SAMPLING AND WEIGHTING ISSUES OVER THREE WAVES OF THE COMMUNITY TRACKING STUDY HOUSEHOLD SURVEY: USING A PARTIAL OVERLAP TO INCREASE EFFICIENCY

John W. Hall and Barbara Lepidus Carlson
Mathematica Policy Research, Princeton, New Jersey 08540

KEY WORDS: Community Tracking Survey; change estimates: partial overlap design

The household component of the Community Tracking Survey (CTS) recently finished its third wave of data collection. Interviews with over 20,000 households collected information on approximately 60,000 persons per wave. The household survey employs a clustered survey comprising 60 sites selected with probability proportional to size (PPS) methods to provide local and national estimates and an unclustered RDD survey to increase the precision of national estimates.[1] Two features of the clustered sample are larger (over 1,000 households) samples in 12 of 60 sites to allow for precise local estimates and an in-person supplement in those same 12 sites to increase frame coverage.

This paper discusses a limited number of issues. The overall design of the CTS sample and selection of the CTS sites are reviewed in Carlson, Strouse, and Hall (2002). More detailed information on sample selection, weighting and data collection may be found in Strouse, et al (1998) and Strouse, Carlson and Hall (2001). Metcalf, et al. (1996), present details of the original sample design including site selection.

Considerations in designing the second through fourth waves included the precision of estimates of change and survey costs. The design for the second and third waves included selection of a subsample of telephone numbers and addresses selected in the previous round. In the third wave, optimal allocation based on cost led to oversampling households interviewed at the time of the second round.

The remainder of this paper comprises a brief review of the sample selection and weighting strategies for each of the first three waves of the CTS, plans for selecting the household sample for the fourth wave, and a brief summary of the issues discussed.

First Wave. The first round of the household survey collected data on over 32,000 family insurance units (FIU)[2] and approximately 60,000 individuals. Ninety percent of the interviews were conducted in the site-based sample and the remainder in the unclustered sample. Ninety-eight percent of the sampled cases were identified using list-assisted RDD sampling techniques. A small sample of households were selected in the 12 high-intensity sites using an area probability approach. The area probability sample yielded data on with 635 FIUs and approximately 900 individuals.

Sites had been selected and the overall sample allocation determined prior to the first wave. The biggest remaining challenge in sample design was selecting the in-person sample. The purpose of the in-person sample is to increase coverage by including persons in households that would be excluded (no telephone service at all) or underrepresented (substantial interruptions in service) in the RDD survey. Only those households with recent significant interruptions in telephone service were eligible for the in-person sample; because of this, it was decided not to include areas where the likelihood of finding eligible households was very low. In addition the measure of size for selection of in-person interviewing areas was the estimated number of households with no telephone service.

Weights were calculated for the first wave of the household survey to allow for both national and site-specific estimates. Design-based weights were employed that adjusted for: differences in probabilities of selection, non-response, multiple telephones, and interruptions in telephone service. After these adjustments, post-stratification adjustments were made, the weights were trimmed, and the trimmed weights again post-stratified.

Four issues were of particular importance in weighting the data for the first round:

- combining the clustered and unclustered samples for making national estimates

[1]The unclustered sample will be dropped from the fourth wave.

[2]A family insurance unit (FIU) is a group of related individuals who would typically be covered by the same health insurance policy. See Strouse, Carlson and Hall (1998, 2002) for a complete definition.

- combining the in-person and RDD samples
- characterizing probabilities of selection given the division of the sites into high- and low-intensity
- using cases from the supplement sample to augment sample estimates for specific sites

Details of the methods used to address these issues may be found in Strouse, et al (1998). A summary follows.

The clustered and unclustered samples were treated as independent samples for national estimates, and were combined using a composite weight. The weight given to the clustered sample (λ) was determined by the ratio of the average effective sample size of the clustered sample to that of the sum of the effective sample sizes of the two samples; the weight given to the unclustered sample was ($1-\lambda$).[3]

Combining the in-person and site-based RDD samples was complicated by the fact that the samples were not independent; any household in a large MSA that had an interruption in telephone service lasting more than two weeks, but less than the entire field period, had a chance of being selected for either sample. Rather then attempting to incorporate the multiple chances of selection into the initial weights (the inverse of the probability of selection) CTS weighting procedures employed what was called the telephone service interruption factor (IAF). The IAF includes components that adjust (synthetically) for both multiplicity and length of interruption of telephone service.

