

DISCUSSION

Robert E. Fay, U.S. Bureau of the Census
Washington, DC 20233-9001

Context of the Papers. Robert Bell and Keith Rust have presented their own work and opinions, but it is clear that the papers reflect the authors' experience as members of the Panel to Evaluate Alternative Census Methods of the National Academy of Sciences, charged to look at options for the 1990 census and beyond.

Scientific panels, such as those of the National Academy or those directly convened by various federal scientific agencies, can be useful for many reasons.

Among the reasons are:

- Panels can represent communities of scientific interest to the government. The scientific community is one of many communities affected by government decisions, and panels can speak for these interests.
- Panels can provide knowledge and skills generally not otherwise available to the government. In some cases, the contribution can be to identify relevant scientific literatures that could be reviewed to gain greater insight into current knowledge about similar problems. In some cases, the interaction of the panel members may result in proposals for promising new research directions.
- Because of their advantage of greater distance from the problem, panels represent an opportunity for improved judgment for complex problems. Panels may identify additional alternatives or have the opportunity to weigh evidence more objectively (1).

The preceding list derives from my observation and is certainly not exhaustive.

The specific context of the panel's work should be noted as well: the panel was convened at the request of Congress to provide an independent review of the Census Bureau's plans and research program. The panel's work was made more challenging by the changing nature of these plans. In essence, the panel dealt with a moving target. To illustrate, the Census Bureau revised the timing of the ICM strategy somewhat, affecting Bell's original remarks on this topic (2). Rust based the original draft of his paper on summaries of Alexander's proposals that Alexander subsequently revised (3).

Both individual authors, and the work of the panel

generally as represented by their interim report (4), deserve substantial credit for their careful discussion of the complex problems involved and for their balanced discussions of conflicting considerations.

Remarks on scientific context of census design.

This discussion is also affected by my own context. My primary professional identity is as a statistician; for example, as a member of the American Statistical Association since 1969, and, in contrast, of the American Association for the Advancement of Science only since 1987. The weekly arrival of *Science*, however, presents an alternative, and profoundly more convincing, exemplar of scientific inquiry, one which provokes unsettling comparisons to practices commonplace in the statistical profession. Indeed, continual exposure to "real science" provides evidence of scientists' underutilization and occasional misuse of statistical tools, but also the much higher standards of scientific practice evident in many scientific disciplines than applied statistics.

General questions that this comparison suggests include:

- Is statistics a scientific profession (i.e., more than just a subset of mathematics)?
- Does the format common in the statistical literature overlook useful devices used by many scientific journals, to a degree that inhibits scientific exposition (5)?
- Are (real) mathematicians, apart from statisticians, the dominant culture in any field of scientific inquiry?
- Is survey research, census taking, etc., akin to a technology based on a science or sciences? Is a parallel to engineering appropriate here?

These general questions help to frame the response to a more specific question also relevant here:

- To what extent is there a peer-reviewed scientific literature on the census undercount or census taking in general, and where is it?

Examples exist of applied statistics papers

recording what was done rather than presenting objective evidence that a particular statistical procedure was correct (6). Indeed, the recent film "Forrest Gump," calls to mind a possible modification to a familiar bumper sticker:

STATISTICS happens

The literature on the census undercount includes a number of scientifically troubling aspects. For example, in 1992 the Census Bureau revised its 1992 estimate of 5.3 million missed persons to 4.0 million, including the effect of partially offsetting computer processing errors in the original 1991 estimates, on the order of 0.9 million persons net. (Thus the gross errors were larger.) By most standards of science, this was a large revision. Except for the Census Bureau's internal work, however, I am unaware of published analyses that were redone with the new data. Thus, many lessons that might be drawn from the experience are now clouded. The statistical literature appears tolerant of exposition from which it is difficult to assess scientific validity (7).

The charge of this panel was to look forward to the next census, which has been clearly their focus. At the same time, one might have wished for a firmer assessment of the current state of knowledge on the census undercount, and the panel's work did not substantially improve the current situation.

Contribution of the papers. I would like to personally thank the authors and more generally the panel for the careful consideration to the topics covered today.

The role of a discussant in this case is less clear because both papers are themselves discussions. Bell reviews the Census Bureau's plans for nonresponse followup estimation (NRFU) and integrated coverage measurement (ICM). Rust comments proposals for continuous measurement (CM) developed by Charles Alexander (8) of the Census Bureau. Both papers are extremely effective discussions, and it is hard to improve on them beyond noting points of detail for further consideration.

Bell identifies major issues surrounding NRFU -

- Should NRFU be based on sampling in 2000?
- Should a unit or block sample be used?
- How accurate do blocks need to be?

