

## DISCUSSION

William Seltzer, United Nations Statistical Office

Let me state at the outset that in participating in this session I will be speaking in my personal capacity as an American statistician and not as an official of the United Nations. In other words, the views expressed are my own and not those of the United Nations.

I was somewhat uncertain how to approach the task of serving as a discussant at the present session. Was I expected to "discuss" the three papers, the evaluation programme planned for the 1980 US Population and Housing Census, or the methodology of census evaluation in general? In the end, I decided that I would, to some extent, have to do all three. Thus, although the three papers provided the starting point for my comments, I have not hesitated to make some more general points where appropriate.

To begin with I think it is useful to ask the basic question: Why do we evaluate the census? In other words, what is the function of a census evaluation programme? The answer, widely recognized by statisticians and embodied in the United Nations Principles and Recommendations for Population and Housing Censuses, is twofold. First, evaluation by providing information on the quality of the census results facilitates the proper use of these data. Second, evaluation by providing information on how well the data collection process and its components functioned assists in the planning of future data collection programmes.

Any census evaluation programme needs to be assessed in terms of how well both these functions are performed. This is particularly true in light of the increasing sophistication of some users and uses of census data and the increasing public skepticism about the benevolence and validity of many governmental activities. Historically, the first function was seen as serving primarily the needs of the data user and the second function, the needs of the data producer. In recent years, however, the extension of census uses and the greater statistical sophistication of many census users has tended to blur this distinction. All of us have, or should have, a stake in both the functions of evaluation.

The three papers presented today were prepared by staff of the Bureau of the Census actively involved in planning the 1980 census evaluation programme. I think we are all appreciative of the time they took away from their day-to-day responsibilities to prepare these papers. In so doing, they continue an important tradition of the Bureau. Collectively, the papers presented today provide extensive information about the current status of census evaluation plans. I shall now comment on each paper in the order they were presented and then conclude with a few summary remarks.

Woltman paper on content error. I am somewhat disappointed with the paper by

Henry Woltman on the content evaluation programme. Actually, most of the problems lie more with the content evaluation programme itself rather than with the paper, which after all, only describes this programme. In any case, I find the content evaluation programme, as described in the Woltman paper, unnecessarily restricted in several important ways. First, the paper presents a restricted definition of error defining it in terms of "a wrong value of the characteristic". However, the rightness or wrongness of a given census statistic depends in part on how that statistic is used. Of course, the uses to which census data are put are so varied that one must consider a range of uses. For example, an "error" of two years in the reported age of a person has a quite different meaning for those uses that require an age distribution by single years compared to those uses where a five-year, 10-year or even broader age groups will suffice. Particularly, in the context of a content evaluation programme it seems important to stress what is considered to be error depends on how the data are used.

Second, the paper contains a useful, but incomplete, review of evaluation tools. The omission of any references to analytically based checks, for example, a comparison of census income aggregates with equivalent aggregates derived from national accounts data, is unfortunate. A comprehensive evaluation programme should be planned with both statistically - and analytically - based techniques in mind. Third, the paper provides no indication of how the content evaluation studies described will be of value to specific users of these data. In other words, I would like to know more about the public policy relevance and research relevance of these evaluation studies.

Finally, as Mr. Woltman indicates, decisions on the size and nature of the content evaluation programme and indeed on the other part of the evaluation programme will depend, in large part, on the funds available for evaluation and the costs of individual evaluation studies. It would have been useful if information on the price-tag associated with each proposed study had been included in the paper. In this way, we all could have obtained a clearer notion of the trade-offs that the Census Bureau staff are now facing. Moreover, such a presentation of cost information could also facilitate the mobilization of additional financial and other support by users for individual components of the evaluation programme.

Bateman-Cowan paper on coverage error. The paper by David Bateman and Charles Cowan on the coverage error programme is a very good paper. It makes numerous useful points. Among these are: (1) The paper stresses the dual functions of evaluation; (2) It brings out the arbitrariness of the distinction between coverage error and content error; (3) It takes a comprehensive view of evaluation that encompasses both analy-

tical and statistical methods of estimating error; (4) It gives explicit consideration of important uses and users of census in planning evaluation activities; (5) It reviews the relevant history of the evaluation research programme and thereby explains how and why the current programme evolved; and (6) It describes the specific methods of evaluation in a balanced way, bringing out the advantages and limitations of the various techniques.