First the IAFs were computed for both the in-person and RDD samples. Using the IAF as a component of a preliminary weight, the RDD sample was then post-stratified so that the weighted total for households with some interruption equaled Current Population Survey (CPS) estimates of households without telephones and those with no interruption equaled the estimated number with telephones. In strata that included the in-person sample, the weights of that sample were adjusted to the total number of estimated households without telephones. The two samples were then combined and weights of those with and without interruption post-stratified to (CPS) estimated totals of households without and with telephones, respectively.

Another issue, pertinent for national estimates made using the clustered sample, was how to calculate probabilities of selection, given the disproportionate allocation of sample to high-intensity sites (the average number of interviews in a high-intensity site is approximately four times as large as in the average low-intensity site). Because the 12 high-intensity sites comprise a random subsample of the 48 sites selected from large MSAs for making national estimates, the probability of selection within the site was calculated based on an expected allocation rather than the actual number of telephone numbers selected: the allocation to any of these 48 sites had a 1/4 chance of being that of a high-intensity site, and a 3/4 chance of being that of a low-intensity site.

The fourth weighting issue in the first wave, that of using interviews from the supplemental sample to augment site-specific weights arose because over half the interviews in the supplement are conducted in areas covered by one of the sites. Further, some very large sites (Los Angeles, New York City) are among the low-intensity sites; adding interviews from the supplement could substantially increase the number of cases available for analysis in these sites. To produce weights for such estimates, we calculated probabilities of selection within each site as if all numbers in exchanges assigned to the site had been selected for the site sample, even if initially selected for the supplement.

Second Wave. For the second wave, the overall design was retained and the sample size was nearly the same as in the first wave. However, in planning for the second wave of the household survey, a major concern was the precision of estimates of change. For estimates of change, overlapping samples (panel or partial panel designs) typically have greater precision than do independently selected samples. The CTS initial design had envisioned a partial panel, with substantial (45 to 50 percent) but not complete overlap between waves. While some consideration was given to selecting independent samples at the household level for the second wave (the sites would remain the same), two designs received the most consideration:

- following samples of individuals or households
- following samples of telephone numbers (RDD) and addresses (in-person)

[3]The average effective sample size for each sample is defined as the unweighted (nominal) sample size divided by the average design effect. The average design effects were based on those computed for a sample of 14 variables for the sample overall and for several subgroups. Separate calculations were made for the site and unclustered subpopulations.

CTS chose to follow samples of telephone numbers and addresses for a number reasons, two of which are discussed here. First the cost of following individuals or households over a two-year period between CTS waves would have been substantial. Second, if we sampled telephone numbers from the first wave, the actual overlap at the individual levels would be high, because the majority of the population keep the same address and telephone number over a two year period. Allowing for non-response, it was estimated that sampling 70 percent of the first wave telephone numbers would result in an overlap at the household level of 40-50 percent.

The sample design for the second wave included a subsample of 75 percent[4] of the telephone numbers sampled and released for calling in the first wave and all of the addresses contacted for the in-person sample. Except for a few hundred households who were adamant refusers at the time of wave one, all wave one telephone numbers were sampled at the same rate for wave two. The remainder of the RDD sample included telephone numbers from working banks not included in the frame for the first wave as well as numbers that had had a chance of selection (but were not actually sampled) for the first wave.

Weighting the second wave used the same components as did the first wave. However, because of the overlap with the first wave sample, calculating the probabilities of selection was more complex (some telephone numbers had two chances of selection for the second wave and some had only one).

We estimate that the design used for wave two reduced cost compared to either a design that would use independent cross-sectional samples or one that would follow individuals. However, preliminary analyses indicate that gains in precision were small.

Compared with using independent samples, the second wave design increased response rates, thus saving on costs of refusal conversion. The increase in response rates is due at least in part to the use of financial incentives in the household survey. Compared to a design that follows individuals, the approach used by CTS saved costs because effectively following individuals would require some combination of contacting persons between waves and locating individuals at the subsequent wave who had moved or changed their telephone numbers for other reasons.

