

ONWARD TOWARDS A 2000 CENSUS DESIGN: RESEARCH RESULTS

Robert D. Tortora, Susan M. Miskura, Don A. Dillman, Census Bureau
Susan M. Miskura, U.S. Census Bureau, Washington, D.C. 20233

KEY WORDS: Census, Response Rate, Coverage

This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.

I. Introduction and Overview

The Census Bureau is well into a program of research and development to select the design of the 2000 census by December 1995. The objectives of the program are to determine a new design for the decennial census that: improves the coverage of the census, with emphasis on reducing the differential undercount; reduces the cost of the census relative to previous censuses; and invites wide public input into selection of a final design. This paper describes some examples of how we are achieving these objectives, and raises issues for future consideration.

A key element of the Research and Development (R&D) Program has been to identify a full set of design alternatives in terms of different ways (options) to accomplish the major "distinguishing" features of a census. Based on intensive consultation with a wide range of stakeholders, we identified a full set of options for taking a census. The Technical Committee of the Year 2000 Task Force then constructed 14 "design alternatives," which formed a framework for different combinations of options for accomplishing the distinguishing features of the final design. The design alternatives were then reviewed in a series of meetings with over 25 groups of stakeholders. The process for identifying and reviewing the preliminary set of design alternatives, and a full description of the final 14 candidate designs, along with research plans to select among them, are described in (Keeley, 1992 and Year 2000 Research and Development Staff, 1992).

Since defining the alternatives last summer, we have been conducting research on each of the designs. We have widely and continually reviewed these results, as well as policy issues, with external stakeholders. Our goal has been to identify the most promising options by September 1993. This timing is necessary in order to allow adequate preparation for a full scale field test of the selected options in 1995.

By Mid-May, our research and consultations had provided most of the information we needed to specify the options to be tested in the 1995 Census Test.

Therefore, we issued 14 Design Alternative Recommendations (DARs) (Year 2000 Research and Development Staff, May 17, 1993) to document the proposed options. This was done to allow wide stakeholder review by September, and so that internally we could begin to develop and refine these options in a large scale set of tests consisting of 3 urban sites and one rural site. While we will be testing one basic design, having diversity in the sites will aid in assessing their likely performance on a national scale.

The content and format of the DARs acknowledged a major theme of our stakeholder consultations. Specifically, we were advised not to be "prisoners" of the specific 14 design alternatives. We were continually urged to select the most promising set of options across all designs, and concentrate development on assuring that they could be implemented in a fundamentally different census design. This has also been described as a "building block" approach to census reform. Therefore, the DARs responded to this input by specifying a workable set of options that, together, result in major fundamental changes to the census design.

The single design for the 1995 Census Test will be a combination of the most promising design options. Some of the options will be "tested" in a developmental sense; that is, our current information about costs, research results, and the like lead us to a single choice and we need to determine the best way to implement it. For other options, the test will incorporate experiments or simulations to determine the best variation of the option, or choice of options, to select. The key concept, however, is that **all 14 designs, as originally described, were eliminated in the Design Alternative Recommendations**, since every one of the 14 has at least one option that does not merit testing or evaluation for the 2000 census. The result is a "hybrid" design consisting of the most promising options among all 14. Most recently, we received a Test Design Recommendation (TDR) proposing more details of the 1995 Census Test (Year 2000 Research and Development Staff, August 1993).

As noted earlier, this paper describes only some of the research results and remaining issues for the Year 2000 census Research and Development Program. In particular, Section II describes a series of tests aimed at

improving the response rate to the census questionnaire mailing. This work illustrates our efforts to reduce the cost of the census by improving the rate of mail response; improving response rates may also indirectly help to improve coverage. Section III discusses issues regarding the concept of a "one-number census"--a census with results that combine traditional counting; the use of administrative records; and estimation. That section raises a number of questions regarding how to evaluate major conceptual and procedural changes in the design of the census.

This paper does not describe a number of other ongoing research efforts, such as those to improve the listing of household members on the census questionnaire; develop new technologies for collecting and processing census data; or improve record linkage techniques to support new design requirements. Information on work in these areas is available from the authors.

