SAMPLING AND STATISTICAL ESTIMATION IN THE DECENNIAL CENSUS

Robert M. Bell, RAND
P.O. Box 2138, Santa Monica, Ca. 90407-2138

Key words: nonresponse follow-up, CensusPlus, Post Enumeration Survey

Abstract: Methods for counting the population in the 2000 census may differ substantially from those used in the 1990 census. In a 1995 census test, the Census Bureau plans to evaluate sampling for nonresponse follow-up–the most expensive and error-prone part of field operations. They will also test CensusPlus, a new form of integrated coverage measurement designed to reduce the differential undercount. Both innovations will rely critically on statistical estimation. Choices about whether and how to implement these changes will involve several tradeoffs: e.g., cost savings vs. accuracy; variance vs. bias; and accuracy of counts for large vs. small areas.

INTRODUCTION[1]

This paper discusses two major innovations that the Census Bureau is considering for producing counts in the 2000 census: sampling for nonresponse follow-up and integrated coverage measurement. These innovations respond to the two main criticisms of the 1990 census: that costs grew out of control and that there was differential coverage among demographic groups and geographic areas.

Instead of following up all housing units that failed to respond during mailout-mailback operations, *sampling for nonresponse follow-up* involves trying to enumerate a sample of such housing units (most likely between 10 and 33 percent). Data from the nonresponse follow-up (NRFU) sample would allow estimation of counts and characteristics of mailback nonrespondents who are not sampled.

The second proposal, called *integrated coverage measurement* (ICM), is designed to measure and correct the differential undercount. During the 1995 census test, the Census Bureau plans to evaluate a new ICM method, CensusPlus, designed to overcome drawbacks of the Post Enumeration Survey used to evaluate, but not correct, the undercount in 1990.

Criticisms of the 1990 Census

The 1990 census was substantially more expensive than previous censuses, even after accounting for inflation and population growth. In constant 1990 dollars, the enumeration cost per housing unit has increased from $11 in 1970, to $20 in 1980, to $25 in 1990. The largest single part of the expense was follow-up of housing units that had not responded during the mailout-mailback portion of the census. Estimates of the total cost of NRFU operations in the 1990 census range from $490 million to $560 million, roughly 20 percent of the $2.6 billion ten-year cycle cost of the census (Bureau of the Census, 1992; U.S. GAO, 1992). Each 1 percent of nonresponse to the mailed questionnaire is estimated to have added as much as $17 million to the cost of the census. Much of the problem in 1990 resulted from a steep decline in the mailback response rate–from 75 percent in 1980 to just 65 percent in 1990.

The high nonresponse rate may have had as much impact on data quality as on costs. The resulting delay in completion of nonresponse follow-up pushed back the beginning of coverage measurement by the Post Enumeration Survey (PES). A long delay between census day and the beginning of coverage measurement compromised the ability of the PES to operate accurately and was one of several factors that prevented the Census Bureau from incorporating the PES results into official counts released by the legal deadlines. Also, the latter stages of census operations suffered degradation of data quality. Ericksen et al. (1991) report that, for the 1990 census, the rate of erroneous enumeration on mailout-mailback was 3.1 percent. On nonresponse follow-up, the rate was 11.3 percent; on field follow-up, the rate was 19.4 percent.

The Census Bureau has used Demographic Analysis to evaluate coverage of census enumerations since 1950 (Coale, 1955; Himes and Clogg, 1992). Demographic analysis combines data from previous censuses, vital statistics on births and deaths, and other

[1]This paper has been informed by discussions and research undertaken by the Panel to Evaluate Alternative Census Methods, Committee on National Statistics, Commission on Behavioral and Social Sciences and Education, National Research Council. The author wishes to thank the study director Duane Steffey, the panel chairman Norman Bradburn, and the other panel members, particularly Cliff Clogg and Alan Zaslavsky, the other members of the sampling and statistical estimation working group. The views expressed here do not necessarily reflect the opinions or recommendations of the Panel or any of its individual members other than the author.

administrative records, such as Medicare data, to obtain national population estimates by age, race, ethnicity, and sex. In recent censuses, dual-systems estimation has used data collected for a stratified sample of households to evaluate coverage. In short, persons "captured" in the survey are matched against the census enumeration in order to estimate the fraction of the population that was included in the census. Similarly, a sample of people enumerated in the census is followed up to determine whether these people should in fact have been included or whether they were erroneously enumerated. Dual-systems estimation allows estimation of coverage–undercount or overcount–by combinations of demographic group, geographic area, and other variables available on the census form (such as owner/renter status).