Preliminary analyses, conducted using a limited number of variables, indicate that the wave two design did not significantly improve the precision of change estimates. Reasons for this finding, assuming it is correct, may be in a lower than expected population correlation over time, the limitation of estimation techniques, or the fact that much of the individual level correlation is captured at the site level. Further analyses are planned.

Third Wave. For wave three, a major change was made in the design of the RDD sample. For wave two, almost all of the numbers from wave one were sampled at a single rate. However, for wave three, three different rates were used to sample numbers that had been selected for the second wave. The basis of the difference in probabilities of selection was the wave two survey status: whether calling the sampled number resulted in an interview, a refusal, a non-working number, or a non-contact.

The experience from wave two was that the telephone numbers where wave one interviews were completed were more likely to yield an interview than was any other category. The wave two per interview cost for this group was thus relatively low. Conversely, telephone numbers where a refusal occurred in wave one yielded few interviews and the cost per wave two interview was quite high. Using the principals of optimum allocation based on cost, we thus estimated optimal sampling rates for each wave two response category. To reduce complexity in the design we grouped categories and set three sampling rates for wave two telephone numbers:

- all the telephone numbers where an interview was conducted at wave two were sampled
- numbers where no contact was made at wave two were subsampled at a rate of 67 percent
- numbers where a refusal of any kind had occurred at wave two, and numbers that were not household numbers at wave two were sampled at a rate of 33 percent

As in wave two, the wave three RDD sample also included numbers not selected for the second wave.

Weighting the wave three data followed the same procedures used in prior waves. The probabilities of selection for wave three were substantially more difficult to compute, and for much of the sample could only be approximated. The reasons for the complexity are two-fold. First, some telephone numbers could have been selected only for wave three, others for waves two and three, and some for all three waves. These factors had to be considered when calculating initial weight for the entire sample. Secondly, because sampling of the wave two telephone numbers was conditioned on the

[4]The proportion sampled was increased to 75 percent because of lower than expected response rates.

wave two survey result, calculating the probability of selection in many cases required estimating after the fact, the likelihood that a number would have resulted in a specific result in the prior round. For example, if a number was selected for the first time in wave three, a determination was first made whether it could have been selected for the wave two sample. Then, an estimate was made of the likelihood that calling the number in round two would have resulted in an interview, a refusal, or some other status. These estimated likelihoods were used to produce an "average" estimate of the number's probability of selection had it been part of the wave two sample.

Fourth Wave. Interviewing for wave four will begin in January 2003. At this time, the design calls for dropping the unclustered RDD supplement. While the supplement has increased the precision of national estimates, the increase was not as great as originally projected. Further, the precision of national estimates based on the clustered sample alone is very good.

A second design change will include updating the in-person sample. The original interviewing areas were selected using 1990 Census data. New areas will be selected, using data from Census 2000.

Summary. The CTS household survey to date includes three waves with interview data on approximately 60,000 person per wave. A fourth wave will begin in 2003. Data can be used for both national and site specific estimates, including estimates of change over time. The CTS has employed a partial panel design. This design has increased the efficiency of the sample, probably more by reducing costs than by increasing precision. Design decisions that have increased efficiency have, however, increased the complexity of the sample weighting process.

References

Metcalf, Charles E., Peter Kemper, Linda T. Kohn, Jeremy D. Pickreign. Site Definition and Sample Design for the Community Tracking Study. Technical Publication No. 01 Center For Studying Health Systems Change. October 1996

Carlson, Barbara Lepidus, Richard Strouse, and John Hall. The Community Tracking Study: Measuring Changes In The U. S. Health Care System Over Time. 2002 Proceedings of the American Statistical Association, Survey Research Methods Section [CD-ROM], Alexandria, VA: American Statistical Association.

Strouse, Richard, John Hall, Frank Potter, Barbara Lepidus Carlson, Peter Cunningham, Joanne Pascale, and Rita Stapulonis. "Report on Survey Methods for the Community Tracking Study's 1996-1997 Round One Household Survey." Mathematica Policy Research, 1998. Available at www.hschange.com as Technical Publication 15.

Strouse, Richard, Barbara Lepidus Carlson, and John Hall. "Report on Survey Methods for the Community Tracking Study's 1998-1999 Round Two Household Survey." Mathematica Policy Research, 2001. Available at www.hschange com as Technical Publication 34.