In addition, this paper does not describe the components of the R&D program that address the level and type of content to be collected in the census--that is, what questions should be asked in the census from the entire population or a sample. The R&D program is actively analyzing data needs and uses and changes in them over time; and looking at alternative ways that may better meet these needs. Again, information on those efforts is available from the authors..

Finally, note that the R&D program is being conducted in a dynamic environment. We continually learn new information from our research; integrate ideas from external stakeholders; and are subject to determinations made in the budget process. Direction and scope therefore changes rapidly. In particular, some of the plans and issues described in this paper are subject to rapid change due to one or more of these factors.

II. Research to Improve Response Rates to the Census

In the process of researching the design alternatives, we have undertaken various experiments and studies in a large number of areas. All of them could potentially result in major changes to the design features or the underlying infrastructure of the decennial census. Previous ASA papers (Miskura, 1992 and Bryant, Miskura, and Dinwiddie, 1992) have described the research in a number of areas. In this section, we describe the major line of research targeted at improving the response rate to the census. We are currently winding up this work, and as a result, have committed to an elaborated implementation scheme, and questionnaire simplification, for the 2000 census. This

illustrates another major theme of the R&D program--that through a series of small scale tests early in the decade, we wanted to provide information to allow larger tests to concentrate on developing methods that require full scale census conditions to draw objective conclusions.

The 1990 decennial census required surveying over 100,000,000 households. The primary means of data collection was intended to be a mail-back census form. The 1990 response rate for this mail-back form was 65 percent, 10 percentage points lower than the 75 percent response rate obtained in 1980. A consequence of this lower response rate was the necessity of spending an additional \$100,000,000 or perhaps more for personal enumerators to complete the census.

During the last two years, four experiments have been designed in order to identify ways that the census mail-back response rate might be improved. Three of these experiments have been completed, and the fourth is in the implementation stage. These experiments have been conceptually linked to one another so that results may be cumulated across experiments.

The Simplified Questionnaire Test (SQT)

This experiment, conducted in the Spring of 1992, tested whether the use of respondent-friendly census forms, reducing the number of questions (from those included in the 1990 short form), and asking for social security numbers would influence response rates. The actual 1990 short form, with dates changed to 1992, was used as a control, and a national sample of 17,000 households was surveyed. The sample was stratified so that differences in response from Low Response Areas in the 1990 Census (primarily central city areas) where 11 percent of the U.S. population resides, and High Response areas (the remaining 89 percent of the United States) could be determined. All households in the five experimental panels received prenotice letters sent a few days before the Census Form, a reminder postcard sent a few days afterwards, and in the case of nonresponse a replacement questionnaire about three weeks later. All correspondence was sent by first-class mail and was addressed to "the resident(s) at: (the address)."

This experiment showed that respondent friendly construction improved response rate by 3.4 percentage points, and shortening the form to include no housing questions and two fewer individual questions (allowing the form to be one sheet of paper on both sides, rather than a booklet of two sheets of paper folded and stapled with the equivalent of eight pages of information) improved response by 4.6 percentage points. Shortening the form further to include only a roster (name and date of birth) did not improve response

further, and the request for social security number decreased response by 3.4 percentage points. (Dillman, et al. In Press)

The Implementation Test (IT)

Because the overall completion rates for the SQT were quite high in comparison to previous noncensus year tests, and the prenotice and reminder card were thought to have contributed significantly to that response, these elements were singled out for testing in the IT. In addition, the effect of including a stamped reply envelope, found effective in other noncensus research, was added for testing. All eight of the possible individual and combined uses of these three elements were tested in a survey of 50,000 households in a national sample comparable to that used for the SQT, including stratification. No replacement questionnaire was sent. This test was conducted during the Fall of 1992.

