Discussion

David Morganstein, Westat, Inc.

It is usual to begin a discussant's comments by applauding the work of the presenters and this discussion will be no exception. My applause is not pro forma, however. We all benefit greatly from the experiences described in and the ideas put forth by these papers. For these efforts, I thank the authors. The burn-out some have described is understandable. The committees have taken on a difficult task and shed light on an important issue.

■ Defining Response Rates

Let me begin with a frequently heard exhortation, expressed in many forums outside the realm of Federal surveys: the need for standards. If there is one clear message from the experiences described by these committees, it is the importance of agreed upon definitions. I will mention Deming's name several times in my comments. Now I will only repeat one of his valued messages: operational definitions. If the committees' work leads to an interdepartmental action to prepare definitions regarding routinely used terms such as "response rates," then their work will have accomplished much.

In our own efforts, at Westat, to improve quality, we too have re-learned this simple lesson. I will mention two efforts: the first to document the process of weighting, and the second, to improve the efficiency of our telephone center operations. In both cases, weeks of somewhat fruitless meetings culminated in a simple realization. Well-intentioned and experienced people were using the same terms to mean different things. After a few additional meetings to develop a glossary of terms, the work took on an accelerated pace and progressed rapidly.

To address specifically some of the difficulties these two committees uncovered in drawing conclusions about the set of surveys they studied, I suggest making a distinction between the screener response rate and the extended questionnaire response rate. Although it is the composite that we need to raise, recognizing the components carries several advantages. For one, it may provide for better benchmarking when comparing several surveys. Second, understanding the elements of the process often uncovers methods for improving the product.

The two committees discussed important differences in measuring the response of residential surveys versus establishment surveys. Because of the large variation in size of establishments, we give them unequal weights when sampling them. For the same reason, it is clearly preferable to use a weighted response rate when summarizing the quality of the results. It is not hard to show the cost-effectiveness of additional effort to ensure the participation of the IBMs and GMs of such frames. A similar weighting of effort should be applied when assessing the impact of their absence.

■ Dealing with Nonresponse: Prevention vs. Re-work

Some see the issue of dealing with response rates as a problem in resource allocation. A lot of energy is spent ex post facto making adjustments and trying to find ways to make them better, as described in the paper from Statistics Canada (Binder et al., 1994). There are several papers at other sessions,as well, dealing with alternate, clever and intricate methods of adjusting for nonresponse (Rizzo and Kalton, 1994; and Witt and Folsom, 1994). Yet we all agree that a dollar spent to prevent low response may save considerably more than what may appear to be less expensive efforts spent afterwards trying to smooth over the results. Reduction in survey cost is a deceptively seductive metric for managers to use as a Key Performance Indicator. What is a percentage point in bias worth? How much should we spend to reduce it?

Let's speak first about prevention. How do we learn about improving response rates? This concern takes on added impact as budget trimmers look for ways to reduce government spending. Federal surveys are the most viable supporter of investment in experiments to test response rate improvement methods. If contractors lose on cost because their proposal includes a suggested experiment to determine a better contact procedure, then in what manner will the sponsoring agency learn to make a better product?

One important task that this committee may choose to address is communicating the importance of research to Congress and senior Federal managers. As Deming has reminded us many times, it is the job of senior managers to provide the resources needed for improvement. Exhortations to conduct better surveys and threats to reduce or cut-off funds for research will not yield improved methods. Spending 1% of a survey's budget to test methods that could raise the response rate by 5% may be worth the investment.

In the National Adult Literacy Survey, we obtained both agency and the Office of Management and Budget (OMB) approval for a pilot study of the effect of monetary incentives. This experiment led to the conclusion that we could raise the response rate more by offering a monetary incentive to the respondent than we could by spending the same amount of money on additional follow-up efforts (taking into account the cost of the pilot study). Had the agency and OMB not been willing to spend additional effort in better preparation, we would not have identified this improvement in efficiency. This is just one more survey related example of the value of an "ounce of prevention."

Perhaps the committee or it's sponsors, certainly the research community at large, should focus on ways of teaching this philosophy and its application in survey design to decision-makers. In his American Statistical Association Presidential address the other night, Ron Iman referred to us when he said, "We are our own best customers." All of us in this room agree on the need to conduct research on better survey methods; unfortunately, we don't make the funding decisions. How can we and the statistics profession move from the stage of recognizing the existence of a problem, to the development of a plan to educate decision-makers, to the implementation of that plan?

Turning our attention now to re-work, what have studies in nonresponse adjustment techniques taught us? The work discussed here by Statistics Canada mirrors the results both Westat (Rizzo, and Kalton) and Research Triangle Institute (Witt and Folsom) reported on in another session. The real issue is not finding a more sophisticated model; rather, it is in identifying the variables most correlated with response. When dealing with a repeating survey and all the experience gathered from its many administrations, it is more likely that the survey designers will better understand the response mechanism. The researchers will have a better understanding of what variables to use in the adjustment models, resulting in reduced bias. It is more difficult to identify those critical variables in new surveys, where these mechanisms are less well appreciated. When good predictors are identified, the choice between simple cell adjustments or logistic models for predicting response is not as important.

■ Summary

These committees have shown us, once again, that conventional wisdom may not always be right. The often accepted notion that response rates have been falling may not, in fact, be so. The committees admit that they did not set out to measure effort expended. They can not determine if the cause of consistent response rates is due to the infusion of additional resources. However, the very high level observed, primarily in residential surveys, should give us some comfort.

Let us spend as much effort in the next year, finding ways to educate decision- makers on the importance of research to prevent errors, as we have in the past year to measure them. The many Federal customer satisfaction studies that are being planned or carried out, as we meet, will continue to suffer from

low response rates so long as the resources available for their conduct remain meager. Are we willing to suggest that surveys not worth doing well are not worth doing at all?

I'll close on an optimistic note for survey statisticians: "All good things come to those who weight."

■ References

Binder, D.; Michaud, S.; and Poirier, C. (1994), "Recent Development in Reweighting for Nonresponse Adjustment," *Proceedings of the Section on Survey Research Methods, American Statistical Association* (in this volume).

Rizzo, L. and Kalton, G. (1994), "Adjusting for Panel Nonresponse in the Survey of Income and Program Participation," *Proceedings of the Section on Survey Research Methods, American Statistical Association.*

Witt, M. and Folsom, R. (1994), "Testing a New Attrition Nonresponse Adjustment Method for SIPP," *Proceedings of the Section on Survey Research Methods, American Statistical Association.* ■