DISCUSSION

Joseph Waksberg, Westat, Inc.

It is interesting that the speakers at this session who should represent opposing viewpoints, for example, the point of view of the Government vs. the contractor or a profit-making organization vs. one presumably mainly interested in research, arrive at essentially the same conclusions. They agree on the fact that the present system is not very good, on the problems that exist, and have approximately the same suggestions on how to improve current practices. I do not have any major disagreements with any of the speakers. However, I suspect that the speakers are underestimating the complexity of the situation and are too optimistic about the ability to make general improvements in a vast Federal system.

At least one reason for this is that the three speakers, although having diverse kinds of affiliations, have one thing in common. They represent organizations that have highly competent and sophisticated technical skills; they are concerned with quality; they understand the trade-offs between quality and cost; and they are aware of the many factors in statistical studies that affect quality and can assess the impact of tradeoffs in expending resources on different aspects of quality. Unfortunately, these technical skills do not exist uniformly, either in the Government or in contracting organizations. This is what makes it difficult to conceive of a general and simplified procedure for preparing RFP's and choosing among bidders.

If Tom Jabine were the typical representative of a Government agency and Gene Erickson and Sol Dutka were typical representatives of contractors, the proposal in the Jabine-Pigman paper to have RFP's clearly state the objectives and funds available and leave the details to prospective bidders would undoubtedly produce the best results for the Government. I have much less confidence in the ability of many other Government agencies to choose the best offer when bidders are given such wide latitude. Erickson has pointed out the paucity of information that exists to help in choosing between high response rate or large sample size, on the increases in variances arising from clustering, etc. For many Government agencies, I suspect it is better for them to specify the major parameters of a survey design than to face a bewildering set of offers, some emphasizing sample size, others high response rates, still others more intensive training and supervision, with the agency staff not really knowing how to assess the relative merits of the different proposals. Also, there are probably contracting organizations with competent operational staff and who can produce work of reasonable quality if a Government agency described the required tasks in some detail, although they might not have the technical capacity to produce the basic plans. I am not sure they should be squeezed out of the possibility of doing some of the Federal statistical work.

Several of the speakers have commented on the desirability of the Federal agencies involving survey statisticians more directly in the preparation of the RFP's and in the choice of contracting organizations. I believe this is really the heart of the matter. Until more technically qualified personnel are involved in the contracting process, I doubt that changes in specifications or rules will have much effect. I am not implying that all Government agencies contracting statistical work are lacking such staff, but it is a fact that many do.

Although I agree with the basic content of the papers presented here, there are few specific issues I would like to comment on. First let me raise a few questions on several points in Erickson's paper.

(1) Nonresponse: I don't believe it is good practice to combine nonresponse and lack of coverage in a single measure. There are a number of reasons for keeping them separate: (a) For many surveys coverage is not under the control of the survey manager whereas response is. A combined measure does not provide information on whether the contractor is doing a satisfactory job. (b) Sometimes substitution is used for nonresponse adjustment. The proportion of substituted cases can be considered a measure of nonresponse. It is confusing to attempt to include undercoverage in the same measure. (c) Independent figures are not always available.

I agree with Erickson that coverage problems may be as important as nonresponse. However, I would suggest that agencies require computation of both nonresponse and coverage ratios (when methods exist for estimating coverage), but that these should be reported separately. This, incidentally, is the Census's practice.

(2) <u>Cluster Sampling</u>: I'm surprised to hear there is a controversy on its use. I have not come across it. What I have found, however, is the difficulty of deciding on a reasonable segment size for a particular study, and

the lack of information to help in such discussions. This situation will not be improved unless a body of information on intraclass correlations is built up. Both Erickson and Jabine have pointed out how rare it is to see an RFP which requires computations of standard errors. I have not seen a single RFP that asks for an analysis of between and within cluster variances, although with modern computational methods this would require little additional effort. Contractors are understandably reluctant to propose such efforts since the additional cost could put them at a competitive disadvantage. If the Government agencies do not specify that such analyses are required, statisticians in and out of the Government will never be able to choose intelligently among alternative sample designs.

(3) <u>Research</u>: Calculations of intraclass correlations are only part of a body of methodological research needed to improve data collection procedures. It is shortsighted of Government agencies not to include provision for some methodological research in large statistical projects. There are some exceptions. NCHS has funded research studies in advance of major studies, and this occasionally occurs in other agencies, but such research is quite rare and tends to be specialized.

Let me turn now to the Jabine-Pigman paper.

(1) Level of quality needed: The Jabine-Pigman paper starts off with the assumption that the major problem in Government-sponsored work is lack of quality. Although I have no quarrel with this emphasis, there is another side of the coin that needs attention. Not all surveys need high quality work as is implied here, and in some cases it is likely the Government is paying more for quality than is justified by the analytic needs of the data.

The main issue I found missing in the discussion today concerns the quality of data needed for a particular study. Possibly the title of the session resulted in a concentration on the lack of quality. However, there are situations when higher quality is built into a survey than needed. This occurs, for example, in decisions to use personal rather than telephone interviews (to avoid the bias of excluding nontelephone households) or decisions to include high-cost areas such as Hawaii and Alaska in sampling frames. It would be useful to give some consideration to assessment of the quality actually needed, in relation to the expected uses of

data for a particular survey.

(2) Providing offerers with data on available budget and survey objectives: This is suggested as a way of improving the selection process. Mr. Dutka recommends a similar approach. Knowing the budget available is certainly essential for an intelligent response to an RFP. Keeping it hidden helps neither the Government nor the bidders. RFP's frequently refer to a "level of effort", but it has always seemed foolish to me to engage in such circumlocutions rather than clearly stating the maximum amount of money available for a study.

Asking offerers to develop survey proposals based only on a description of survey objectives is a sensible proposal for the larger agencies, with reasonably competent technical staffs. As I indicated earlier, I am not sure how this would work for smaller agencies. Possibly OMB should explore the feasibility of some kind of centralized system for smaller agencies.

(3) <u>Probability sampling</u>: Explicity stating that probability sampling is expected, and that use of nonprobability methods need special justification is obviously an important improvement. However, the agencies should accept the fact that under some circumstances nonprobability methods are appropriate. It should be noted that the pilot study on survey practices carried out by the Subsection on Survey Research Methods of the ASA used a purposive sample of projects. I assume there was a good reason not to use a probability sample.

If probability sampling is listed as a specific requirement in RFP's, then we may need to be more careful of our definition of probability samples. Will deliberate exclusions from the frame disqualify some sample designs if the words "probability sample" are taken literally? Some typical exclusions are: Alaska and Hawaii, group quarters, nontelephone households if random-digit dialing is used. How about if a Federal agency wants a study in a few locations - one county, four metropolitan areas, etc. Do the areas have to be selected on a probability basis as well as the units within them? There may be some legal ramifications if definitions are not carefully stated.

After hearing the three papers presented, I would like to summarize my own recommendations for improvements in the contracting process. In approximately priority order, they are as follows:
(1) RFP's should indicate the maximum
funds available.

(2) Biders should be provided with flexibility to trade-off different factors affecting quality, e.g., sample size vs. response rate.

(3) A method should be found for involving survey statisticians in the writing of RFP's and the choice of contractors. This is particularly critical for the smaller agencies. Perhaps some type of pool can be established for statistical assistance. (4) For large projects, more use should be made of RFP's requesting preliminary proposals only, with the Government paying for more detailed designs for the two or three best initial proposals.

(5) Uniform and standard definitions need to be established for such concepts as response rate, probability sampling, what constitutes acceptable primary sampling units, etc.
(6) Some part of the funds for large projects should be set aside for methodological research.