## DISCUSSION

William E. Winkler, Energy Information Administration

The common theme of the papers in this session is that of editing, weighting, and nonresponse imputation. A good overall paper on editing and imputation is provided by Fellegi and Holt (1975). The basic theme of their paper is that one should determine what analysis will be performed, determine what data and accuracy are required in the data to allow proper analysis, and design the editing system to assure sufficient accuracy in the data to allow adequate analysis.

The paper of Czajka describes a methodology for evaluating an existing edit/imputation procedure that is used by the Statistics of Income Division of IRS. The basic idea appears to be that of drawing a sample of records to receive full-scale edits on major fields of interest and use the fully-edited records as donor records for those records that will receive partial edits. I perceive the basic difference between full-scale and partial edits as being the amount of manual review. If this is the case, then such a strategy is a well-conceived method of reducing the primary cost of long-term edits, that of manual review.

My chief criticism of this paper is that it needs more background description so that the reader can better follow what appears to be a thorough, well-conceived analysis. As edit/imputation strategies, by their very nature, are data-base specific, a more complete description of the analyses to which the data are subject, the key characteristics of the data base, and the problems that the edits are designed to solve is needed. Without the description, I find it difficult to understand why a strategy of allowing two donor records (from the fully-edited ones) per 20 recipient records (from the partially-edited ones) is sound. If the imputation classes are relatively homogeneous and only aggregate totals (not relationships of such totals) are being considered, then such a strategy might be adequate.

The conclusion that preselection of items for editing needs to be improved is strongly supported by the analysis presented. This would seem to necessitate a separate project to evaluate the original edits. Also, the strategy of substituting group means (within properly defined groups) is sound.

The first need, however, is determining better methods of preselection. The major costs of such an undertaking involve the cleanup and validation of the data base. After cleanup, evaluation of imputation using both a modified hot deck procedure and group means can be performed. If the stratification strategy is sound (i.e., imputation groups are relatively homogeneous), then either hot deck or group mean should yield similar results.

The paper of Huang presents a method of adjusting for nonresponse under the assumption that data are missing at random (Little, 1982, see also Rubin, 1977). The paper presents considerable theoretical background and an empirical analysis. The empirical analysis is sufficiently detailed that it presents a convincing argument that the methods the author introduces for application to the Monthly Retail Trade Survey should be given serious consideration. The analysis is sufficiently thorough that it serves as a good tutorial on how these types of analyses should be done.

The paper of Schultz, Huang, Diffendal, and Isaki presents an analysis of different synthetic estimation methods that might be used in adjusting for census undercount. The notable examples show the changes in allocations of the numbers of congressional seats in States under different estimation methods.

It is not clear how the artificial populations used in the analysis are constructed. If characteristics of the artificial populations do not reasonably correspond to the actual data used in making the current allocations, then the techniques considered in the paper may not be applicable. To allow better understanding, the estimators syn 2 and syn DA should be described in detail.

The paper of Copeland, Kwok, and Hoy provides a thorough description of the current method of adjusting CPI (consumer price index) weights (possibly to account for nonresponse or noncoverage) and describes two proposed methodologies for weight adjustment that may perform better.

The first method, raking ratio estimation (RRE), generalizes the existing method of adjusting the imputation weights to two independently derived population totals to adjusting to three independently derived totals. The second methods uses generalized least squares (GLS) to adjust the weights.

The authors currently fit under the assumption of complete independence. As there are three sets of constraints, they can also consider fitting under various conditional independence assumptions as in Bishop, Fienberg, and Holland (1975). By modelling the data and the weighting adjustment more efficiently, they may be able to determine better methods of nonresponse adjustment.

I suggest that the authors provide a more complete description of what are the "reasonable" properties that assure convergence of the iterative fitting (RRE) procedures.

Due to the fact that the overall analysis is still incomplete (i.e., the RRE was not allowed a sufficient number of cycles to determine if and to what it converges), the authors have made no compelling argument for preferring RRE or GLS. They might consider developing a number of situations in which one or the other is to be preferred. For instance, GLS requires computation of variance-covariance matrices. Much computation is involved in obtaining them and such matrices can be quite unstable (different samples yield different results).

The measures MA (raking), MB (maximum likelihood estimation), MC (minimum chi-square), and MD (GLS), used in evaluating the fitting in the RRE and GLS procedures will often yield

similar results. In the complete-data situation presented here, MA and MB will yield the same limiting distribution. To a first order of approximation (using Taylor expansions), MD will agree with MA. Thus, to obtain situations in which MA and MD can distinguish between various limiting distributions, the authors will have to allow their procedures sufficient time (given successive supremum differences of at most 0.01) to converge.

REFERENCES


Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975), Discrete Multivariate Analysis, MIT Press, Cambridge, MA.

Fellegi, I. P., and Holt. D. (1976), "A Systematic Approach to Automatic Edit and Imputation," JASA 71, 17-35.

Little, R. J. A. (1982), "Models for Nonresponse in Sample Surveys," JASA 77, 237-250.

Rubin, D. B. (1977), "Formalizing Subjective Notions About the Effect of Nonrespondents in Sample Surveys," JASA 72, 538-543.