DISCUSSION

Harold Nisselson, Westat, Inc.

If it was one of Dr. Bishop's objectives to make it more widely known that her area of EIA has challenging problems and is trying interesting methods for dealing with them, she certainly succeeded. Problems of incomplete data are pervasive in surveys and censuses. EIA's problems are especially severe because of time constraints.

In reviewing Dr. Bishop's paper several points stand out. First, EIA is looking at its data carefully. This is essential for trying to assure the quality of the energy data base. It is particularly important when dealing with universes which show positive skewness, so that the larger units account disproportionately for aggregate statistics of interest. Second, EIA is not content with quick fixes for current statistics but is trying to improve its basic methodology. The practice of looking at the problems in current data is also invaluable for that effort.

As Dr. Bishop makes clear, data quality assurance requires effective techniques for identifying and resolving instances of suspect data and for dealing with incomplete data. She refers at various places to efforts to follow-up with respondents to resolve questionable reports. However, a certain amount of imputation for incomplete data becomes necessary to meet EIA's time schedules for the publication of weekly and monthly data. A question naturally arises as to how much imputation to accept in data for publication. For example, for one of the items in Table 2 -- "low temperature, metallic" collectors -- 40 percent of the estimate shown was imputed; and for another, -- "medium temperature, liquid" -- 30 percent. In the case of the latter item, which accounted for over a third of the total square feet for all types of collectors, the early raw data was much closer to the later raw data than the imputed figure.

In principle, item imputation and adjustment for nonresponse should be controlled below the point where the potential bias in the survey estimates could obscure the measurement sought. It may be difficult to apply this principle in practice -- in part because the survey may have multiple objectives, and the data may be used by a number of analysts for a variety of objectives. However, a minimal conclusion is that information about the adjustment of data for nonresponse and item imputation should be published to the fullest extent possible, for the quidance of data users. I believe that Dr. Bishop would concur in this.

The reference to bias brings me to a point on which I would differ from Dr. Bishop; that is, her broadening the definition of imputation to mean any "statistical procedure that takes raw data from a subset of the universe and estimates a statistic from the complete universe." It seems to me that her proposed definition tends to blur the distinction between variance and bias in an estimator which is necessary both for efficient survey design and for applying statistical distribution theory to the results of a survey. I believe that a better starting point is that of a probability sample with a consistent estimator. Perturbations can then be introduced in the context of a total survey error model.

Turning to the individual examples, Example 1 the case of Domestic Crude Oil Production is an interesting illustration of how to try to live with an unsatisfactory set of reporting systems while improvements are being introduced. The interim steps benefitted both by careful structuring of the estimation approach and sophisticated methodology. Each of the forecasts for January 1980 shown in Table 1 is above the corresponding actual figure. I have a question as to whether this was a chance event, or an artifact of the methodology used. For example, are the forecasts adjusted in such a way as would cause a discontinuity in the current estimates for January, the first month of the new calendar year?

Example 2, has several points worth comment. As I compute it, the utility reports in the test batch -- considering the three reported items and one derived one -- have a 6 percent item reject rate in editing. This seems high for a continuing survey. It suggests the possiblity the some improvement could be obtained by laying out on the report form some data checks for the utility to make before sending in its report. On a more technical note, the use of Gnanadesikan's influence function as an outlier detector is an interesting application and picked up an additional 10 percent of outliers. I am curious as to how, given an outlier point, one of the variables in the pair (or perhaps both) is determined to be the offender. The method used in the Annual Survey of Manufacturers may be of interest in this connection. With regard to the plan to evaluate the tests in terms of false positives and false negatives, the approach of trying to detect all large errors should help to minimize the variance of the net error after edit. However, focussing just on them may not minimize the expected value of the error after edit; for example, if large errors and small errors tend to be in opposite directions.

In Example 3, weekly data from a selected set of large companies is inflated to an estimate for all companies by assuming that the ratio of stocks for large and small companies for a given week is the same as in the last month for which complete data are available -- generally two months back. Here, typically, the adjustments are on the order of 10 to 20 percent. An extension of this approach would be to use a composite estimate of the last two monthly ratios. More generally, however, it is not clear that cut-off methods are always best.

The treatment of incomplete data is an area of very active interest among statisticians concerned with survey data. A Panel on Incomplete Data of the Committee on National Statistics (CNSTAT) is just completing its work, which will be available within a few months in the form of a 5-volume report dealing with both current theory and current practice. Some of you may remember

the Symposium organized by CNSTAT in connection with the annual ASA meeting last year in Washington. The Panel was able to find only a few cases in which the imputation methods used were evaluated with reference to the missing data from respondents. The work and future plans that Dr. Bishop described represent an almost unique and exciting opportunity to contribute to the empirical assessment of methods for treating incomplete data.