

DISCUSSION

Eugene P. Ericksen, Temple University
Department of Sociology, Philadelphia, PA 19122

In 1980, the Census Bureau spent \$1.1 billion to collect the census, more than four times the actual and two times the inflation-adjusted cost of the 1970 Census. The 1980 estimated undercount was 1.4 percent, down from the 2.9 percent experienced in 1970 (U. S. Bureau of the Census, 1988). Unfortunately, the "improvement" was more apparent than real. Erroneous enumerations, thought by the Bureau (1975) to be few in 1970, numbered 5 to 6 million in 1980 (Fay, 1988). Subtracting these from the 1980 count increases the undercount to 3.7 to 4.1 percent, well above the 1970 level. More importantly, the racial differential in undercount was scarcely affected, even by the Bureau's official figures. In 1970, the Black undercount was 8.0 percent, for "all others" it was 2.2 percent, and the differential was 5.8 percent. In 1980, the Black rate was 5.9 percent, the "all others" rate was 0.7 percent, and the differential was 5.2 percent, only slightly below the 1970 figure.

The persistent racial differential caused intense pressure for an adjustment to be placed upon the Bureau. They were sued by numerous cities and states, and many technical experts supported the plaintiffs. The Bureau considered the data and methods available for an adjustment in 1980 to be unreliable, and resisted the pressure. At the same time, they realized that spending vast sums of money may not eliminate the undercount. Accordingly, the Census Bureau created an Undercount Research Staff, and directed a substantial sum of money toward research that would improve the data and methods available for an adjustment.

The five papers given in this session illustrate the "two-headed" approach taken by the bureau for the 1990 Census. On the one hand, it is committed to an effort to eliminate the undercount, and the 1990 Census budget will probably exceed \$3 billion. On the other hand, the Bureau recognizes that it is unlikely to eliminate the undercount, and it has put procedures and data in place for an adjustment.

The first three papers given, by Kobilchark, Oliveto, and Gore, by Hazard, Sledge, and Tenebaum, and by Treat and Edson, describe some of the procedures that the Bureau hopes will lead to a complete count. The papers focus on tests and procedures designed to improve the address lists that comprise the Master Address Register (MAR). This is very important, because census questionnaires are mailed or delivered only to those addresses in the MAR, considered to be the "key" to a good census. In 1980, there were two main problems with the MAR: (1) some housing units were omitted, especially in urban areas with high poverty rates, and (2) other housing units were duplicated, particularly in rural areas. Whole household omissions were a major component of the undercount in 1980 (Fay, 1985), but an estimated 2.1 million people live in housing units that were counted twice (U. S. Bureau of the Census, 1985). In discussing plans to improve procedures for 1990, we would expect Census Bureau research to focus on how to reduce housing unit omissions in cities, reduce duplications in rural areas, and reduce costs overall.

Kobilchark, Oliveto, and Gore describe plans for a 1988 dress rehearsal which took place in Missouri and Washington. Because there are as yet no results, the paper should be regarded as "work in progress." Even

so, I find it striking that there is little discussion of the problems that new and modified procedures were designed to fix. They say, "The dress rehearsal employed the full array of methods, techniques,... intended for use in the national census." They do not tell us how they would decide if the procedures were working. For example, there was an extensive "Census Awareness and Products Program (CAPP)." How will the Bureau know if the CAPP was effective? Did it reduce the racial differential in census awareness observed in 1980 (Moore, 1982)? Next, the paper spends considerable effort describing methods to obtain a complete listing of addresses. Eliminating "whole household" omissions, though, is only part of the problem. What are the plans for reducing the number of "within-household" omissions which continue to plague the Bureau? Finally, it would be helpful to know more about how the authors intend to evaluate their results. Will they focus simply on the overall level of the undercount, or will they also be interested in the racial differential?

Hazard, Sledge, and Tenebaum focus on compiling address lists in the rural areas of Washington and Missouri. Again, I am struck by the lack of context for their research. I expected that their paper would start with a review of the problems experienced in rural areas in 1980, and the Bureau's plans for reducing these problems. The paper does provide good detail about proposed methods for 1990, but there are some hints that all is not going well. For example, the authors tell us that they intended to check on the quality of address list updating by intentionally deleting some addresses from the original lists and then checking to see what proportion of these deletions were spotted by the fieldworkers doing the updating. Unfortunately, we are told that "it is not possible to make precise estimates of the suppressed unit add rate because a number of quality control records were incorrectly completed by the postal supervisors and data was not received for about 12 percent of the suppressed units." These are the same types of problems that plagued coverage improvement efforts in the 1980 Census, described in detail by the Bureau (1987). Finally, while the authors do report on some omission and duplication rates in the lists they compiled, they omit the key statistic. When all was said and done in Missouri and Washington, what proportion of housing units were listed twice? Did housing unit duplication contribute importantly to an overcount as occurred in 1980?