The prenotice letter, stamp, and reminder postcard individually improved response rates by 6.4, 2.6, and 8.0 percentage points, respectively. The 2.6 point increase was not statistically significant, although the magnitude of the increase was similar to what has often been found in previous research (Dillman, 1978, 1991; Armstrong and Luske, 1987). The effects of the elements were also found to be mostly additive, and did not interact with one another. In comparison to the control group, the combination of letter-stamp improved response 9.8 percentage points, the stamp and reminder, 9.5 percent, and the letter and reminder, 12.7 percentage points. All three elements together improved response by 14.2 percent. Each use of the letter and reminder added significantly to response, but the stamp only added significantly when used with a preletter and no reminder. The most important conclusion from this experiment was that neither a prenotice letter or reminder postcard eliminates the effect of the other; both seem important to achieving a high response. (Dillman et al. 1993)

By comparing the response rates for this test which did not use the replacement questionnaire mailing, and the SQT which did, it was reasoned that the replacement questionnaire could have improved response by about another 10.4 percentage points. This estimate compared to about 11.0 percentage points obtained for the SQT by noting when response from the original mailout had dwindled to a few per day, and the effect of the replacement questionnaire started coming in.

The Mail-Telephone Mode Test (MTMT)

The MTMT, for which data collection occurred in Spring 1993, is designed to determine whether response rates can be improved by offering people the alternative of calling in their responses by telephone, rather than sending them in by mail. This test measures pure preference for responding over the telephone by offering that possible in each of the contacts (prenotice, questionnaire mailing, reminder postcard, and replacement form mailing).

However, it seems undesirable to switch many people who would respond by mail to telephone, for reasons of costs as well as the logistics of setting up telephone call-in facilities large enough to handle the first days of response when most calls would probably be made. Therefore, three other panels, in this study of 22,500 households stratified as were the previous experiments, introduced the telephone later in the data collection process. Two panels introduced a special telephone invitation after the reminder postcard used in the SQT, IT and control for this experiment, to determine if the response improve of 10.4 percentage points from the replacement form mailing, can be obtained without having to send that mailing which is both expensive and logistically difficult because of the large numbers of households involved in the census.

This experiment related to the other experiments by utilizing the respondent friendly booklet form and exact implementation procedure used in the SQT as the control group.

Results of this experiment show that offering with each contact the alternatives of responding either by mail or telephone, about eight percent of those who respond choose telephone. However, the response rate is not improved. The main effect of the telephone invitation is that some who would respond by mail switch to telephone. This finding suggests that offering mail or telephone response alternatives to those who receive the form by mail will not improve response rates to the census. (Clark et al. 1993)

However, a special fourth contact of a letter only (no replacement questionnaire included) improved response by 2.8 percentage points, 0.2 percent of the increase being from mail and 2.6 percent being attributable to the telephone. To the extent that offering the telephone improves response, it appears to happen in a follow-up situation when a replacement questionnaire is not also being provided.

The Appeals and Long Form Experiment (ALFE)

This experiment is an ambitious one which combines together for reasons of efficiency two quite different tests.

One part of the test involves four treatment groups which will test the application of respondent friendly questionnaire design procedures to long questionnaires. The 1990 long form serves as the control group. One respondent friendly design is modeled after the "separate space for each individual" procedure used in all of the other tests. The second design is a graphically improved row (for each question) column (for each individual) design, which results in a form that eight pages shorter. We hope to learn for long forms, what we learned in the SQT for short forms--whether respondent friendly design concepts improve response, and in particular for item nonresponse, which is an especially significant problem on long forms. This test will again use the prenotice, reminder and replacement form mailings of previous tests.

The second part of this experiment tests whether the addition of certain appeals, i.e. emphasizing that response is mandatory vs. emphasizing the benefits of responding in combination with strong vs. less strong explanations of confidentiality will improve response. The "stick" approach of emphasizing mandatory response was found effective in Census Bureau Industry surveys when stamped onto envelopes sent to businesses. The "carrot" approach of emphasizing benefits in a "motivational insert" was found effective for improving household response to test censuses in the 1980s. In this test, the outside of the envelope and motivational insert will be used jointly to present strong mandatory and benefits messages. Because of past research that has found that too much emphasis on confidentiality can discourage response, two versions of such messages will be tested with each of the appeals.

This experiment is being conducted during the summer months of 1993, again using the prenotice, reminder, replacement questionnaire procedure with the respondent-friendly booklet questionnaire of previous tests in this series. Results will be available in early Fall.

What have we learned?