All previous coverage evaluation programs have demonstrated the existence of an overall undercount–more omissions than erroneous enumerations. More important, they have found that there is a differential undercount–i.e., that certain groups, such as African American males, and certain areas, such as central cities, are systematically undercounted relative to the rest of the population during the main census enumeration. Despite improved overall census coverage, the African American/white coverage differentials have remained remarkably constant at about 4 percentage points since the 1950 census.

In 1990, dual-systems estimation based on the Post Enumeration Survey, estimated a net undercount of about 1.6 percent, not far from the demographic analysis estimate of 1.8 percent. The PES estimated differential undercounts of 3 to 4 percentage points for African Americans, Hispanics, and Native Americans and substantial differential coverage related to other demographics and to geography.

Looking Forward to the 2000 Census

Among many changes that the Census Bureau is considering for 2000, some of the most promising are methods aimed at improving mailback response: a prenotice letter, a respondent-friendly questionnaire design, a reminder postcard, a replacement questionnaire, and a statement that response is mandatory. Experimental tests of these innovations have produced combined effects on mailback response of 23 percentage points or more under non-census conditions (Dillman, Clark, and Treat, 1994), but there is no way to know how large an impact they will have on response in the 2000 census. Considering that trends thought to have contributed to the declining response rate–increasing diversity of the population and growing distrust of government–are continuing, response-improvement measures may serve mainly to reverse the trend toward even lower response rates (Committee on National Statistics, 1994). Even if the overall mailback response rate improves substantially, the assessments of response improvement methods do not suggest that they will make any headway on reducing differential coverage. Other programs that target special, hard-to-enumerate subpopulations might reduce differential coverage to some extent, but it is also unlikely that these innovations will close the gap substantially. Thus, the problems that plagued the 1990 census still loom large for the 2000 census.

SAMPLING FOR NONRESPONSE FOLLOW-UP

Plans for the 1995 Census Test

The Census Bureau has made evaluation of sampling for nonresponse follow-up a major component of the 1995 census test. Households that do not respond to the mail questionnaire by 6 weeks after the initial mailout (14 days after mailing of a replacement questionnaire) will be considered mailback nonrespondents, and one-third of these households will be sampled for NRFU. Current plans call for the collection of only short-form data during NRFU. No attempt will be made to obtain information from the other two-thirds of mailback nonresponding households. An attempt will be made to identify vacant housing units before selection of the nonresponse sample. Interviewers will visit units for which a postmaster returned the prenotice to the first mailing. Confirmed vacancies will not be included in the NRFU sample.

A major goal of testing NRFU sampling in the 1995 census test is to learn more about the relative merits of sampling individual housing units (a unit sample)[2] versus whole blocks (a block sample). Non-ICM blocks in the four test sites will be evenly split at random between the two types of samples. Within blocks in the unit sample, the Census Bureau will sample 33 percent (one-third) of nonresponding housing units. In the other non-ICM blocks, they will use block sampling. That is, all mailback nonrespondents will be followed up in one-third of the block-sample blocks, while no NRFU activities will be conducted in the remainder of the block-sample blocks. Complete nonresponse follow-up will be conducted in all ICM blocks.

I will address two major questions about sampling for NRFU in the 2000 census:

1. Should sampling for nonresponse follow-up be used in 2000?
2. Is a unit or block sample better?

[2]Census Bureau documents refer to this as a case sample design.

Advantages of Sampling for Nonresponse Follow-up

The Census Bureau's cost models estimated very large cost savings from sampling for NRFU. For NRFU sampling rates of 50 percent down to 10 percent, they estimated cost savings compared with the 1990 ten-year cycle costs ranging from approximately $300 million to $750 million, even after increasing the sample size for ICM measurement (Bureau of the Census, 1993).

Sampling for NRFU would also offer timing benefits compared with complete NRFU. A reduced workload could move up the schedule for initial and repeat attempts to contact residents, which might reduce the number of errors in the latter stages of follow-up. Because one of the potential problems with CensusPlus is difficulty with retrospective identification of census day residency, moving up the completion of NRFU could also be an important benefit for the quality of ICM. Earlier completion of ICM would also make it easier for the Census Bureau to produce final counts in time to meet legal deadlines.

