Estimation Issues for the Continuous Measurement Survey

Lynn Weidman, Charles Alexander, Gregg Diffendal, Susan Love

U.S. Census Bureau
Washington, DC 20233

The paper was presented at the Annual Meeting of the American Statistical Association (ASA), August 1995.

This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform interested parties of research and to encourage discussion.

I. Introduction

The Continuous Measurement Survey (CMS) is a proposed monthly survey whose content will be roughly equivalent to the decennial census long form or sample data. During each calendar year it will include sample housing units and special places from all block groups in the nation. Monthly sample size will be 400,000 for the first three years, then decrease to 250,000 in year four. The data collection for each month's sample will take place in three phases: self-response by mailout/mail return; computer-assisted telephone interviewing (CATI) for mail nonresponse addresses whose phone numbers can be obtained; and computer-assisted personal interviewing (CAPI) for a subsample of the remaining nonresponse units. Approximately one month would be allowed for each of these phases. Housing unit (HU) status will be defined as of the time of first response: either mail return, CATI contact or CAPI contact. Nonrespondents are defined as of the time of attempted CAPI interview.

Two separate sets of estimates will be available for geographic areas, depending upon population. For areas of at least 250,000 persons, annual estimates will be produced, and for smaller areas averages over several years will be estimated and updated yearly. Initially the averages will be over three years, then progressively lengthening out to five years when the sample size decreases. Because we view this survey as continually updating the characteristics of a changing population, we want to reflect the variability in the resident population of an area over the year. (As an example, the number of students in a city with a large university varies as a function of whether or not school is in session.)

The purpose of this paper is to present issues that arise from CM's design and must be addressed in determining its estimation procedures. In some instances we have made decisions about the methods we will initially put in place to produce estimates during the 1996 test. Since we do not yet have data from CM on which to base any subtle improvements, our initial methods will be simple in most respects and will be guided by experience with decennial census data. The paper describes possible modifications that will be considered and evaluated with CM data. The tentative decisions for the initial 1996 test estimates are indicated on bold type.

In the following section we present the main features of CM and how they differ from other Census Bureau household surveys. Section 3 includes a short description of each of the consequent issues, its potential impact on the estimates and decisions made for the 1996 CM operational test.

II. CM and the Basic Weighting Approach

There are several aspects of this survey that vary from those of the Census Bureau's traditional monthly household surveys. These include:

- Data collection by three modes.
- Data collection for a sample unit can occur at any time over a period of 3 months, rather than during a fixed shorter time span, like one month or one week. This introduces the possibility of defining a reference period as a function of time of interview rather than as a fixed time interval related to month of mailout.
- Making estimates for geography as small as block group.
- Estimating both housing unit and person characteristics.

The basic weighting approach for many of the Census Bureau's traditional household surveys includes three main steps:

- an initial weight of the inverse probability of selection;
- adjustment factors for responding units to account for nonresponding units;
- person level adjustment factors so that specified national or state age x race x sex estimates equal control totals.

There are additional steps for many of the surveys, but these are the steps they have in common. (e.g., U.S. Bureau of the Census, 1978.) CM will also follow this approach.

The structure of CM introduces a large number of weighting issues that need to be addressed when applying these steps. We separate them into the following major areas.

- Defining and capturing the target population.
- Defining the reference dates.
- Defining the noninterview adjustment procedure.
- Use of HU controls for small geographic areas.
- Obtaining population controls for all counties.
- Effects of weighting options on annual estimates.

- Combining annual estimates to form multiyear estimates.

III. Estimation Issues

Target Population

What people CM wants to include in its estimates for a particular area is the first thing that must be decided. Generally, this is the entire population living in the U.S. except for the transient homeless and crews of merchant vessels. The major issue within this population is where to count a person if he/she spends a substantial amount of time living in more than one location during the year. Do we define one of these locations as the place of residence (as the decennial census does) and only count the person if they fall in sample there, or do we use the one at which the person falls in sample? This decision determines whether an area has a stable population during the year or one with seasonal variation. Students who live during the school year in their college town and with their parents the rest of the year, and people who spend the winter months in a warm climate (e.g., Florida or Arizona) in a second home make up a large percentage of persons with multiple residences.

At this point we want to be able to capture the seasonal variation and have it reflected in the annual average. Thus for the 1996 Test we are counting people with multiple residences where they are currently staying if they are there for more than two months. Otherwise they are counted where they "usually" reside.