- Cost analysis is an important consideration in choosing between a unit and block sample.

His discussion emphasizes the importance of the 1995 test results in addressing some of these questions, but also provides advice based on general principles. For example, he finds the case for sampling convincing. He points out that the importance of the block is as a unit of aggregation; thus, he would evaluate the accuracy of block estimates primarily on this basis. He favors a block sample while recognizing the usefulness of an experimental test of its advantages and disadvantages relative to unit sampling; in particular, lower unit costs for block sampling may help offset the variance effect of drawing a clustered sample.

Similarly, he raises key concerns about ICM:

- Possible interaction of CensusPlus with census
- How effective would ICM be?
- How does a CensusPlus compare to the 1990 approach, using a Post Enumeration Survey (PES)?

This list serves to identify key issues that should help guide planning for the test and interpretation of the results. Overall, I find his remarks supportive of the basic Census Bureau research objectives in this area.

Rust's review identifies several key points worth emphasizing:

- Development of CM will require advances on several fronts simultaneously.
- Many measurement issues arise from the continuous form of data collection.
- The question of user acceptance of the products is critical, but more research is required to determine what the actual acceptance might be.
- The issue of outdateness depends on the variables measured.
- Evaluating costs will require further modeling and research.
- There is conflicting and inconclusive evidence on the changes in the quality of the statistics compared to the census long form.

This list and others raised by Rust represents a helpful review of the current status of this research.

I will make one complaint about both papers, namely, that there are too few citations to actual evidence for some of the claims made. Thus, I link at this point my comments about science and my discussion of these papers. Without references at several strategic spots, I find it difficult to identify the evidence on which the authors have based their analysis (9).

In hindsight, my own work has frequently lacked sufficient citations to reflect the necessary science. As one example, the Census Bureau relied heavily on reinterview to measure bias in the 1990 Post Enumeration Survey, yet I failed to insure that an adequate literature review calibrating the performance of reinterview under controlled conditions.

I do not wish to single the authors out for this general problem, however. I again thank them for their generous contribution towards major issues facing the Census Bureau.

REFERENCES AND NOTES

1. An article in *Science* a few years ago provided extreme examples of the social psychological phenomenon of commitment, include the Lockheed Tristar disaster, in which an inside group becomes increasingly isolated from outside criticism or common sense. Although the cases cited are far more drastic than any issue here, there is evidence for the importance of outside viewpoints from those less committed to a particular course of action. (I regret the loss of citation to this source.)
2. The original proposal was to begin reinterviewing households responding by mail shortly after the first receipts; the potential interaction with the ongoing census favored postponing the initial wave of reinterview to reduce such interactions.
3. C. H. Alexander, "A Prototype Continuous Measurement System for the U.S. Census of Population and Housing," presented May 5, 1994 to Pop. Assoc. meetings, is a more current reflection of his views.
4. Nat. Acad. of Sci., *A Census that Mirrors America, Interim Report* (Nat. Acad. Press, 1993).
5. Few statistical journals appear to tolerate footnotes or endnotes. For illustration, I have attempted to follow the style of *Science* here. Among other advantages, notes permit a multi-level discussion of an issue.
6. I invite the reader to assess my claim. I believe that it would be counterproductive to single out a few examples, although a future review of several examples of this sort (lessening the effect of criticism of any one) could be productive.
7. S. Fienberg, *Chance* 5, 28-38 (1992) offers an example of this sort. I have commented elsewhere [R. E. Fay and J. Thompson, 1993 Annual Research Conference Proceedings 71-91 (Bureau of the Census, Wash. DC.)] that the testimony of one witness, Franklin Fisher, was technically flawed, and his results can only be reconstructed by attempting to determine what mistake was made. Fienberg's description is inadequate to suggest even to the alert reader this problem.
8. C. H. Alexander, "A Continuous Measurement Alternative for the U.S. Census (DOC. # CM-10), unpublished Census Bureau report (1993); C. H. Alexander, "Progress on the Continuous Measurement Prototype (DOC. # CM-12) unpublished Census Bureau report.
9. To cite a few examples, I would have preferred to know the specific evidence for Bell's remarks on the relationship between quality of data and time of collection. I do not claim that he is wrong, but should cause and effect be inferred from the observational data on which this may be based? Does he know of analyses of which I am not aware? When Bell recommends the use of a cognitive lab without citation, I cannot know whether this is a blanket endorsement of a developing technology or whether more specific methods are intended. Rust's discussion of the census environment would have been enriched by references to studies, if any, that have analyzed this.
10. This paper reports results of research undertaken by a staff member of the Census Bureau. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.