Disadvantages of NRFU Sampling

In contrast to the cost and operational advantages, sampling has negative implications for the precision of counts and other results, especially for small areas. Counts and attributes of persons in nonsampled nonresponding housing units would need to be estimated, producing sampling variability roughly proportional to the number of cases being estimated (although the exact relationship would depend on the sample design and estimation method). As results are aggregated to larger geographic areas, errors would diminish in size relative to the population of the area.

Another concern that has been raised about the use of NRFU sampling is that publicity about it may reduce the mailback response rate. If sampling is used in the 2000 census, that fact would certainly become public knowledge, which might dilute any positive effect that the mandatory nature of the census has on the mailback response rate. Although I doubt that the public takes enough interest in the census for this phenomenon to produce a practically significant problem, the possibility should not be ignored. Unfortunately, there is no way to learn from census tests what impact such reactions might have.

Should NRFU Sampling be Used in 2000?

Whether to use sampling for NRFU in the 2000 census is mainly a policy decision as to whether the expected cost saving due to sampling outweighs the likely decrease in the accuracy of counts and other data, particularly for small areas. The 1995 census test will provide valuable data to help inform that decision, more recent input to the NRFU pieces of the Census Bureau's cost model and data on the relationship between NRFU and ICM. In particular, it will be important to identify all fixed components of the cost of NRFU sampling, in order to obtain accurate estimates of the cost savings during the 2000 census. However, the most complete information about the impact of sampling for NRFU on the accuracy of the census is still to be gained from additional simulations with 1990 data.

In the end, whether to sample for nonresponse follow-up is likely to come down to the question: How accurate does the 2000 census need to be for small areas? Although that is more a policy question than a statistical one, I offer a pair of comments. First, counts and other tabulations are needed at the block level primarily to allow flexibility for redistricting and for aggregating results to various political jurisdictions and other territories. Thus, the success of the 2000 census should be measured by the accuracy of these aggregate statistics rather than by the accuracy of block-level data.

Second, sampling variability is not the only source of error in census results. Incomplete counts and erroneous enumerations occur during both the mailback stage and even a 100 percent NRFU operation. While sampling for NRFU would certainly contribute the largest part of the error in block- and tract-level data, sampling error may be small compared with systematic error for larger levels of aggregation. If NRFU sampling frees resources for steps to reduce other sources of error in the final results, it may produce a more accurate census–at least for larger geographic areas.

As both comments indicate, deciding whether to use NRFU sampling will involve not only weighing the importance of accuracy versus cost savings but also the relative importance of accuracy for small geographic areas versus larger ones.

Is a Unit or Block Sample Better?

The choice between a unit and block sample for NRFU involves mainly a tradeoff between greater statistical efficiency for a unit sample versus operational and cost advantages for a block sample. In addition, a block sample would be easier to combine with integrated coverage measurement.

Sampling for NRFU necessitates estimating the attributes of nonsampled housing units in a block from the information obtained about responding units (mailback or sampled during NRFU) in that block and in nearby blocks that are demographically similar. It is likely that a unit sample would generally produce more accurate estimates than a block sample of the same size

because of within-block correlation in household size and other attributes of mailback nonresponse housing units, even within carefully selected strata.

Suppose, for illustration only, that information from sampled housing units in a 100-block area (roughly 1,000 nonresponding housing units) is used to estimate the characteristics of nonsampled mailback nonrespondents in the same blocks. To the extent that there is within-block correlation in the 100 blocks, data on a sample of nonrespondents spread evenly among the 100 blocks would be more valuable than data from the same number of housing units concentrated in a smaller number of blocks, by a ratio known as the design effect (Kish, 1965). A unit sample would also provide the opportunity to use information from sampled mailback nonrespondents in the same block to improve the estimates for nonsampled housing units in that block.

Certainly, heterogeneity among blocks is to be expected for characteristics such as race and ethnicity. However, the critical quantities to estimate may be *differences* among race-ethnicity/age groups in mailback response rates, which may be relatively homogeneous among blocks. Initial Census Bureau simulations with 1990 census data have found advantages to both unit and block sampling under various circumstances (Fuller, Isaki, and Tsay, 1994), but further investigation is needed to separate the possible effect of the estimation procedures from the design. Also, these simulations have been limited to a few district offices. More comprehensive simulations with more fully developed estimators are needed to pin down the size of the unit sample advantage.

