DISCUSSION

Seymour Sudman, University of Illinois at Urbana-Champaign
385 Commerce West, 1206 S. Sixth Street, Champaign, IL 61820

The papers in this session deal with two of the components of total survey error, sample biases and response errors. I shall first discuss the papers relating to response errors and then those dealing with sample biases.

One of the most widely cited facts about social class in America is that most people are reluctant to classify themselves as upper-class. The Blair and Williamson paper tells us that they may well be a response phenomenon that relates to a broad range of autobiographical questions, at least those that are related to consumption.

The phenomenon is much stronger when people are asked to group themselves into thirds. On average, only about 15 percent put themselves in the upper-third. When asked to group by quartiles, 18 percent put themselves into the upper quartile. In a closed version, only 9 percent put themselves in the upper fifth while in an open version 14 percent put themselves into the upper fifth. This ignores the effect of compensating errors. I suspect if Blair and Williamson had used the absolute information on frequencies to classify respondents' own classifications, they would have found even lower agreement.

It is not especially surprising that respondents do not do very well with this task. It is a difficult one that requires respondents not only to estimate their own behavior, but that of others like themselves.

Given that most respondents don't really know where they are in the distribution, giving information about the distribution is not likely to be helpful and, indeed, may simply increase the response error in the data. Schwarz and Hippler have demonstrated that one way the respondent can be given this information is simply by the closed-end answer categories provided. Blair and Williamson provide this information more directly. It would have been interesting if they had replicated the Schwarz-Hippler studies by simply devising categories based on responses to the early waves.

In general, one would conclude that providing information to respondents about others is not an effective way of improving reporting about routine behavior.

The Hubbard, Lessler and Forsyth paper is similar in focus and, unfortunately, in results to the Blair-Williamson paper. That is, most respondents used the cognitive process they were most comfortable with, mainly estimation, and changing the conditions by use of anchoring and instructions on how to answer, i.e., counting vs. a general rule did not affect reporting.

In retrospect this may not be surprising. Many studies suggest that respondents use a satisfying strategy to come up with the first answer they think reasonable, and getting them to switch to a more cognitively complex strategy is very difficult. The key factor in determining whether counting or estimation are used is the number of events, as has been shown by Blair and Burton.

An issue not addressed in the Hubbard, Lessler and Forsyth paper is that of self-presentation. Some of the questions are very sensitive and respondents may well be editing the answers.

The Cox, Elliehausen and Wolken paper as well as some discussion in the Hubbard, Lessler and Forsyth paper remind us of the importance of the respondent's understanding the question, thus echoing the long-standing work of William Belson.

Thus, financial terms such as "initial capital" and "credit" were unclear to many small businessmen. Many respondents thought that money market mutual funds accounts with checking privileges were checking accounts. The new cognitive methods such as use of thinkalouds in a laboratory setting help reduce errors caused by misunderstandings.

A key finding in the Cox, Ellenhausen Wolken paper is that 28% of all small firm balance sheets did

not make sense. It would have been useful for the researchers to followup with respondents to determine why this occurred.

Turning to issues of sample bias, the Thran and Willke paper is concerned with the sample biases that result from the six percent of physicians who cannot be located. They identify a series of factors that differentiate non-contacted physicians from others. From their Table 1, it can be seen that gender is important since there is some likelihood that women may change their names. AMA membership is another obvious factor as is age since younger physicians are more mobile and hospital based physicians are harder to find than those who are office based.

Their Table 1 would be easier to read if Thran and Willke followed the more standard procedure of percentaging across to 100 percent rather up and down. Table 2 shows essentially the same results using a logistic regression. I found it puzzling that AMA membership and Board Certification that are strongly significant in Table 1 do not show up as significant in Table 2.

Tables 3-5 follow up with estimates of the effects of the non-located physicians on three key variables-income, hours worked and patient visits. BAsed on the regressions, there are large differences, more than 20 percent, in income between contacted and non-contacted physicians and smaller differences on patient visits and hours worked. This would suggest that a weighting methodology be used to account for this group. The authors conclude somewhat optimistically that weighting may not be required since the size of the group is likely to shrink in the future, but my reading of their paper leads me to suspect that the non-locates will continue to be troublesome for the same reasons have been in the past.

The paper is an excellent sample of the kinds of analysis that are necessary when a researcher is faced with real-world problems of non-cooperation non-contact.

I assume that essentially the same motivation drives the paper by Hing and Bush on sample bias issues related to the inability to obtain interviews from next-of-kin in nursing homes. The main problem again is inability to locate.

Hing and Bush in their Table 3 properly percentaged for easy comparisons show that the percentage of next of kin who responded drops in non-certified facilities, for patients under 65 and for those with certain mental and physical ailments. Their Table 4 is simply Table 3 percentaged the other way. Tables 6 and 7 put the same variables into a logistic model using only variables found significant in a stepwise regression. The odds ratio in the final column is a useful way of summarizing the size of effects.

One problem I had with the paper is that a clear distinction is not made between statistical significance and importance. For example, because of large sample sizes, a difference of 2.4 percent in response rates between men and women is statistically significant, but it is not clear how important this is.

There is no attempt made in this paper to look at some key dependent variables to see how much, if any, effect the sample biases have. I would have found this helpful. As it is, one doesn't really know what to make of the biases.

Obviously using next of kin greatly increases the complexity of data collection as well as introducing some potential sample biases. I would assume that the usefulness of the data obtained outweighs these problems. From this perspective, I would hope that either Hing and Bush or someone else at NCHS is looking at the quality of data to see how able next of kin are to provide what is required. Again, the use of record check information could be very revealing.