Treat and Edson report on address list development in urban areas. Problems with the commercial address lists in cities were notorious in 1980, and the General Accounting Office (1980) reported that lists in many large cities covered only 59 percent of housing. Treat and Edson tell us that the lists will be better in 1990. Unfortunately, the criterion they use is inappropriate. Outside of urban minority neighborhoods with many poor people, it appears that the 1980 Master Address Registers were pretty good. Therefore, one might expect that the focus in 1990 would be to reduce the differential coverage. It is nice to know that the commercial lists for the 1990 Census are better than the analogous list for 1980. What we need to know is whether coverage is better in poor areas like the South Bronx and the West Side of Chicago. Treat and

Edson convince me that lists will be better in general. I am disappointed at their lack of interest in reducing the differential undercoverage.

The last two papers describe research on census adjustment. Woltman, Alberti, and Moriarity tell us about Census Bureau sampling plans for the 1990 Post Enumeration Survey. In 1980, it was difficult to study the geographic detail of the undercount because of the nature of Census Bureau estimation units, usually whole states but sometimes large cities and metropolitan areas. Several statisticians, myself included, called for the definition of more appropriate sampling units and strata which would reflect the nature of the undercount. These sampling units and strata would be defined by factors such as race and ethnicity of local populations, population density, economic factors, and whether the area was a city, suburb, or rural area.

The Census Bureau has clearly heeded this advice, and I believe that their sampling plan is good, though not perfect. I do not believe that the "allocation" variable is a good proxy for the undercount. Allocations create people in housing units where the Bureau is not able to contact an occupant. As a consequence, allocation rates are higher for renters than owners, but they only occur in housing units listed on the MAR. Allocations therefore ignore addresses not included in the MAR as well as omitted people living in households where others are counted. The allocation rate is correlated, but not identical with, the undercount rate. It is too low in areas where the undercount rate is high. Because the sampling rates are to be higher where the undercount is expected to be higher, reliance on the allocation rate causes sampling rates in high undercount areas to be lower than optimal. Secondly, I am not convinced that the owner-renter distinction does a very good job of identifying the uncounted. I would prefer to see a greater emphasis on economic conditions, because I believe that urban poverty is the number one cause of the undercount. Owner-renter status is probably not very important once wealth and income have been taken into account.

The fifth paper by Mulry and Spencer documents the 1986 Test of Adjustment Related Operations in Los Angeles county. The persistence of the undercount problem is documented by the estimated undercount rate of 9.02 percent. This occurred although most of the 1990 Census procedures were in place.

Mulry and Spencer show us that even with all the sources of error taken into account, the reliability of the undercount estimate is still very high. Many of the problems such as matching error and missing data which were thought to plague the 1980 Post Enumeration Program data have been substantially reduced. For example, in 1980, because missing data rates were high, varying the assumptions used to impute missing data could have a big effect. In 1986, there were fewer cases which needed imputation, so Mulry and Spencer found that varying the imputation model had almost no effect on the level of error. I would have found it useful if Mulry and Spencer had listed the important sources of error in the 1980 PEP,

and indicated which ones now appeared to be important and which others no longer seemed to matter.

As I see it, correlation bias is now a major source of error. It is interesting to note that Mulry and Spencer do not consider it reasonable to suppose that it might be negative, perhaps because the data collection period for the Post Enumeration Survey took place well after Census Day. We need to develop models for considering the correlation bias, and estimating the most likely range in which it may fall.

In sum, it appears that a large and differential undercount in 1990 is inevitable. Once erroneous enumerations are subtracted from the count it is likely that the undercount rate will continue to be high. Blacks and Hispanics will be missed at rates higher than Whites. With such a great likelihood of differential error, we need good methods of adjustment. It is heartening to see that the Census Bureau has developed such a productive research program on the adjustment.

Fay, Robert E. II (1985), "Implications of the 1980 PEP for Future Census Coverage Evaluation," *Proceedings of the Section on Survey Research, American Statistical Association*, pp. 413-418.

Fay, Robert E. II (1988), "1980 Census Preliminary Results Memorandum No. 119." unpublished Census Bureau report, March 10.

Moore, Jeffrey C. (1982), "Evaluating the Public Information Campaign for the 1980 Census--Results of the KAP Survey," unpublished Census Bureau report.

U. S. Bureau of the Census (1975), "Coverage of Population in the 1970 Census and Some Implications for Public Programs," *Current Population Reports, Series P-23, No. 56*, Washington D.C., U. S. Government Printing Office.

U. S. Bureau of the Census (1985), "The Coverage of Housing in the 1980 Census," *Evaluation and Research Reports. 1980 Census of Population and Housing, PHC80-E1*, Washington, D.C., U. S. Government Printing Office.

U. S. Bureau of the Census (1987), "Programs to Improve Coverage in the 1980 Census," *Evaluation and Research Reports. 1980 Census of Population and Housing, PHC80-E3*, Washington, D.C., U. S. Government Printing Office.

U. S. Bureau of the Census (1988), "The Coverage of Population in the 1980 Census," *Evaluation and Research Reports. 1980 Census of Population and Housing, PHC80-E4*, Washington, D.C., U. S. Government Printing Office.

U. S. General Accounting Office (1980) "Problems in Developing the 1980 Census Mail List," Washington, D.C., U. S. Government Printing Office.