There are several other cases where a person does not have one fixed residence and it must be determined if he/she should be included as a resident of a specific sample housing unit. An important example is persons with no fixed residence, including: i) those who regularly stay at several places during each week or month; ii) those who do not have any regular or predictable pattern of residence. Conceptually our goal for such people is to establish a unique location where they "live" at each point in time during the year, and to average across the year counting them each day at the location where they live. Residence rules are discussed more thoroughly in Love et al(1995).

Once we know what people we want to include as members of a housing unit, we must develop a process which comes as close as possible to capturing this set during an interview. This is done through construction of a household roster by the (primary) respondent. For CM we must have both a self-rostering procedure for mail forms and an interviewer-assisted rostering procedure for CAI (computer assisted interviewing). It is possible to come much closer to the "ideal" roster for CAI than on a mail questionnaire. This is because CAI can provide the interviewer with probes and followup questions that can be asked of the respondent. We must count on respondents who complete the mail questionnaire to pay close attention to written instructions and to understand them. Only the most general instructions on who is to be considered a household member will be included in the rostering section of the questionnaire, to prevent confusing the respondent.

Our estimation procedures will inevitably proceed on the basis that the questionnaire has assigned a valid residential status to each person in sample as of the time that information is collected for that person. Estimates of the number and characteristics of the people in a particular area will be made by assigning weights to the people residing in that area and summing the weights. Designing the household roster questions to minimize undercounting and double counting of people is a major challenge. Our approach to this is discussed in Love, Dalzell, and Alexander (1995).

Estimates of the total number of housing units are not affected by rules for establishing residence, but estimates of the number of occupied and vacant units are. This is discussed at greater length in Love, Dalzell, and Alexander (1995).

Our current procedures for including the literally homeless population would be to include them during any portions of the year when they are staying in a "noninstitutional group quarters" (such as a shelter) and leave them out of scope when they are not in a "living quarters" that we can include in sample. We are not considering trying to sample street sites, encampments, or other nonresidential locations. However, we are following progress on service-based sampling methods that use "multiplicity weighting" to represent homeless people who use services taking into account the time when they are staying at nonresidential locations. (See Schindler, et al, 1994; National Institute on Drug Abuse, 1991.) If useable methods are developed, we will consider incorporating them into the CM design.

Reference Period

CM data collection will take place in 3 phases - mail, CATI and CAPI - over a period of 3 months. We considered two options for combining respondent data from these 3 phases for processing.

- Fixed reference period. The occupancy status of a housing unit and its list of residents are defined as of a specified day of the sample month, regardless of the response month. All time frames used in data collection are related to this fixed date.
- Floating reference period. The day on which the household residents or the occupancy status of a housing unit is determined is the reference day, regardless of the sample month. All time frames used in data collection are related to this reference day.

Using the fixed reference period would require at least one month of recall from CATI respondents and at least two months of recall from CAPI respondents to construct the household roster, plus whatever is needed for other questions. The floating reference period does not require the additional recall back to the sample month as a starting point for these respondents. In addition, CATI or CAPI respondents using the fixed reference period may not have lived in the sample unit in the reference month, so the unit would have to be treated as a nonrespondent, thereby causing a higher nonresponse rate for this option. On the other hand, use of the floating reference period requires the assumption that the samples from each month represent the same population and can introduce selection bias in each month's sample when this is not the case.

For the 1996 test, we decided to use the floating reference period. There were three main reasons for the decision.

- Comments from various advisory panels and potential data users seemed to be most concerned about the recall bias, among the sources of bias

just mentioned.
- With a goal of keeping the test questionnaire simple, we preferred not to add the complication of specifying reference periods and giving special instructions for respondents who had recently moved into the sample address.
- We have no specific suggestions for how we would measure the recall bias, where as we think we can get some rough bounds on the selection bias by measuring month-to-month variation in characteristics of units interviewed by the different modes.

We plan to look into how changes over time, both short and long term, in the differences between each mode's respondents might introduce bias into estimates because of the use of floating reference periods. This will be accomplished through determining possible types of changes and how they are treated differently when obtaining estimates from the two reference period options.