First, from the SQT and IT, it can be argued that by simultaneously using five elements: respondent-friendly construction, a shorter form, a prenotice, and reminder postcard, and a replacement form we have shown that completion rates in a noncensus year can be improved by 31.4 percentage points. Joint use of the prenotice and reminder improves response by 12.7 percentage points, respondent friendly construction by 3.4 points, slightly shortening the form by 4.6 points, and the replacement form by 10.4 percentage points.

It is important to recognize that because of the way that the experiments were designed, assumptions are

required to reach this conclusion. First, while respondent-friendly construction was individually tested, the shorter length was only tested using such construction, so that no independent measure of length was obtained.¹ Second, it is assumed that the additional increment of response obtained from testing the combination of prenotice and reminder on a shorter, respondent-friendly form is no less than it would have been if tested on the regular 1990 short form which lacked either of these procedures.²

Based on this assumption, completion rates have been improved by 123 percent for Low response areas vs. 73 percent for high response areas. It can also be seen that the test elements have been differentially effective in the high and low response areas. Whereas respondent-friendly construction was especially effective in low response areas, the combination of prenotice and reminder were especially effective in high response areas.

As of yet, we do not know whether invitations to call answers in by telephone, the use of mandatory vs. benefits appeals, or different confidentiality explanations will improve response. It's important to note that our designs for these experiments will test each of these elements on top of the use of elements already tested, so that we will continue to emphasize understanding the cumulative effects of several elements in combination, rather than elements as individually applied alternatives to other elements.³

This section has summarized the techniques which, based on past mail research literature, seemed likely to improve response. For policy reasons, the Census Bureau determined that we would not pursue the use of a sweepstakes (Lucas, 1993).

Next Steps?

There are at least two significant barriers to the direct application of this research to conduct of the 2000 census.

First, it is important to recognize that these tests are being done in noncensus years. In the past, the Census Bureau has obtained somewhat lower response rates in noncensus years than in census years. For example, the 1986 content test, obtained only a 49.6 percent response, while the 1990 census achieved a 65 percent response rate. The usual explanation for this difference is "census climate," a succinct explanation of the combination of media attention, advertising, and cultural sense of participation that seems to build during each census year.

The response rates obtained in our tests with the use of the five elements found to increase response are much higher than normally obtained in noncensus years,

but are close to the same, or perhaps a little lower, than those obtained during with none of these elements in the last census.⁴ We do not know whether the existence of a "census climate" will substitute for the effects of these elements or add to the response likely to be obtained in a census year. Certainly a 30 percentage point increase will not be realized in the 2000 census since that would suggest a response of nearly 100 percent. That leaves us with considerable uncertainty with respect to the exact implications of the present findings, and those yet to be obtained from the ALFE test. The use of promising methods in the 1995 Census Test will give us some reading on this.

Second, there is a need to make some of the elements we have tested more practical for use in a mailing to more than 100,000,000 households. For example, the replacement questionnaire has utilized a letter that was addressed to a particular household which had to be put into the same envelope with a questionnaire containing the same address. This "double-imaging" requirement would be enormously difficult to assure against error during mass production for a census. Also, the current respondent friendly form requires keypunching for processing. Efforts to build a respondent friendly form which can be optically scanned during processing have not yet been undertaken. These are only two of the many adaptive elements that need further investigation.

By-Products of the Research

It is not unusual that research aimed at predetermined objectives results in new paths of development. Our work on questionnaire simplification and implementation falls in this category.

One of the new aspects of the implementation strategy is to use first-class mailings. When first-class mail is used, the Postal Service returns material that has been addressed to vacant or nonexistent addresses (Postmaster Returns or PMRs). In the past, without first-class mailings, we were not able to rely on these designations; but with first-class mailings, we have an opportunity to redesign one of the large cost components of census data collection. Specifically, in the past two censuses, two interviewer visits were made to vacant and nonexistent addresses: the first, during nonresponse follow-up, made the initial classification; the second verified it.

