DISCUSSION

Martin H. David University of Wisconsin - Madison

The SIPP was conceived as an instrument for measuring transitions and change. The conception has yet to reach full term, and these papers indicate some of the major technical problems that must be solved to use the longitudinal aspects of the design successfully.

Estimation of persons' characteristics from the panel is essential. The scheme proposed by Kobilarcek and Singh has several strong points:

a. COHORT ORIENTATION. The representative sample is a sample of the household universe at the time of the first wave of the panel. The longitudinal change data are estimated with respect to that initial population.

b. DIFFERENTIAL TREATMENT OF WAVE 1 AND SUBSEQUENT NON-RESPONSE. The non-response at the first wave is a typical cross-sectional problem for which well established techniques already exist. Subsequent non-response can be conditioned on Wave 1 data; less is known about these conditional non-response rates.

c. WEIGHTING FOR MISSING WAVES. The absence of one or more waves can not yet be modelled with an appropriate imputation scheme for longitudinal estimation. Thus weighting should be used to deal with partial response over the period of interest.

Point (a) implies that births into the sample universe are ignored and data for persons who leave the universe must be measured or imputed. Ignoring births is not of great importance, since a new representative panel is available every 12 months, and provides a source of information on additions to the household universe. When (a) is followed to its logical conclusion, the implication is that after 8 waves of measurement data on change refer to the experience of the initial cohort, and can not easily be reconstructed into a measure of retrospective reports of change from a representative sample of the universe on the terminal date of the panel.

Point (b) implies that it is appropriate to use different variates to predict probability of "complete" response, given response to wave 1, than the variates used to predict probability of response to wave 1. My main quarrels with the proposed non-interview adjustment are threefold: evidence, variance estimates for the weighting procedure, and modelling. It is inexcusable that no empirical evidence is offered for the variates that are chosen to provide non-interview adjustment. Work by McArthur and Short [1] shows that at least two of the variates chosen are not significant in a study of attrition to the fifth wave, conditioned on response to the first. I would prefer no non-response adjustment to an adjustment that is not substantiated by strong, published, empirical evidence.

Second, I am concerned about mean square error of estimates. Nothing in the argument presented suggests that the weights calculated are stable, even if they are unbiassed. Indeed, the collapsing procedure used by the Bureau is subject to unknown sampling variation so that one can not be assured that the technique for weighting is not distorting measures of change or transitions.

Third, the appropriate statistical tool for smoothing the weights and understanding the statistical properties of the procedure is a model of non-interview. Multi-variable probit or logit would appear to be the preferred technique. Such a technique would require the Bureau to structure the model according to some well-articulated hypotheses rather than jumping from one data-fishing excursion to another. I do not believe that adjustment by classification is a technique that will be acceptable to most analysts of the SIPP data. They will want instruction in general techniques from the Census.

Point (c) realistically declares that, given present technology, a complete observation must contain all waves for the period for which analysis is desired. The Bureau proposal is overly restrictive. Interviewed persons with complete data are rejected if they reside in households where one or more observations are missing. Those persons tend to live in unusual households (3 or more persons over the age of 18), and the non-interviewed persons tend to be younger persons with some economic independence from the others in the household. A preferable procedure is to include such persons in the interviewed population. Their data will be complete, except for variables that are defined on the household as a whole.

The proposal before us is not sufficiently general to meet present analysis needs. The concept of a longitudinal panel of waves 1-3 has already been made obsolete by Williams [2] who is linking waves 2-5 for analysis of annual poverty and work by McArthur [3] which links waves 1-5. Their studies use w_i from the first wave or no weights on samples that are not clearly described with respect to inclusion of imputed waves or entrants to the sample (Williams). It is extremely important that the Bureau issue a paradigm for constructing weights to include all of the available public use data, because analyses will be done on the longer panels. The Bureau has been negligent in failing to produce this document at the same time that it released the public use version of Wave 3.

Some other blemishes on the proposal need to be mentioned. The Bureau is discarding data for people who marry (or move in with) other members of the sample, because its data processing programs are inadequate. This is only worth noting because the Bureau intends to propagate this stupidity through the remainder of the 1984 panel, the 1985 panel, the 1986 panel, and the 1987 panel! Second, as I understand it, no weighting adjustment is made for sample loss due to movement out of the universe. Imputations are required for such cases. For the deceased, appropriate imputations to arrive at 12-month totals can be made by assuming zero income in the months after death. However, no such technique exists for persons who move out of the household universe. The proposal implied by using an indicator variable for months after leaving the universe, is that the Bureau wishes to censor data from such cases. I can not see a better alternative, but I think the problem should be explicitly described, and the distinction between deceased and others should be clearly drawn. Third, the raking of the SIPP to the CPS totals appears to be out of place. Much is to be learned from treating these two measures as independent, and I do not see the need for imposing sampling error from the CPS on the weigths for SIPP. Last, I should mention that the proposal only covers 8-months of data for rotation group 4. 12-months of data are available only for 75% of the sample.