Noninterview Adjustment

A common approach for noninterview adjustment is the so-called "weighting class" adjustment (Kalton, 1981). Assign HUs to cells according to some set of characteristics that can be measured for nonresponse units. Then multiply the weights of the respondents in a given cell by the factor (total basic weight for all HUs in cell)/(total basic weight for all respondent HUs in cell). The cells must be collapsed so that there are a non-zero number of respondents in each cell; additional collapsing is usually done to put a limit on the noninterview factor, or to have some minimum total number of unweighted HUs in the cell.

If CM uses this approach, there are three main issues about how to carry out this adjustment.

- Which response mode cases are used to calculate the adjustment factors and to which response mode cases should they be applied?
- What characteristics will be available to define the cells?
- How should cases be combined geographically and over months for adjustment together?

The decennial census has used the weighting class of approach to separately adjust for nonresponding HUs and for nonresponding persons. Each adjustment is carried out at essentially the tract level, where the geographic entities are referred to as "weighting areas." The total weight in the denominator of each adjustment factor is the 100% census count for the area. It is feasible to adjust at this level of geography because of the large sample taken at one point in time over all block groups. Since CM spreads its similar sample over three or five years, use of some larger geographical areas may be required.

The household surveys, which do not include a mail response mode, use the weighting class approach and collect characteristics of nonrespondents, such as race, units in structure, and even household size, by observation or information from neighbors. Recall that in CM, nonresponse occupied units are only identified during the CAPI phase; all units eligible for CAPI did not respond by mail, and those contacted by CATI did not respond a second time. It is only for the CAPI sample that personal visits are required to obtain response and that characteristics of nonresponse HUs could be obtained. In addition, we know from decennial census data that the characteristics of HUs that respond by mail differ from those that must be followed up with personal visits. We must take this information into account if we are to design a weighting class adjustment for CM.

For CM, which shares aspects of surveys and census, should the noninterview adjustment factors only be calculated for and applied to the CAPI sample? Should both the CATI and CAPI sample be used, since both may differ from the mail respondents? Should mail respondents be included as part of the nonresponse adjustment, perhaps receiving a fraction of the factor that is applied to the CAPI respondents? As always for this adjustment, we must consider both the characteristics and the response probabilities of the respondents from the three phases and how they relate to those of the nonrespondents. For the 1996 test we will probably be applying the adjustment to both the CATI and CAPI respondents, in effect assuming that they are more similar to each other than to the mail respondents.

There is a relationship between the number of variables used to define adjustment cells and how cases are combined over time and geography for use in a given set of cells. The more variables used, the more combining is needed to reach minimum cell sizes. Since we will be making estimates down to the block group, it makes sense to try to keep the need to combine block groups to a minimum if response rate is more homogeneous within block groups than across them. Thus for 1996 we will not be using any variables in addition to geography and mail response status, will be combining cases over a year, and will be doing the least combining of block groups needed to reach minimal required cell sizes. What criteria should be used for deciding which block groups are to be combined? Should we use 1990 census characteristics when applying these criteria?

Giving geographic detail priority over interview month also makes sense as an initial strategy. We know block group is important because estimates are actually made at that level. The relative importance of interview month is likely to be smaller, although in areas with large seasonal populations, if response rate is seasonal, we may need to give more importance to interview month.

We must look at the impact these decisions have on the final weight of individual cases, especially in combination with the decision to use the month of first contact as the reference month. Since cases are combined over a year to perform noninterview adjustment, the total weight assigned to cases interviewed in a given month will vary with monthly mail response rate. Are we willing to accept this variation or do we need to modify the weighting process to obtain equal total weight per month?

There are other approaches that need to be considered for possible future use. One of these is to model response propensity for each case based on a small set of characteristics, then multiply the weight of each response case by the inverse of its propensity. This allows more variables to be taken into account than for the weighting class approach. How to assign a probability of non-response to mail respondents requires more discussion.

Use of MAF Address Counts in Weighting

Since we are making estimates for block groups and a master address file (MAF) is available with the "current" address counts for each block group,

the weighted estimate of total addresses should equal this count. (Alexander and Wetrogan, 1994.) However, after nonresponse adjustment is complete, the total weighted number of addresses in a block group will not be equal to the number on the "current" MAF. This is because the adjustment was not applied at the block group level and the sampling may have been done (depending upon timing of operations) using an earlier version of the MAF that had different counts.

We can apply the ratio adjustment factor (MAF count)/(total noninterview adjusted weight) to the weights by block group so that we obtain the "current" MAF counts. Again, this process gives rise to some questions.