Another apparent advantage of unit sampling is that it would spread imprecision due to sampling and estimation among all blocks, thereby reducing the maximum amount of block-level error. On the other hand, because block sampling would eliminate the need for estimation in some other blocks, the two methods would not differ in the total number of housing units where estimation is needed. Consequently, unit sampling would not necessarily improve the accuracy of aggregate estimates beyond that due to within-block correlation.

In contrast, block sampling appears to offer certain operational advantages. Enumerators would need to spend less time traveling between blocks, and they might also be able to use their time in each block more effectively. For example, while visiting a complete sample of mailback nonrespondents in a block, enumerators might frequently observe occupants entering or leaving other units on the NRFU list. With a unit sample instead, enumerators might tend to finish and proceed to the next block too quickly for that to occur. Based on very preliminary assumptions, the Census Bureau has estimated that a block sample would save from $14 million (for a 10 percent sample) to $42 million (for a 50 percent sample) more than the corresponding amounts saved by a unit sample of the same size. Therefore, it is not obvious in advance whether the unit or block sample is more efficient in terms of accuracy for samples of *equal cost*. Operational data from the 1995 census test should allow the Census Bureau to estimate the relative cost advantage more accurately.

Block sampling would fit better with any likely method of ICM because 100-percent NRFU would be required in the ICM blocks (and, perhaps, in surrounding blocks). Complete NRFU is needed so that the block total from the ICM operation can be validly compared with the total from preceding census operations. Thus, even if unit sampling is the primary strategy for NRFU, it may need to be mixed with some block sampling for ICM purposes.

A related consideration is whether the choice of sampling design affects coverage in NRFU housing units. For example, with the more concentrated effort involved in following up a block sample, enumerators might be more likely to discover housing units that had been omitted from the frame (e.g., garage apartments). And if they do, it will be easier to use the results because such housing units will automatically be part of a block sample. Enumerators may also be able to collect better proxy information for difficult-to-complete cases under block sampling.

The Census Bureau plans to perform statistical tests for whether the average household size differs systematically between unit and block sampling in the 1995 census test. However, the size and design of the planned test are such that it could easily miss a coverage difference of 0.05 persons per housing unit (about 2 percent of persons in sampled units) between the block-sampling and unit-sampling design (Bureau of the Census, 1994); a difference of this magnitude would be important to the decision on which sampling plan to use. If coverage differs under block sampling and unit sampling, then the viability of unit sampling for NRFU operations would be compromised, because ICM would measure coverage for block-sample NRFU and there would not be an adequate corresponding measure for unit-sample NRFU. Consequently, the Census Bureau should investigate other ways to compare the validity of the two methods, such as comparing the numbers of added housing units.

Based on the last concern, I think that block sampling is likely to be the better choice unless simulation results turn up a substantial efficiency advantage for unit sampling.

Due to space limitations, this paper does not consider several other important questions related to sampling for nonresponse follow-up (Committee on National Statistics, 1994):

- What proportion of units or blocks should be sampled?
- Should the sampling probability be uniform across blocks (for a unit sample) or across areas (for a block sample)?
- How should the Census Bureau treat mail returns received after the beginning of NRFU?
- Should any nonresponse follow-up operations be conducted for all households before sampling for nonresponse follow-up?

INTEGRATED COVERAGE MEASUREMENT

The 1990 Post Enumeration Survey

The 1990 Post Enumeration Survey (PES) was designed not only to measure coverage of the census enumeration, but also to allow adjusting the 1990 census counts if it was judged that PES data could be used to improve their accuracy. The PES consisted of two surveys conducted in identical samples of 5,300 block clusters (165,000 housing units), one to measure undercoverage and one to measure erroneous enumerations. Methods were developed for adjusting census data for all subnational geographic units and for demographic groups (Hogan, 1993; Mulry and Spencer, 1993; Belin, Diffendal, Mack, Rubin, Shafer, and Zaslavsky, 1993). In 1991, the Secretary of Commerce not to carry out the adjustment, a decisions that was originally upheld in federal court, but reversed on appeal (U. S. Department of Commerce, 1991; Fienberg, 1992; Bryant, 1993).