This proposal does give a concrete framework for building longitudinal weights. It should be generalized immediately to assist those who are doing analyses on the 1984 calendar year and to provide the appropriate extentsions for analyses of the full panel and year-to-year change.

HUBBLE AND JUDKINS

The model developed by Hubble and Judkins is exciting. The execution raises questions. First, nothing is done to establish that the reinterview observations are more valid than the originals. The references to unpublished memoranda and 1968 CPS material do not inform the reader. Lack of documentation makes it unclear why false negatives and false positives from the reinterview imply the conditions C1 = C2 C3 = C4 that are asserted. Moreover, since the timing of the reinterviews is not described it is impossible to know why C2 = C5 and C4 = C6. Part of my confusion about this may have to do with the absence of a definition of period, which I have taken to mean the reference period for the SIPP (4 months prior to the month of interview). Without more information about the reinterview program for SIPP it is impossible to judge whether the approximations made late in the paper are appropriate.

I believe that attention to the role of errors in estimating annual gross flows is also misplaced. Considerable interest relates to the instantaneous probabilities that persons (households) will enter the Food Stamps program (or some other state) and the probabilities that they will leave the program (or state). Such probabilities have already been estimated by Carr, Lubitz, and Doyle [4], using a discrete time (monthly) Markov model and data from the ISDP. Such models may or may not include duration dependence of the transition probabilities. Whether or not there is duration dependency, the statistic of interest is not a year-to-year gross flow, the statistic of interest is the conditional probability of remaining in a given state given the date at which the individual entered the state. Year-to-year change will be net of all persons who make two transitions in that span of time, and will therefore not relate to the cumulative effect of a monthly Markov process on the distribution of the population by state.

If this opinion is correct, then the Hubble-Judkins model in equations (2)-(4) can usefully be applied to correct flows measured between waves for month-to-month change. Such flows will affect only one rotation group for every change to be estimated. However, application of the technique to this problem raises other questions. The measurement from month t+1 is four-months' recall; the measurement from month t is one-month's recall. Applicability of the proposed technique requires that such differences in recall are not associated with differences in error, or that the model be extended for systematic effects from recall that could conceivably be estimated from the other rotations.

An intuitive argument that leads to the same conclusion is that (2) will have different arguments for every choice of elapsed time for measurement of gross flows, implying different estimates for M. If the error process is incorporated into a discrete Markov model, the implications of parameters of the error process for measurement of change for different intervals can be computed.

While correlated errors over periods as long as one year appear quite plausible for those situations in which an episode has been forgotten (whatever the cause -- length of recall, proxy, or interviewer errors), more specific models of the correlation structure might be appropriate to the problem of telescoping. For example, if my last spell of unemployment terminated less than four months ago, I should report unemployment in the current SIPP interview. If I fail to remember the one week of unemployment in the first reference month that is the only unemployment to be reported, I have telescoped the termination of unemployment back in time. We would expect the probability of such telescoping to be small and to diminish, the longer the period for which telecoping occurs. In the limit, we would expect the probability of telescoping yesterday's unemployment to be infintesmal. This conceptualization leads to a rather different model of the correlation of errors at two points in time than the model proposed for estimation. Both need to be tested.

The authors are suitably sceptical of their estimates. I would argue that the false negative probabilities estimated from the reinterview ought not to be arbitrarily increased from evidence in the ISDP validation. That validation was certainly encumbered with matching error that is not present in the present context. How one should discount the ISDP rates for matching error is any one's guess, but the ISDP error rate is not necessarily a better indication of truth than the reinterview. (Again, it is hard to appraise the author's judgement because no data about the reinterview are provided.)

The authors should be congratulated for a promising start on an important problem.

WEIDMAN

Weidman's paper underscores my comment about the importance of a Census document on the frameowrk for longitudinal samples. It is not obvious that his selection of data represents any meaningful universe. Weidman professes to be interested in entry and exit probabilities for a number of different income types. However, analysis of the rate with respect to income type 1 is conditioned on receipt of income type 1 at some point in the sample or receipt of selected other income types. In addition, the sample is conditioned on complete data and continuing membership in the household population. These conditions vitiate meaning that might be assigned to demographic differences, per se. Weidman maintains that the comparison of such differences across rotation groups to examine the difference between a measurement of status change that is within a wave and a measurement of status change that is between waves will be informative. Given sampling error that attaches to any estimator from a roation group and the fact that error processes may differ between the included and the excluded population, I find this argument doubtful. When the argument is coupled with a totally ad hoc fishing expedition and no statistics I am dismayed. One would hope to find behaviorally motivated hypotheses about response error and a thoughtful statistical structure to deal with the problem.

References

1. McArthur, Edith and Kathy Short. 1986. Life events and sample attrition in the SIPP (This conference)

2. Roberton Williams. 1986. Poverty rates and program participation (This conference)

3. McArthur, Edie, Jeanne Moorman, and . 1986 Following children in SIPP (Paper presented to the Population Association of America, May, 1986.)

4. Carr, Timothy, Irene Lubitz, and Patricia Doyle. 1984. Measuring entry and exit rates to the Food Stamp program (Food and Nutrition Service, USDA, Washington, D.C.)