DISCUSSION

Albert J. Reiss, Jr., Yale University

These stimulating and highly competent papers on redesigning the National Crime Survey (NCS) should in no sense be regarded simply as efforts to redesign an important national survey intended to produce social indicators. Rather, these papers address more general issues and problematics in survey design and in statistical modeling and estimation.

ESTIMATING VICTIMIZATION PREVALENCE IN A ROTATION PANEL SURVEY

The Eddy, Fienberg, Griffen paper presents five estimators of victimization prevalence and empirical results for them including an upper and lower bound for the ad hoc estimator based on the proportion of housing units not victimized by crime in a year. Given the conclusion from much research on the NCS that in the short term the probability of victimization for a person or household is much closer to 0 than to 1, it is not surprising that the empirical values for the ad hoc estimator in Table II are close to the upper bound for which it is assumed all missing values are zero. Nor is it surprising that of the estimators $\hat{\theta}_2$ is closest to the lower bound of the ad hoc estimator since $\hat{\theta}_2$ is the simplest model among those assuming that every missing monthly value is a victimization. For, if we exclude the repeated events of a series victimization, as these models do, repeat or multiple victimization is relatively infrequent in the short term, even in a population of victimized persons.

What does seem of interest in comparing the various models is that the more assumptions one builds into the model to approximate the prevalence of victimization in a population of housing units, the closer the prevalence estimate to the lower bound of the ad hoc estimator (though there is some variation among the years for which prevalence estimates are made). These estimators are the result of initial efforts to model prevalence of victimization in a population using housing units rather than households or persons, but the results suggest that differences in model assumptions have important implications for the treatment of missing values. Above all, patterns of multiple or repeat victimization can have important implications for prevalence as well as incidental estimates. Though not apparent in these examples, estimates of crime free months as compared with estimates of annual prevalence are prone to overestimation of victimization when victimizations are assigned to missing values.

Since the model based approach of this paper will be enhanced by increased understanding of the structure of missing data in the NCS, I shall address some matters concerning patterns of missing values based on preliminary research with my NCS longitudinal file. Although my own work has focused on accounting for missing values for persons and single person households by dynamic processes within a population, that investigation also suggests that there are a number of patterns accounting for missing values on households and locations.

First, when each location is visited at a six-month interval, the interviewer must determine whether a housing unit is vacant or occupied and, when occupied, whether the family is temporarily away or can't be contacted within the prescribed number of call-backs. Past studies by the Bureau of the Census of undercounts discloses that at least 10 percent of the dwelling units enumerators conclude are unoccupied are found occupied on a subsequent visit for a reliability check. Assuming the NCS interviewers do not respond solely to information on the control card but rather make an independent judgement about the occupancy/status of a dwelling unit upon each visit, one would expect at least some proportion of missing values for any panel rotation to be due to interviewer judgements about vacancy. Although it is plausible to assume such judgements are randomly distributed within a population, undercount studies show they are closely tied to population and class densities of settlement.

Another source of missing values for housing units is related to their size and that of an occupying household. Single person households for a variety of reasons are more likely to have missing values. Among the more important reasons for these missing values is that single persons, particularly younger ones, spend less time at home and are thus more often unavailable for interview.

There is reason to infer that the size of households also interacts with the mode of interview, i.e., whether in person or by phone. Although all initial contacts with a household are made in person, that is not always the case for subsequent contacts, since the proportion of households interviewed by phone has grown enormously. One suspects that phone interviewing may result in higher missing values in a panel rotation sample survey, particularly due to refusals to be interviewed.

What often is forgotten in longitudinal studies is that low refusal rates in a cross-section cumulate over time across different as well as the same households. This seems to be especially the case for single person households. A 3 to 4 percent refusal rate in a cross-section can easily lead to 10 percent of the households having one missing interview in seven interviews. Indeed in the NCS since one does not follow households when they move the refusals from one panel wave to the next may be even less likely to cumulate across the same households. Moreover, events like vacations, hospitalizations (which differentially affect the elderly), and other temporary absenses from home may well characterize a large number of households during their time in sample. If so this might be observable in patterned variation by months. Households whose interview dates fall in summer months should have more missing values due to vacations. Parenthetically I note how little we know about ways that the dynamic features of a population and of the behavior of its members affect sample designs. Understanding these effects is an especially high priority in utilizing information collected in longitudinal designs.

What seems apparent from considering only these sources of variation for missing values is that assigned values are more consequential for estimating prevalence for some models presented in this paper than for others. For example, the assumption of the Homogeneous Bernoulli model that every household has the same probability of being victimized in any month seems untenable for the missing data pattern of absence due to vacation, given seasonal variation in victimization rates.

