DISCUSSION

Brenda G. Cox, Research Triangle Institute

To derive unbiased estimates from survey data, selection probabilities for each sampling unit must be incorporated into the estimation procedure. Survey statisticians typically meet this requirement by constructing sampling weights that weigh survey data to reflect differential sampling of the population of interest. Nonresponse and undercoverage also lead to a distortion of the sample with respect to the population. To compensate for the resultant bias, survey statisticians partition the sample into groups related to response to study variates and then adjust respondent weights within groups so that they sum to the group's known or estimated population total.

The three papers in this section deal with how to use person-level population counts to construct a family-level analysis weight. Lemaitre and Dufour evaluate alternative methods by comparing Canadian Labour Force Survey results with Census tabulations. Alexander and Roebuck generate artificial data sets with known properties to compare six minimum distance weighting methods. Finally, Zieschang presents a generalized least squares weighting approach and compares it to the principal person weighting technique currently used in the U.S. Consumer Expenditure Survey.

1. The Weighting Context for the Three Papers

These papers focus on household surveys with multi-stage nested designs where households are the ultimate sampling unit. CE analyses use consumer units (CUs), which can be loosely described as economic families. The Labour Force Survey uses families as the analysis unit, where a family is defined as related persons living in the same housing unit.

The surveys begin by constructing sampling weights reflecting the multi-stage selection probabilities. Both surveys compensate for total household nonresponse via weighting class adjustment of the responding household's weights.

Household undercoverage is not accounted for by such nonresponse adjustments. Undercoverage simply means that the frame used for sample selection did not contain all population members. The mechanism to reduce undercoverage is post-stratification adjustment of the household weights. The best counts to use are household-level counts for categories related to (1) propensity to be undercovered or fail to respond and (2) response to survey variates of interest. Presently, such household counts are not available, either in the U.S. or Canada.

An alternative procedure implemented by the Census Bureau and Statistics Canada uses person-level counts for categories defined by age, sex, and race. (These counts are projected by aging the last census and updating these results based upon birth and death records.) The weighting procedure begins by attaching the nonresponse-adjusted household weight to each person within a sample household. Then these person-level weights are adjusted within cells so that the cell's adjusted weights sum to the population projection for that cell. At this point, a family member is arbitrarily chosen to represent each family. The post-stratified person weight for this "principal person" becomes the family's analysis weight.

Note that this procedure is an imperfect form of post-stratification. If these family-level weights are applied to each family member, tabulations by age/race/sex categories will not sum to population projections.

2. The Lemaitre and Dufour Paper

Lemaitre and Dufour consider three alternate methods of creating family weights for the 1981 Canadian Labour Force Survey:

(1) using the head as principal person,
(2) using the spouse as principal person, and
(3) using the harmonic mean of the weights.

These weights were evaluated by comparing estimated totals to 1981 Census tabulations.

Regardless of weighting method, when age and sex were used for post-stratification, families of two or more were usually overestimated while unattached persons were consistently underestimated. One interpretation of these results is that nonresponse and undercoverage produces a sample that contains proportionally more multi-person families than the population. Post-stratification to person-level counts will cause person-level weight sums to total to population counts but will not remove the sample distortion unless these person characteristics are directly related to family size.

It is unfortunate that Lemaitre and Dufour did not include estimates based upon the nonresponse-adjusted household weight. Comparison of these results to census data might provide an indication as to whether the various family weighting procedures were reducing, increasing, or having no effect on the distortion of the weighted sample.

Their next analysis tested the three methods of family weighting when marital status counts were used. Post-stratification using marital status produced less biased estimates of unattached persons and typically less biased estimates for multi-person families, regardless of whether the weight was derived from the head, female spouse, or harmonic mean.

The authors note that in practice the gains may not be as great as observed in this study (where 1981 Census totals were used) as post-census projections must make assumptions that become progressively less true as time since the census increases. A worse case scenario should be investigated where age/sex/marital status projections from the previous census are used for post-stratification and the resultant family estimates compared to 1981 Census data.