Current plans for the 2000 Census are predicated on the use of integrated coverage measurement as an essential part of census taking, not just an evaluation of other census operations. Therefore, ICM is not regarded as a method of producing a "second set" of population estimates that competes with population estimates obtained without the use of ICM. Instead, ICM would integrate the use of samples, statistical estimation based on these samples, and statistical modeling with the other census-taking operations. This new use of ICM as an essential component of census taking defines the one-number census concept.

Five main concerns have been expressed about the 1990 PES and the resulting estimates (e.g., Freedman et al., 1993, 1994):

1. *Correlation bias.* Dual systems estimation assumes statistical independence between "capture" in the census enumeration and "recapture" in the PES samples within each post stratum. If some people are harder to count than are others in the same post stratum (i.e., the capture and recapture probabilities are correlated), that fact leads to estimated counts that are biased downward.

2. *Processing or nonsampling errors.* Certain errors, such as errors in matching people between the main enumeration and the PES, can bias the estimated counts.

3. *Failure of the model used to smooth counts.* Critics question two assumptions–that the PES estimates true adjustment factors unbiasedly and that true adjustment factors can be described by a linear combination of stratum characteristics (including interactions).

4. *Sampling error.* The size of sampling error depended mainly on the size of the PES sample.

5. *Inability to produce counts by the legal deadline.* Under current law, population totals for states are due by December 31 of the census year. Counts for blocks by age, race, and Hispanic origin must be available by March 31 of the following year for use in legislative redistricting. Due to the late initiation of PES field operation and the complexity of matching names, the 1990 PES missed those deadlines by several months.

CensusPlus in the 1995 Census Test

The Census Bureau has decided to evaluate in the 1995 census test an ICM method called *CensusPlus,* which is designed to overcome some of the concerns about the PES. CensusPlus uses intensive enumeration methods and highly trained interviewers with the objective of obtaining a complete enumeration of the true population in the blocks sampled. As with the PES, regular census operations–including precensus notification, mailout of census forms, and NRFU–also take place in the blocks sampled for ICM.

The assumption of complete coverage replaces the independence assumption implicit in use of the DSE after a PES, although it would still be important that the CensusPlus operations be conducted in a way that does not make the sample blocks atypical with respect to the conduct of primary census operations. The ICM enumeration involves adding people found in ICM who were omitted from the census and deleting people who were included in the census but found by ICM to have been erroneously enumerated. Then the coverage rate for each group may be estimated as the ratio of the count obtained by pre-ICM operations in sample blocks to the corresponding count after completion of ICM.

The logic of the method can be illustrated as follows. We compare CensusPlus counts for a given group (say, Black males aged 20-34 in urban areas in the Northeast who rent instead of own their home) in

the sample blocks with the estimated number of such persons obtained from the census enumeration, assignment, and NRFU for those blocks. The ratio of these numbers measures coverage or gives a factor that can be used to prorate estimates from non-ICM blocks similar to the sample blocks. CensusPlus will be tested in the 1995 Test Census.

CensusPlus has been designed to distinguish errors associated with errors in Master Address File (MAF) from mis-enumeration within housing units. Early in the year, prior to the census mailout, interviewers canvass the ICM sample blocks to construct an independent listing of housing units (and addresses). This list is then matched to the Master Address File (MAF), the frame for enumeration and NRFU in non-ICM census operations, generating two lists: housing units that were found by the ICM canvass but missed in the MAF and housing units that were included in the MAF. The two lists of units are followed up in the housing unit coverage and within-housing-unit coverage portions of ICM, respectively.

The housing unit coverage operation is designed to check the completeness of the MAF and estimate (1) the number of housing units that were omitted from the MAF (and therefore from the frame for mailout and NRFU) and (2) the number of persons omitted because they were in these housing units.

The original design called for running CensusPlus just behind the regular census operations in those blocks, to facilitate identification of residency on census day and to improve the ability to meet legal reporting deadlines. Housing units in the within-housing-unit sample would have been followed up as their census returns came in, whether by mailback, from NRFU interviews, or from unaddressed questionnaires. However, concerns about contamination of the primary census operations (see below) and implementation difficulties led to a redesign where CensusPlus reinterviews will begin after completion of nonresponse follow-up.