On the basis of the paper I have three concerns and one question about the coverage evaluation programme. My first concern relates to the complexity of some of the methods and adjustments described. I think there is a danger that some of the evaluation procedures proposed are so complex that they will create more uncertainties in the minds of users and policy makers than they resolve. Can the results of an evaluation study, once they are available, be presented in a way that they will be comprehensible to most users? My second concern relates to the time scale of the evaluation studies described in the paper. Will the results of the evaluation studies be ready when they are needed? My third concern is confined to coverage improvement programme. As described in the paper, it is not clear to me how the individual contribution of specific coverage improvement efforts to changes in over-all coverage can be identified. My one question is this: why not use other types of administrative records to test census completeness, for example, court records or records of public assistance programmes?

Bailor-Miskira paper on the experimental programme. This is an interesting and comprehensive exposition of a many-faceted experimental programme. I will confine my comments to three of the six activities described in the paper. First, with regard to the Updated/List/Leave experiment I have two questions: (a) Won't the procedure result in some re-introduction of correlated response errors, which self-enumeration was designed eliminate? and (b) What about the people participating in this experimental programme who receive the census form three weeks prior to April 1 and promptly return it? How will they be treated? Second, although I consider the Alternate Questionnaire experiment an important activity, it is not clear to me why this experiment could not have better been done as part of the census pretest programme. In particular, I am concerned about the confused reactions of those getting the non-FOSDIC forms and seeing all the publicity in the media about the FOSDIC forms.

Third, the student intern experiment is, in my view, a very promising effort. If successful, it will help the Bureau carry out the census more effectively and efficiently and at the same time contribute to the training of future statisticians and users of census data. Moreover, this experiment also presents the American Statistical Association with an opportunity to foster this co-operative undertaking between "official" statisticians and their "academic" counterparts. I hope full use will be made of the American Statistician and Amstat News to

inform those teaching statistics courses of the existence of this programme. One potential difficulty, beyond those discussed in the paper, is that a number of interns participating in the programme may over-react to the real-world errors associated with any large-scale data-collection effort. On the other hand, these reactions to errors can provide an excellent opportunity for learning. For this reason, I would stress the importance of the training materials the Bureau indicates it will provide for use in the classroom component of this experiment.

Concluding comments. By way of conclusion there are four points I would stress. The first point is the importance of explicitly involving uses and users in developing any census evaluation programme. However, the goal of evaluating the suitability of data for sophisticated uses and the goal of demonstrating the essential fairness of the census may present somewhat conflicting demands on a census evaluation programme. In these circumstances it is important to emphasize that not only must uses and users be taken into account in planning the evaluation programme but also uses and users must be taken into account in considering the most suitable methods for presenting the results of evaluation.

The second point I would emphasize, and it is linked to the previous point, is the time-scale involved in completing evaluation studies. A population census takes some time to complete and a census evaluation programme takes even longer. But the timeliness of evaluation results, if these results are to be made full use of, is as important as the timeliness of the census itself. How to achieve adequate timeliness without undue sacrifices in quality, cannot be answered simply. Nevertheless, it is an issue that must be addressed.

The issue of the cost of census evaluation efforts generally and the costs of individual evaluation studies is the third point I wish to emphasize. All three papers stressed the importance of costs in making decisions about the evaluation programme but not one paper gave any quantitative information about the actual costs of evaluation. I strongly believe that the wider availability of such information, even in rough terms, would contribute to the usefulness for the Bureau and for the rest of us of public discussions about the census evaluation programme.

Finally, I would conclude with the question, How can ASA members be of help to the Bureau in connexion with the 1980 Census generally and the evaluation programme in particular? For I think it is important that we statisticians, as a profession, be as of much assistance to the Bureau in this work as possible. Such assistance can take a variety of forms: fostering the intern experiment, participation in professional discussions such as the one today, remaining informed about progress achieved and difficulties encountered in the 1980 Census, responding to requests by the Bureau and users for technical advice, providing informed interpretations to non-specialists of the goals and methods of the

evaluation programme, etc. A population census is after all one of the most highly visible statistical activities of any nation. Thus, we are all involved professionally in the 1980 Census and concerned about its proper evaluation. I am confident the Bureau can count on the professional expertise and understanding of ASA members in carrying out its important and difficult responsibilities in connexion with the 1980 Census.