit appears in the sample. I suppose it must take about four to six months before data are received from the Permit places reporting annually and the data are processed. The sampling, listing, and assignment to CPS or other current surveys must take another few months. Generally, it appears that the average lag will be close to one-to-two years. Is this correct? If so, are the rather marginal savings worth the small, but troublesome bias?

5. Revisions in HIS

I have several questions on the plans for revision in the HIS. The reported purpose of moving from a list to area sample is to make the sample households and persons available to NCHS for use as a frame for other surveys. I assume that an area sample would be more expensive to operate because its sampling could not be integrated with the work done for other household surveys. It would be interesting to find out what the additional cost is over the next decade, and parallel to this, what are the expected savings to NCHS in carrying out other surveys.

Secondly, there is a curious history of segment size in this survey. The HIS started out with a segment size of about 9. As part of the redesign efforts in the early 1970's, the segment size was first reduced to 6 and then to 4 because this appeared to be closer to the optimum. A new analysis of cost and variances presumably indicates that the optimum is close to 9 again. It is possible that the increase in gasoline prices and other true changes in unit costs have had substantial effects on the cost function and have led to a significantly different optimum. However, it is also possible that alternate segment sizes arise from different estimates of unstable variances and unit costs, which are notoriously hard to estimate accurately, and which may simply have been estimated poorly in both revisions. It would be instructive to find out the extent to which this may hold.

Thirdly, I think there's a good chance the Bureau is overstating between PSU variances in the HIS by including in it what may be a significant amount of a between-interviewer component. Several years ago, in connection with another project, I examined HIS data for the large SMSA's in the U.S., those in which HIS has a decent sample size so that the reliability is fairly good. I was struck by the large variation among the areas for items in which one would expect more stability. The variability existed even after the differences in the sex-age-race-income composition of the areas were taken into account. It is reasonable to speculate that much of this may be caused by interviewer variability rather than real differences among PSU's. I believe earlier studies also indicated that interviewer variability had a significant effect on the total error of HIS. It would be useful to explore this effect in greater detail. It may have a significant effect on the calculation of the optimum number of PSU's.