Finally Lemaitre and Dufour defined a family-level weight that exactly reproduces specified person counts. Their proposed estimator is a regression-based approach that minimizes the squared deviation between the final adjusted weight and the nonresponse-adjusted household weight. The proposed estimator produced improved estimates of unattached persons and some improvement in estimates of total families. However, more than 5% of the sample had weight changes in excess of what would be typically encountered. This may be indicative that the regression-based approach changes the underlying distribution of the sample within cells in a pervasive and ill defined manner.

3. The Alexander and Roebuck Paper

Alexander and Roebuck discuss how to revise the household weight (1) to incorporate the person-level counts while (2) reducing the bias and variance of survey estimates and (3) producing one weight that can be used for both household and person level analyses.

Constrained minimum distance methods force weight totals to add to control totals while minimizing the distance between the adjusted and unadjusted weights. Household and person level versions of three distance functions were created: generalized least squares (GLS), minimum discriminant information (MDI) and maximum likelihood estimation (MLE). The authors provide algorithms, some new, to obtain solutions to the constraints and describe properties that simplify computation.

The authors prove that, for each method, all households of the same type--where type means the same person counts by age, race, and sex for instance--receive the same weight adjustment factor. This property means that the underlying sample distribution within household type is not disturbed in weight adjustment. While somewhat reassuring, it should be noted that a large number of post-stratification categories are typically used which results in a very fine partition of the sample. For the 48 age/race/sex cells used in CE weighting, household type is defined by a vector in 48 space.

Alexander and Roebuck also characterize the performance of the estimators under various assumptions about undercoverage bias. Model 1 assumes that most undercoverage is within-household loss. Model 2 assumes that the undercoverage is all at the household level. Finally, Model 3 assumes 10% undercoverage and that undercoverage is evenly split between whole-household and person-within-household undercoverage. The authors consider this a "more realistic model." Evidence to support this conjecture would be of interest.

For the eight family types, the least biased estimate was most often associated with the principal person method (5 of 8) and next most often with the MLE person method (2 of 8). The best estimate for total households was associated with the MLE person method. Further, when the various distance measures were themselves used to evaluate the accuracy of the weighting approaches, the MLE person method was uniformly best. When expenditure estimation was considered using Models 1 and 2, all seven estimators overestimated the mean with the principal person method least biased. When Model 3 was used, overestimates were produced when it was assumed that the reference person reported for the entire household and underestimates when it was assumed that the reference person reported for listed members only.

Note that these results are for artificial data sets only. The findings are revealing and suggest that before implementation of constrained minimum distance weighting, candidate estimates--including the nonresponse-adjusted household weight--should be evaluated using actual survey data and comparisons to the Census.

4. The Zieschang Paper

The Consumer Expenditure Survey (CE) has two components: (1) a survey where each household completes a diary of expenditures for a two week period, and (2) a rotating panel survey where consumer units report expenditures in five interviews conducted every three months. CE deals with undercoverage via a person-level post-stratification where each consumer unit is assigned the adjusted weight of the principal person (Alexander, 1986). The principal person is the female spouse of the householder (if present) and otherwise the first listed householder.

As an alternative to the principal person weighting methodology, Zieschang proposes an extension of the household-level generalized least squares (GLS) weighting procedure. This approach minimizes a weighted function of the squared deviations of the weights before and after adjustment subject to meeting constraints imposed by the control totals. The extended procedure allows simultaneous weighting of the two survey components--diary and interview--to produce family-level weights that reproduce the person-level counts used as control totals in post-stratification.

The GLS procedure was implemented using CE data and evaluated with respect to the principal person weights. For variance estimation purposes, weights were computed for the full sample and 20 sample replicates. Control totals were used for 24 age/race/sex cells. Domains chosen from the composite group were region, sampling frame, tenure, and family type.

Basically, GLS reproduced the control totals and "integrated" the composite totals as it was designed to do. It improved the precision of population counts but one has to ask, "Is this an important component of CES reporting?" For family income before tax, a correlate with what CE is really interested in, improvements were minor except for the population total.

Zieschang notes, "noninterview adjusted person counts fall short of the controls in a pervasive if differential pattern across person types and consumer unit size classes." The latter pattern is not allowed for in principal person or GLS weighting. How effective the proposed GLS method is in improving the quality of survey estimates is directly related to the relationship between the GLS weight adjustment applied within each household type (a vector in 24 space for this application) and the effects that undercoverage and nonresponse produce for that household type.