- Do we apply this adjustment once for a full year's responses, for each month's responses, or something in between? Does it make sense to be on a different schedule than the annual noninterview adjustment?
- Could this adjustment serve to eliminate the monthly variation in total counts introduced by performing an annual noninterview adjustment?
- Should it be done by block group or some combination of them? Is some minimum number of sample HUs required for combining?
- If block groups need to be combined, what criteria should be used for defining the combination? e.g., minimize the sum of absolute differences between estimated and control HU counts over all constituent block groups or minimize the difference for the combined groups as a single unit.

For 1996 we expect that CM interviewers will be providing information to update the MAF when they see noticeable changes in HUs, such as conversion of a vacant building to apartments or new housing now occupied, during their CAPI visits. Our decision about the operational details of applying this adjustment depends on the answers to these questions, as we get more experience with actual MAF files.

Housing Unit Control Counts for Small Areas

A research question for the 1996 test will be to see if the "demographic methods" traditionally used to give intercensal estimates of housing and population (see reference) can be applied to produced reasonably accurate housing unit controls for block groups or other small areas. The general approach would be to update the number of housing units from the previous census, based on changes in the MAF since the census.

This could give an estimate of total housing (occupied + vacant). Possibly some crude models would be applied to estimate occupancy rates based on CM data from similar areas.

The Bureau's intercensal estimates staff will prepare such estimates. They will be evaluated by comparison with independent listings for blocks recently in sample for some other household surveys.

Population Controls for Counties

The adjustment to population control totals for specified demographic groups assigns weights to persons. Each person is put in the appropriate cell, defined by a combination of demographic characteristics, with his/her housing unit weight and these weights are summed. Then the weights are multiplied by (control total for cell)/(sum of weights for cell), with the result being the final person weight. The controls used are projections made based on the previous decennial census, updated with data on births, deaths, migration, etc.

In the 1996 Test we plan to perform this adjustment annually at the county level. These projections will be produced once a year and we will cumulate cases over a year in order to make the number of cases in each cell large enough so that the adjustment does not inflate the variance. Even then for small counties it may be necessary to collapse cells together or group counties to attain needed cell sizes.

Among the issues concerning these controls are the following.

- The controls by age, race and sex and for counties below a certain population are being produced for the first time by the Census Bureau for this test (Wetrogan, 1995). How well will this new methodology work?
- The projections are based on the decennial census which uses residence rules that differ from CM's. How do these differences in residence rules affect the comparison between the controls and CM estimates? Can we suitably adjust the controls for these differences?
- The projections represent an annual average population, but we are trying to measure seasonal fluctuations. What sources of information might be available for this?
- How can we feed CM information back into the process of obtaining county projections? Other than administrative records, CM provides the most current and detailed source of small area data for the entire country.

Use of Person Controls in Household Estimates

The population controls just described can help adjust for differential coverage bias by race and other person characteristics. Household surveys have employed several methods of using these person adjustments to attempt to correct household weights for this bias; the adjustment may also reduce the variance of some estimates. A traditional approach is the "principal person method" (U.S. Bureau of the Census, 1978). More recently, generalized least squares and raking-ratio-adjustment have been used to produce household weights that can be used as person weights for each person in the household while still satisfying the controls. (See Zieschang, 1990; Deville and Sarndal, 1992.) The appropriate method depends on the nature of the undercoverage, as discussed in Alexander (1990).

Not having experience with data from the full CM system, we will not make any such adjustments to household or family weights in the 1996 test. We will use separate weights for person estimates and housing unit estimates.

Annual Estimates for Large Areas

CM can produce annual estimates with a specified target coefficient of variation relative to the 1990 census for all geographical areas of at least 250,000 population. They are obtained by adding estimates for all constituent block groups derived from responses over the twelve months of the year.

- What effects do noninterview, HU count and population control adjustments have on the bias of these annual estimates?
- Are short term changes in characteristics accurately represented in these estimates?
- Is there other information available (e.g., administrative records) that can be used to improve these direct estimates?
- Can time series models be used to improve estimates through identification of monthly correlations and long term trends?

Multiyear Estimates for Small Areas

The sampling fraction for CM is such that data for areas of less than 250,000 must be cumulated over several years to obtain estimates of desired accuracy. In 1996 we will be able to make these estimates with one year of data by reducing the sampling interval. Since data are processed on a yearly basis, the multiyear estimates will be some type of average of the individual year estimates. Obviously the question to ask is how this average should be defined. There are many possible patterns of change over the long term, and we would like our average to be able to reflect them all when updating these estimates yearly. Of course, no one average can represent all patterns equally well, so we must establish criteria for determining the one that is most desirable.

