

DISCUSSION

Seymour Sudman, Survey Research Laboratory University of Illinois at Urbana-Champaign
1005 West Nevada Street, Urbana, Illinois 61801

In a contributed paper session, the chair and discussant have the task of attempting to integrate papers that may be generally related, but not obviously connected. Let me express some of the general themes that I see coming out of these papers and then comment on these themes. I recognize that in doing this I am omitting discussion of other interesting and important design issues that the papers raise, but in my limited time I want to focus on issues related to estimating response errors.

One way of doing this is to obtain information on data quality by asking respondents and interviewers for their evaluations. This is done in the Hubbard et. al. and the Cox et. al. papers and in the Lavin Census Bureau paper.

Another alternative is to compare to other data sources at either the individual level, as in Cohen's paper, or at a more aggregate level, as in the paper by Cox et. al. Another theme of the Cox et. al. paper is that still another way of determining data quality is to observe the procedures used by the respondent to answer a question, such as one dealing with financial issues.

Both the Hubbard et. al. paper and the paper by Bourke deal with methods for asking sensitive questions using randomized response and item counts. The Bourke paper establishes a procedure for estimating non-sensitive demographic distributions for sensitive sub-populations, but assumes that the randomized response procedure has no response error, an assumption that is questioned by the Hubbard paper and by earlier research. An interesting extension of the Bourke paper would be to cases where there is measurement error in the randomized response to the sensitive question either correlated or uncorrelated with the demographic variable.

Let me return to the notion that useful information may be obtained by asking interviewers and respondents about their attitudes toward the survey instrument. I believe that there are some useful facts that derive from such questions, but that any such results must be treated with very great caution. I found repeatedly in studies I conducted at NORC and SRL at the University of Illinois that interviewers' perceptions of the cooperativeness of respondents-how hard they tried and how accurate their answers were had a zero correlation with outside measures of data quality.

To summarize, I don't find anything very useful in the comments by Census interviewers related to the external environment and survey work in general. Thus, the interviewers' comments that pressures to obtain a high cooperation rate lead to increased survey error obviously ignore the tradeoff between sample biases and response errors. There is a fairly extensive literature on this topic with very little evidence to suggest that higher cooperation rates lead to higher response errors.

Interviewers and respondents provide much more useful information when asked specific questions. Thus, I would listen carefully when

the majority of interviewers on the HIS and the CE indicate that the interview is too long and the questions are too complex. These comments would suggest to me that these are areas that need additional exploration using cognitive laboratory methods or through the more traditional methods of debriefing interviewers.

Similarly, the rejection by respondents of randomized response for measuring AIDS related activities is useful information. I would not assume, however, that the item count lists that are being tested will produce error free data because these lists were more acceptable in the focus groups. It may still be that in the actual interview respondents will significantly underreport sensitive behavior even though they understand how the lists work. This has been the experience with randomized responses to very threatening questions. Item count lists are an ingenious procedure that are worth testing, but their value in studies of AIDS and other sensitive areas is yet to be proved.

The most traditional method for measuring response error and improving survey estimates is the use of additional more valid measures for comparison. Thus, the use of medical care providers, as described by Cohen et. al. is a standard technique in surveys of medical utilization and expenditures. The provider has information that the patient or respondent does not have, although it cannot be assumed that provider data are totally error-free. The two major disadvantages of using provider data are well discussed in the Cohen et. al. paper. They are the cost of gathering this additional information and the possible sample biases caused either by respondents' or providers' non-cooperation.

In this paper, the high costs are recognized by obtaining provider information on only a quarter of the cases. It is my impression that this sampling rate was determined from experience in the earlier NMES study and balances improved data quality and sampling variances.

Fortunately for the use of this procedure, the cooperation rates of both respondents and physicians have been high. Cohen et. al. indicate that the overall cooperation rate of households responding to all the waves and signing the provider permission forms was 88%. The data collection from providers has not yet been completed, but a proposed cooperation rate of 85% is anticipated. Thus, cooperation does not appear to be a major issue. Note, however, that achieving high cooperation from physicians is the major reason for the high costs.

The Cox et. al. paper demonstrates the use of a variety of procedures for measuring data quality. One of the easiest methods is to ask respondents how they answered the question; for this study, the questions related to the use of worksheets and records. The fact that only 20% of respondents used worksheets and between 34-50% used records should alert the researchers to the possibility of significant levels of misreporting, primarily about financial data.

I must admit to being struck by the fact that

many of these small business respondents seem to behave in about the same way that household respondents behave. The overall cooperation rates as well as the rates of refusal to answer financial questions seem to be roughly what one gets in household surveys. We know that household income is underreported and one might expect similar results for small businesses.

Finally, in an earlier version of their paper, Cox et. al. discussed editing procedures for uncovering possible response errors. These include internal consistency checks which are especially critical for financial data where it is easy to misplace a decimal or mis-write a figure. In addition, they propose to look at previously compiled accounting ratios to compare to those obtained from their sample.

To sum up, there are a variety of methods for getting a fix on the quality of response to a survey, but they are not all equally useful and precise. The most powerful methods using outside sources for validation are also the most expensive. These provide estimates of the degree of response error and thus make it possible to adjust for response error in the final estimates. Even if this is not possible for cost or other reasons, it is still useful to know about potential response problems so that these can be corrected in the future and so that unwarranted uses of the survey data are avoided.