5. Concluding Remarks

The first step in attacking undercoverage is preventing its occurrence. Household undercoverage reflects deficiencies in either the sample design or its execution. The undercoverage encountered in the two surveys described in these papers -- the U.S. Consumer Expenditure Survey and the Canadian Labour Force Survey -- results in the loss of 5 to 10 percent of the households as compared to projected population counts.

Some of this undercoverage may be attributable to the failure of the sample designs to properly account for new growth. An example from the 1977 National Medical Care Expenditure Survey (NMCES) illustrates this point. The U.S. civilian, noninstitutionalized population in 1977 was projected to be 212 million. After adjustment for household and person nonresponse using weighting class adjustments only, the NMCES analysis weights also totaled 212 million. (This can be backed out of Table 4-2 of Cox and Cohen, 1985). In short, NMCES did not encounter the 5 to 10 percent undercoverage that the surveys in these papers encounter when compared to current population projections.

It is well known that as time since the last census increases the accuracy of projections decreases. For instance, 1980 projections indicated a civilian, noninstitutionalized population of 218 million while the Census recorded 223 million, for an undercount of 2 percent in the projections themselves. This indicates that, even for domains included in the counts, post-stratification to population projections can only partially compensate for undercoverage.

The second step in dealing with the problem posed by undercoverage is to characterize the nature of the undercoverage. Lemaitre and Dufour show that single person families tend to be lost more often due to undercoverage. Hence, counts by size of household are clearly needed to compensate for bias due to household undercoverage.

To date, household level projections are not available, either in the U.S. or Canada. To get around this deficiency, the household level post-stratification step is skipped and a family weight is backed out of the person post-stratification process. It is difficult to define what the undercoverage model is for such a procedure. The principal person weighting procedure is an unsatisfactory solution at best for household undercoverage and should be avoided if household projections are possible.

The alternative procedures investigated by the three papers in this section are also unsatisfactory. Basically the problem with these estimators is that a mathematical criterion is allowed to change the underlying distribution of the sample in ways beyond simple inflation to achieve desired population counts. As shown by Alexander and Roebuck, constrained minimum distance techniques preserve the distribution within household type where household type is a vector with as many elements as there are cells in the person post-stratification. The sample distribution across household types, however, is altered to satisfy the minimum distance constraint. It is difficult to envision what impact such procedures will have in part because it is even more difficult to say what undercoverage model is implied by the weight changes.

My recommendation would be to use a variation of the current principal person weighting procedure until an alternative exists with improved results for a range of reporting domains and survey statistics. The Lemaitre and Dufour results suggest that marital status counts should be included in the post-stratification and that the geometric mean of the family member's weights should replace the use of a designated principal person.

My next recommendation is that better direct household weighting methodologies be explored. This includes energetic measures to obtain household-level data that can be used to develop improved nonresponse-adjusted household weights. Development of household projections for post-stratification is also included under this recommendation.

In my opinion, person level counts should be used for post-stratification as a last resort and only after having constructed the best possible household weight to use as input into the process. Before using the alternate procedures described in this session, the methods should be reevaluated to test whether they produce superior survey estimates. Except for the expenditure evaluation of Alexander and Roebuck, these papers have focused on the accurate estimation of demographic counts.

When means and proportions are of interest, the nonresponse-adjusted household weights may produce a better estimate than the methods investigated in these papers, in spite of its undercoverage. Investigations of alternate family weighting procedures should include the nonresponse-adjusted household weight (perhaps adjusted to sum to total households should such a count be available).

Ideally, investigations of alternate estimators should also include a measure of the truth. The approach of Lemaitre and Dufour could be used -- to compare survey data for a census year to census results. Some survey results will not be estimable from the census (e.g., expenditures for food). An adaptation would be to use the census results to create a very efficient post-stratification for the survey weights and then use the resultant survey estimates to evaluate the quality of the alternate post-stratification procedures.

6. References

Alexander, C.H. (1986). The Present Consumer Expenditure Surveys Weighting Method. Population Controls in Weighting Sample Units. Washington, DC: Bureau of Labor Statistics, U.S. Department of Labor.

Cox, B.G. and S.B. Cohen (1985). Methodological Issues for Health Care Surveys. New York: Marcel Dekker, Inc.