Barbara A. Bailar, U.S. Bureau of the Census

First, let me say that I found all four papers to cover interesting, sometimes overlapping, problems with adjusting for total or unit nonresponse as opposed to item nonresponse. All of the papers were directed toward nonresponse in demographic surveys, not establishment surveys. Nonresponse in establishment surveys can have serious differential effects, so it would have been useful to see researchers taking an interest in that related problem.

All of the papers made reference to an individual having a probability of response or belonging to a response or nonresponse stratum. I do not believe that this is a correct model of nonresponse. No survey organizations practice as if this is a correct model either. The response of an individual to a survey is more of a dynamic process, depending heavily on the structure of a survey. We have seen big differentials in nonresponse rates by interviewers. In fact, interviewers with low nonresponse rates are often asked to convert refusals from other interviewers.

None of the papers addressed the problem of when a questionnaire is treated as a partial response and when it is a unit nonresponse. What are the rules that make it a unit rather than a series of item nonresponse?

Obliquely, each author looks for concomitant variables that will have a high correlation with the survey variable of most interest. These variables will be used for adjusting for nonresponse. Yet none of the authors mentioned computing the correlations directly. In some cases, the highest correlation coefficient would need to be done.

Turning to the specific papers, I found the paper by Drew and Fuller to be an interesting example of modeling nonresponse, but I question the assumptions as being unrealistic.

The assumption that the proportion of hardcore nonrespondents is identical in each category is inappropriate. We know that some categories of persons are more likely to be nonrespondents. I also found the later assumption that the probability that an individual responds on any particular call is independent of the characteristic of interest to be unrealistic. In an employment survey, an employed person is much more difficult to find at home to interview than a person not in the labor force.

Is there any emperical evidence to suggest that the model for the response probabilities is appropriate--$q(x) = B_0 + B_1x + B_2x^2$? What is the rationale for this model?

It would seem important to include some kind of cost data in the model. Later calls might be much more costly with much lower payoff.

The paper by Bailey gave some interesting comparisons of different weighting methods for nonresponse. However, I did not get a good idea of how to choose among the methods. This was mostly caused by the bias expressions all involving population values that are generally not known.

The author mentioned a weighting scheme that uses a classification by number of calls. Applications of this procedure have generally not been effective in extrapolating results beyond the last call.

Another suggestion was to adjust by the reason for nonresponse, separating refusals from not-at-home, and others. To do this, we need good evidence that the classifications by nonresponse category and between nonresponse and vacants are accurate. We know that occupied units are occasionally misclassified as vacant. There are also misclassifications by reason for nonresponse. How would the misclassification problem affect the adjustment procedure?

Finally, Table 3, showing the number of contacts with a unit to get a completed case was surprising. The number of units for which information was collected after the third visit was large. Survey organizations that stop with three calls might find this table of particular interest.

The paper by Schore presented an interesting modeling technique for nonresponse. However, the author assumes that nonresponse is a function only of characteristics of respondents. But, as numerous studies have shown, nonresponse is also a function of the interviewers.

The main problem I had with the paper was accepting the probability of response as being normally distributed. I would assume that the probability of response would follow a very skewed distribution for some characteristics. However, the response probabilities were not well defined in the paper. They were expressed in terms of an undefined B_3.

The paper did give a useful description of nonrespondents to a later survey when baseline data were available. Low response rates were correlated with greater employment and earnings. This finding is consistent with other survey data.

Finally, the paper by Shimizu, Gonzalez, and Jones presented a comparison of four alternative variables used for weighting. The results were not unexpected. This is one case where a plot of the correlations would have been extremely useful.

The authors showed that there was a wide disparity in response rates by region with the South 12 percentage points less than the Northeast. Has there been any investigation of this disparity?

The pattern of results in Table A raised additional questions. In the Northeast region, the variable "1965 beds" gave the worst estimates with the simple inflation method giving the next poorest results in almost every case. In the West, the pattern was just the opposite. Why?