Here are some of the factors we must consider in our research in this area. For the sake of this discussion assume that the averages are taken over five years.

- Equal weighting of the years gives minimum variance. The difference between successive yearly estimates of totals is just 1/3 of the difference between the latest year estimate and the estimate from three years age.
- Increasing the proportion of the weight for more recent years gives multiyear averages that are closer in expectation to the current characteristics.
- For unequal weighting of years, a successive years difference of averages represent a linear combination of annual estimates rather the difference between the year added and the year dropped.
- What patterns of change have highest priority for being well-represented?
- What about changes in HUs as opposed to changes in persons?
- Can time series models be used to improve the use of year-to-year correlations?
- Is there any consistency desired between the sum of multiyear estimates for components of a large area and the latest annual average for that large area?

We will be looking into these questions before deciding on the average to use for the 1996 test.

IV. Conclusion

There are clearly a large number of issues that must be considered and investigated before finalizing estimation procedures for the implementation of CM as a national survey. We are sure that we have not presented an exhaustive enumeration of them, but have included the most important. Our challenge is to design a research program to look at both the theoretical and practical aspects of all relevant issues. As decisions are made they must be sufficiently documented so that we can explain clearly both the procedures and their rationale.

V. References

Alexander, C.H. (1991). Incorporating Person Estimates in Household Weighting Using Various Models for Coverage." Proceedings of the 1990 Census Bureau Annual Research Conference, pp. 445-462.

Alexander, C.H. (1993). "Method to Determine Sample Size for the Intercensal Long Form Survey." Internal Census Bureau Report #CM-8, June 8, 1993.

Alexander, C.H and S.I. Wetrogan (1994). "Small Area Estimation with Continuous Measurement: What We Have and What We Want." Proceedings of the 1994 Census Bureau Annual Research Conference, pp. 359-381.

Deville, J.C. and C.E. Sarndal (1992). "Calibration Estimates in Survey Sampling." Journal of the American Statistical Association, 87, pp. 376-382.

Edmonston, B. and C.L. Schultze, eds. (1995). Modernizing the U.S. Census. National Academy Press, Washington, DC.

Herriot, R., D.B. Bateman and W.F. McCarthy (1989). "The Decade Census Program -- New Approach for Meeting the Nation's Needs for Sub-National Data." Proceedings of the American Statistical Association Social Statistics Section, pp. 351-355.

Kalton, G. (1981). Compensating for Missing Survey Data. University of Michigan, Ann Arbor, MI.

Kish, L. (1990). "Rolling Samples and Censuses." Survey Methodology, 16, 1, pp. 63-79.

Love, S., C. Alexander and D. Dalzell (1995). "Constructing a Major Survey -- Operational Plans and Issues for Continuous Measurement." Presented at the 1995 Annual Meeting of the American Statistical Association.

National Institute on Drug Abuse (1991). Prevalence of Drug Use in the Washington, DC. Metropolitan Area Homeless and Transient Population: 1991. Technical Report No. 2, Appendix A.5.3.1.

Navarro, A. and R.A. Griffin (1990). "Sample Designs for the 1990 Decennial Census." Proceedings of the American Statistical Association Survey Research Methods Section, pp. 176-722.

Schindler, E., R. Griffin and A. Navarro. (1994). "Sampling and Estimation for the Homeless Population. Proceedings of the American Statistical Association Survey Research Methods Section, pp. 468-473.

Steffey, D.L. and N.M. Bradburn, eds. (1994). Counting People in the Information Age. National Academy Press, Washington, DC.

U.S. Bureau of the Census (1978). The Current Population Survey: Design and Methodology. Technical Paper No. 40, U.S. Department of Commerce.

Weidman, L. (1995). "Use of Month of Interview as the Reference Month for the 1996 Continuous Measurement Test." Internal Census Bureau Memorandum for the Record, June 5, 1995.

Wetrogan, S.I. (1995). "Administrative Records and Continuous Measurement -- Role of Population Division." Internal Census Bureau Memorandum to Larry McGinn, June 16, 1995.

Zieschang, K.D. (1990). "Sample Weighting Methods and Estimation of Totals in the Consumer Expenditure Survey." Journal of the American Statistical Association, 85, pp. 986-1001.