Both papers raise another issue--that of effects of time in sample on victimization reporting. I note in passing that while a conclusion of the Michigan redesign is that a shorter time in sample will increase aggregate productivity of victimizations, Table I in the Carnegie-Mellon paper and my more detailed tabulations suggest that the longer the time a household or person is in sample, the less likely assigning 0 weight to missing data underestimates the victimization rate.

NCS QUESTIONNAIRE AND SAMPLE REDESIGN ISSUES

Let me turn next to consider but three of the conclusions in the Michigan redesign paper.

First a number of problems in design--recall of victimizations, for example--hinge upon the length of reference period. Though cognitive psychologists disagree on what is short- compared with long-term memory and on what affects retrieval or recall of information, they seem agreed that, on the average, the longer the time between an event and its first recall for whatever reason (including one's own stimulation to recall information), the more difficult it is to recover that event and the less the accuracy in reporting about it. What does seem to be a safe conclusion is that our highest accuracy in victimization reporting will occur for the shortest inverval of time. Were we to limit reporting of victimizations say to the day previous to interview, we would expect far fewer errors from individual retrieval of information than with the current six-month reference period. What this suggests is the need to explore designs using the recall of events in very short periods of time and their cost-efficiency, particularly for RDD-CAI surveys.

When examining the effect of length of reference period on recall of victimizations, I repeat here my perrenial suggestion that in the current NCS we should inquire about victimizations that occurred between the date of interview and the start of the reference period. Those data would be useful both to assess recall accuracy--since those events in a controlled experimental design should appear in the last month of the next panel interview for the experimental group--and to provide prevalence incidence estimates of victimization for short intervals of time, particularly if data are cumulated across several panel waves.

I likewise suggest that NCS redesign efforts link reference period research to their investigation of the cost-efficiency of sample designs.

My second core issue deals with the matter of determining when respondents within the same household are reporting the same as contrasted with different events of victimization, of how one resolves which of the respondent reports provides the more accurate information about the event, and of how one matches current with past household reports of that event. These problems arise in several contexts in the survey--in reverse record checks, in bounding victimization reporting, and in eliciting information on both household and person victimizations from all numbers of a household.

We need to do much more research on sources of variation in reporting in each of these contexts before evolving decision values about which events and what about them is to be counted. Proxy interviewing, by the way, is a special case of this same problem.

In a rotating panel design, for example, the bounding and matching problems pose special problems in households with more than one member. There we face issues about whether treating everyone as a household respondent can affect reporting in subsequent waves because at least in some households each respondent becomes aware of the fact that others reported the same event. Is it possible that at the next interview some respondents will fail to report the household events because they presume others will or have done so or that some heard a previously interviewed member do so and hence don't report them (which can lead to erroneous conclusions about who reports what)?

The more general problem here is that a cross-section strategy of design testing such as that currently being undertaken in the Michigan redesign--and properly so--doesn't necessarily tell you what is critical in its implementation in a rotating panel survey of a given length.

Finally, let me offer a comment on multiple-frame designs. There typically are two, albeit related, reasons for choosing multiple-frame designs. The first is that one adds other frames to achieve some level of efficiency at a reduced cost. The other is that one adds frames to increase the productivity of critical information, here the addition of crime victimizations, and assesses costs per bit of information. The multiple-frame design advanced in this paper--the joint use of telephone and area frames--hinges on the question of how much productivity of victimizations will be reduced when one reduces costs by relying primarily on a telephone (RDD) sampling frame. That is a reasonable and worthwhile effort.

Although I refrain from tackling issues that affect judgements about cost-efficiency advanced in this paper, e.g., the issue of how one reaches conclusions about the overlap in "unreachables" in both RDD and area frames--a missing value or a "nonsampling error" problem as it comes to be called, parenthetically let me note that I decry the terminology in vogue of referring to "nonsampling sources of error" when discussing errors in sample survey designs. The terminology blinds us to the problems of estimating error in surveys since sampling error is probably the most trivial among the major sources of error in sample surveys.

How the addition of frames can increase productivity and accuracy of information seems especially worthy of investigation. Comparing estimates of victimization by crime using different sampling frames, we know that school surveys provide considerably higher rates of victimization for school-age respondents than do household surveys. What would happen to victimization rates in the joint use of school and household frames? Or with the joint use of police jurisdiction and area frames? Each frame poses serious questions about the accuracy of information as well as of its cost. But since a core issue in redesign of the NCS is the under- rather than over-estimation of victimization by crime, our concern should be on multiple-frame designs that increase productivity of victimizations. Is there a multiple-frame design then that will provide reliable estimates with increased productivity of victimization at relatively little, if any, increase in cost over that of the current area probability survey?