Computer-assisted personal interviews (CAPI) will be used to collect information during the CensusPlus enumeration. Each household that had responded during the primary census enumeration will be given a two-part reinterview by an ICM interviewer. First, the interviewer will construct a roster of persons living in the household using more detailed and probing questions than on the regular census form. Then the computer will reveal to the interviewer the roster from the original census response, showing discrepancies from the reinterview. In the "reconciliation" phase of the reinterview, the interviewer will attempt to resolve these discrepancies in order to come up with an accurate roster using information from both the original response and the reinterview.

Some housing units will be resolved as vacant by NRFU; these will be rechecked by ICM interviewers in order to verify that they are in fact vacant, or to conduct an interview in order to obtain information on the household living there when they are not vacant. Conversely, ICM interviewers may determine that some households enumerated by mailback or NRFU were erroneously enumerated and should be removed from the roster. The end product of these operations is a *resolved roster* of both housing units and persons in the ICM sample blocks, from which resolved counts of units and of persons by age, sex, race/ethnicity and other variables would be calculated.

Although the redesigned CensusPlus methodology shares many features with the PES, there are two important distinguishing features. First, the ICM interview will not be independent of other census information, because names from the previous response will be available for matching and reconciliation on the spot. This should reduce the need for an additional contact to resolve discrepancies as compared with the 1990 PES.

Second, the ICM interview will be directed at establishing an accurate roster for Census Day. In contrast, the 1990 PES was defined to include the people at the sample address at the time of the PES, which could be different due to people moving in and out of the sample block in the intervening months. CensusPlus will need the capability to follow up people who move out of the sample blocks after Census Day, which may be facilitated by the shorter interval between Census Day and the CensusPlus interview compared with 1990.

Issues for Evaluation of CensusPlus Methodology

The CensusPlus procedures proposed for the 1995 census test have some very attractive new features. I commend the Census Bureau for developing this innovative design. At the same time, two critical issues about CensusPlus methodology must be evaluated in 1995 before it should be adopted for use in 2000:

1. Can CensusPlus be conducted without affecting the results of the regular enumeration in the CensusPlus sample blocks?
2. Can CensusPlus attain adequate coverage in the sample blocks?

The answers to these two questions will determine the accuracy of the inputs to the denominator and numerator, respectively, of the ratio estimator, and consequently the validity of final population estimates.

Can CensusPlus Avoid Contaminating the Regular Enumeration?

The census coverage rates measured in the ICM sample blocks can only be regarded as valid estimates of the coverage rates in other blocks if the conduct of the regular census is essentially indistinguishable in the ICM sample and nonsample blocks. Because the original Census Plus design had called for overlap in time of the reinterviews with the regular enumeration process, several forms of contamination were possible: e.g., if residents or NRFU interviewers became aware of ICM interviewers in the CensusPlus sample blocks or if CensusPlus interviewers inadvertently approached some housing units housing units that had not yet responded to the regular enumeration (Committee on National Statistics. This type of contamination would bias CensusPlus estimates, because coverage rates measured in sample blocks would differ systematically from coverage rates in other blocks.

A major benefit of the CensusPlus redesign is to avoid these types of contamination. However, one form of contamination remains a concern. The precensus canvass for housing units conducted in ICM blocks may affect awareness of the census and consequently response to the regular census in those blocks–particularly if census personnel knock on doors to verify the existence of housing units.

Can CensusPlus Attain Adequate Coverage in the Sample Blocks?

The proposed ratio estimator, described above, is based on the assumption that the resolved roster in ICM blocks can be treated as the truth for those blocks. First and foremost, there is the problem that many individuals in our society are difficult to count. Comparisons of 1990 coverage measurement results to demographic analysis suggested that at least in some groups, a substantial number of people were missed by both the PES and the regular census, and that the number of these people was underestimated by dual-systems estimation (Bell, 1993). CensusPlus may be no more successful at finding the very toughest households and individuals; hence, the resolved roster will probably be incomplete.

We see several other challenges to this coverage assumption that apply more specifically to CensusPlus.

- Can the CensusPlus methodology identify erroneous enumerations? In particular, one form of erroneous enumeration is duplication, i.e., the listing of one person in more than one place. Detection of duplications must involve at least some searching in blocks adjacent to the blocks that were sampled.
- Is it possible to resolve place-of-residence ambiguities, for example, when true residence is "close" to a block included in a given CensusPlus sample block, or when a person could plausibly be regarded as resident at any of several addresses?
- Can the ICM instrument find the people who lived in the sample blocks on Census Day, even if they have moved since then? And can it distinguish them from people who moved in after Census Day?