By analyzing operational data from the SQT and IT, we believe that we may be able to rely on PMRs from the prenotice and/or questionnaire mailing to replace one visit to these units. We would send enumerators out only once to verify the "vacant" or "nonexistent" status and conduct an interview as

appropriate. The single visit can be made during the few weeks between the questionnaire mailing and obtaining the results of the replacement questionnaire--time that would be used only for checking in questionnaires and preparation for nonresponse follow-up under the previous procedures. This means that we can now complete a significant portion of the work load much earlier than before; it would be completed close to Census Day which means that quality should be improved; we eliminate the second visit to these units; and we will have an experienced crew of interviewers for nonresponse follow-up which may improve the quality and timeliness of that operation.

III. Integrated Coverage Measurement and the One-Number Census

Definition

The term "One-Number Census" names the concept that **the decennial census is designed to produce the best possible single set of results by legal deadlines, and that those results are based on an appropriate combination of counting, assignment, and statistical estimation techniques.**

By results, we mean all totals, characteristics, derived measures, and the like that are based on the decennial census. In particular, the totals that are certified as official for various governmental units are a critical subset of results.

The one-number census concept applies to persons, units, and households. For simplicity and to recognize the importance of population totals, the term "persons" is used as a generic reference unless otherwise noted.

The definition of this concept is not meant to eliminate the possibility that evidence of specific, localized error documented by a local official would be used to correct results.

Methods of Enumeration

There are three ways someone can be enumerated in the census--counting, assignment, and estimation. In the past, census results have been dominated by the contribution of counting, although there have been small, sometimes ad hoc, applications of other techniques. In a one-number census, the fundamental approach to the use of assignment and estimation is changed from one of supplemental and selected use, to one of large-scale integration into the census process.

By **counting**, we mean the full array of techniques by which we attempt direct contact with respondents--by mail, by personal visit, by telephone, or by other means. We also mean data obtained by proxy for another household, housing unit, or person.

Historically, we also have added persons to the census by obtaining information about their existence on administrative records, and verifying this information. This is also considered counting.

By **assignment**, we mean the use of indirect evidence from administrative records, to add persons to a specific geographic location. The sufficiency of evidence to do this without field verification will be determined based on an assessment of the accuracy of administrative records during earlier research.

By **estimation**, we mean the application of an array of statistical techniques to account for persons or units not directly counted or assigned. Most likely are some combination of the following techniques:

1. Performing all or part of the enumeration on a sample basis and using statistical techniques to estimate totals and characteristics.
2. Selecting a sample from administrative records for direct verification or assignment; and using statistical techniques to add others.
3. Performing some appropriate coverage measurement procedure and integrating resulting estimates into the final census result.
5. Modeling to produce estimates for or of persons and geographic areas.

This definition of "estimation" does not predetermine, preclude or eliminate any technique to provide or create information about a person's or unit's existence or characteristics. In particular, it accommodates **imputation**--the computer allocation of codes in place of unacceptable entries or blanks--and **substitution**--the allocation of a full set of characteristics for a person or housing unit--which is a standard statistical practice that enhances the usefulness of the data. The exact mix of methodology is a research issue to be developed before the 2000 census.

In all cases, the contribution of any individual technique or process will not be separable within the final result, although we could provide a "source code" to say where any record or result came from.

Relationship to Design Alternatives, Features, and Criteria

In the Design Alternative Recommendations, we propose, for testing in 1995, the option "Integrated Coverage Measurement" that underlies the concept of a one-number census. That option integrally means that there is a single best census result incorporating

counting, assignment, and estimation. **The premise of "Integrated Coverage Measurement" is that all three of these techniques are designed to complement each other to meet legal deadlines.**

A one-number census is desirable for many reasons. A major motivation is to avoid the problems, controversy, and disruption caused by more than one set of numbers being widely available. Note the following desirable criterion from the Federal Register Notice:

(It is desirable to) "Provide a single, best set of census results produced by legal deadlines for apportionment and redistricting."

The designs under consideration vary in combining direct enumeration and estimation techniques to produce the census results. It is highly desirable to specify a design that has an appropriate combination of these techniques and integrates the results on a time schedule that provides a single set of counts and characteristics by mandated statutory deadlines. Each design must be assessed in terms of the likelihood of meeting these objectives. While the Census Bureau does not necessarily assume that current deadlines will be unchanged, it expects that a convincing case must be made for proposing the legislative actions that would be required to change them.