The CensusPlus redesign will tend to exacerbate these problems because it will allow more people and whole households to move between the regular census response and the CensusPlus reinterview. However, I believe that the reduced risk of contamination will outweigh this drawback of the redesign.

Creative thinking may be needed to adequately evaluate the coverage of the resolved CensusPlus roster. One method would be to find a third source of names and perform "triple system estimation" (Marks, Seltzer, and Krotki, 1974; Zaslavsky and Wolfgang, 1993) to evaluate the number missed by both of the original lists. Possible sources would be an administrative list that was not used either in the original enumeration or in constructing the ICM roster, or a list from a particularly intensive form of enumeration such as observation by a resident ethnographer.

It would be useful to conduct some experiments to evaluate the effect of the design of the ICM reinterview. For example, the reconciliation phase for some fraction of cases could be carried out by experts different from the original ICM interviewers, and the results compared with those obtained in similar households when the interview and reconciliation are carried out in the same session. A careful study of these dynamics under cognitive laboratory conditions may also be helpful.

It may be possible to deliberately "salt" some of the data with information that is incorrect (enumerations or deliberate omissions) but plausible, to measure the success of the ICM reinterview in detecting and correcting these cases.

What if CensusPlus Fails in the 1995 Census Test?

Despite the positive steps that the Census Bureau is planning to improve the main enumeration, the differential coverage of certain groups and areas is unlikely to become less in 2000. Thus, the 2000 census must include integrated coverage measurement as an integral part of the process. Although no ICM method is likely to eliminate the differential undercount, properly-implemented ICM should still improve substantially on the results from the regular enumeration. Thus, any method should be judged on how well it compares with the regular enumeration, not with some

unattainable ideal.

If the 1995 census test uncovers problems that invalidate the tested version of CensusPlus, I believe that there are two reasonable options. First, it may be possible to address those problems through changes in the details of CensusPlus operations. Even if unsolvable problems are identified, CensusPlus should not be dismissed out of hand unless the problems are major.

What if CensusPlus cannot be fixed? Then, I believe that with modifications, dual-systems estimation based on a PES would work very well in the 2000 census. The assumptions and estimation methods used with this method have been subjected to much scrutiny. With the lessons that have been learned, many of the operational and estimation concerns can easily be overcome. Timing could become much less of an issue if NRFU sampling is used in 2000 and if Census Day is moved to the beginning of March. Although concern about correlation bias will certainly remain, that concern should not dictate against use of the PES. The most likely consequence of correlation bias would simply be that the PES would not adjust enough, although it would be a step in the right direction. Finally, it is imperative that the ICM sample size be increased by at least a factor of 2 to reduce concerns about sampling error. That concern would apply to CensusPlus as well.

COMMENTS

The Census Bureau has taken seriously the criticisms leveled at the 1990 census and is considering some of the most important innovations in the history of the U.S. Census. I believe that the 1995 census test will provide much of the information necessary to decide what changes to make in 2000. At the same time, there is much still to learn from 1990 data–especially in relation to the impact of NRFU sampling and ICM on the accuracy of census counts at various levels. It is critical for the Census Bureau to learn more from that source.

All four questions that I raised here involve tradeoffs: e.g., cost savings vs. accuracy; variance vs. bias; accuracy for counts of large vs. small areas; statistical efficiency vs. ease of implementation; and accuracy vs. political, public, and judicial acceptance. Of the four questions, I feel that three depend mainly on getting the right technical input–statistical, economic, and operational. In contrast, whether to use sampling for nonresponse follow-up comes down to a policy question. Even there, however, statisticians need to frame the issues correctly so that resources are not wasted on inconsequential goals.

REFERENCES

Belin, T.R., G.J. Diffendal, S. Mack, D.B. Rubin, J.L. Schafer, and A.M. Zaslavsky (1993), "Hierarchical Logistic-Regression Models for Imputation of Unresolved Enumeration Status in Undercount Estimation," *Journal of the American Statistical Association*, 88, 1149-1159.

Bell, W.R. (1993), "Using Information from Demographic Analysis in Post-Enumeration Survey Estimation," *Journal of the American Statistical Association*, 88, 1106-1118.

Bryant, B.E. (1993) "Census-Taking for a Litigious, Data-Driven Society," *Chance*, 6, 44-49.

Bureau of the Census (1992), "1990 Census Cost Components," Year 2000 Research and Development Staff Memorandum Series, Book I, Chapter 30, No. 4. Memo from Jay Keller to Susan Miskura (August 6, 1992), Bureau of the Census, U.S. Department of Commerce.

——— (1993), "Potential Cost Savings by Sampling for Nonresponse Follow-up," Year 2000 Research and Development Staff Memorandum Series, Book 1, Chapter 30, #11. Memo from Janice Pentercs to Jay Keller (September 17, 1993), Bureau of the Census, U.S. Department of Commerce.

——— (1994), "1995 Test-Experimental Design for Sampling for Nonresponse Followup," Memo from John Thompson to Susan Miskura (March 7, 1994), Bureau of the Census, U.S. Department of Commerce.

Coale, A.J (1955), "The Population of the United States in 1950 Classified by Age, Sex, and Color--A Revision of Census Figures," *Journal of the American Statistical Association*, 50, 16-54.

Committee on National Statistics (1994), *Counting People in the Information Age*, Panel to Evaluate Alternative Census Methods, National Research Council, National Academy Press, Washington D.C.

Dillman, D.A., J.R. Clark, and J.B. Treat (1994), "Influence of 13 Design Factors on Completion Rates to Decennial Census Questionnaires," Paper presented at the Annual Research Conference of the U.S. Bureau of the Census, Arlington, Virginia, March 1994.

Ericksen, E. P., L.F. Estrada, J.W. Tukey, and K.M. Wolter (1991), "Report on the 1990 Decennial Census and the Post-Enumeration Survey," Report of the Special Advisory Panel to the Secretary of the U.S. Department of Commerce (June 21, 1991).

Fienberg, S.E. (1992), "An Adjusted Census in 1990? The Trial," *Chance*, 5, 28-38.

Freedman, D.A., K.W. Wachter, D.C. Coster, D.R. Cutler, and S.P. Klein (1993), "Adjusting the Census of 1990: The Smoothing Model," *Evaluation Review*, 17, 371-443.

Freedman, D.A., K.W. Wachter, D.R. Cutler, and S.P. Klein (1994), "Adjusting the Census of 1990: Loss Functions," *Evaluation Review*.

Fuller, W.A., C.T. Isaki, and J.H. Tsay (1994), "Design and Estimation for Samples of Census Nonresponse," Paper presented at the Annual Research Conference of the U.S. Bureau of the Census, Arlington, Virginia, March 1994.

Himes, C.L., and Clogg, C.C. (1992), "An Overview of Demographic Analysis as a Method for Evaluating Census Coverage in the United States," *Population Index*, 58, 587-607.

Hogan, H. (1993), "The 1990 Post-Enumeration Survey: Operations and Results," *Journal of the American Statistical Association*, 88, 1047-1060.

Kish, L. (1965), *Survey Sampling*, John Wiley and Sons, New York.

Marks, E.S., W. Seltzer, and K.J. Krotki (1974), "Population Growth Estimation: A Handbook of Vital Statistics Measurement," *Population Council*, New York.

Mulry, M.H., and B.D. Spencer (1993), "Accuracy of the 1990 Census and Undercount Adjustments," *Journal of the American Statistical Association*, 88, 1080-1091.

Singh, R.P. (1993), "2000 Census: Coverage Measurement Prototype Methodology," (December 15, 1993), Memo from Rajendra Singh to John Thompson, Bureau of the Census, U.S. Department of Commerce.

Thompson, J. (1993), "Integrated Coverage Measurement and Sampling for Nonresponse Followup," (December 13, 1993), Paper presented at the meeting of the advisory committee for the design of the year 2000 census and census related activities for 2000-2009.

U.S. Department of Commerce (1991), "The Decision of the Secretary of Commerce on Whether a Statistical Adjustment of the 1990 Decennial Census of Population Should Be Made for Coverage of Deficiencies Resulting in an Overcount or Undercount of the Population," *The Federal Register*, Docket 91281-1181, July 15.

U.S. General Accounting Office (1992), "Decennial Census: 1990 Results Show Need for Fundamental Reform," U.S. Government Printing Office, Washington D.C.

Zaslavsky, A., and G.S. Wolfgang (1993), "Triple System Modeling of Census, Post-Enumeration Survey, and Administrative List Data," *Journal of Business and Economic Statistics*, 11, 279-288.