Integrated coverage measurement and the "one-number census" differ fundamentally from past practice in the census. In particular, they are in contrast to the 1990 "dual strategy approach" where the census and coverage measurement method were implemented independently; the results compared; and a decision made about adjusting the census or not. The one-number strategy starts out by saying the results from measurement of coverage will be incorporated into the official census results.

While a "one-number census" has only one, best set of official results, its definition does not preclude release or use of partial or preliminary data, either for census programs such as Local Review, or at points in the steps to combine information across counting, assignment, and estimation methods.

Issues

There are a number of questions and issues that go beyond just defining a "one-number census." They cover both technical concerns, and matters of public policy and acceptance that need to be addressed in

developing the concept into a feasible methodology for the 2000 census.

These issues include:

1. Methods to do estimation-based coverage improvement will require indirect estimation to produce final, official counts by small geographic areas. (Indirect estimation is used as a generic term that includes "synthetic estimation;" the key issue is that direct estimates of coverage will be available only for high levels of geography; since census results are needed at the block and intermediate levels, we need a method to "indirectly" estimate totals for those lower levels.) It would be highly desirable to promote the understanding of this requirement, and gain consensus on the technique(s) to be used, early in the decade. The major aspect of getting such agreement is whether the use of indirect estimation and the specific technique to be used is reasonable. Put another way, given that we will never be able to develop or afford a technique that "guarantees" improvement at all geographic levels or individual entities, we need to decide early on the level we need to demonstrate accuracy for, and research the way to do so.

An early selection and definition is desirable so that 1) it can be used to inform stakeholders of the implications; and 2) that it can be fully developed for use in the census. The Census Bureau has begun to surface possible techniques, and has requested the CNSTAT Methods Panel to advise on this issue.

2. Related to the above point, the ability to produce results by mandated deadlines means that the techniques must be fully tested and developed prior to the census. We cannot go into this process without solid and proven plans. This means that the integrity and feasibility of operations, procedures, statistical methods, and the like must be fully specified, programmed, and supported going into the census. On the other hand, we must balance prespecification of methods and methodologies with the need to assure that changes that will improve the final results can be made. A process for external agreement to such changes should be put in place so that it is available in real time if needed.
3. Census results at all geographic levels must be additive. It is not acceptable to have different values for high geographic levels that differ from

values obtained if smaller level results are added up.

This does not imply that all data need to be available at one time; for example, redistricting data under current deadlines are produced after reapportionment counts. However, it does require that those data, when produced, are additive to the reapportionment counts.

4. The criteria for selecting a best methodology for the one-number census will be subject to wide public input throughout the design and planning of the 2000 census. We expect much of this debate to center on the optimal way to assure accurate shares of the population at various geographic levels, constrained by the additivity requirement. In addition, we will be proactive in educating stakeholders on the limitations and expectations of this methodology. For example, for small areas, whether statistical or political, we will not be able to assure absolute improvement. We will be able to demonstrate improvement at higher levels and the integrity of statistical techniques to "carry down" results to lower levels.
5. The definition of this concept specifies that the single set of results be available by legally mandated deadlines. The current legal deadlines are December 31 of the census year for reapportionment, and April 1 of the following year for redistricting (with many data coming out earlier on a flow basis.) The Census Bureau will strive to have ALL census results come out earlier than required, but complexities in producing the single set of results may make these deadlines tight. Therefore, we must begin now to identify methods that can adhere to those deadlines:

It is possible that in performing this research, we identify a methodology that can meet these deadlines, but will not produce data as accurately as another methodology that cannot meet the deadlines. If this arises, a policy decision about selecting the less accurate method, vs. the "better method" that requires changing the deadlines, must be made.
6. We will evaluate the quality of the final census results. Our plan is to evaluate the efficacy of the process used to produce it. The methodology for this coverage process evaluation is a subject for future research. Even if it does result in estimates of error in the official results, these data will NOT be used for any further correction of the census. They will be used for future census planning.

7. The definition of "one-number census" is idealized in that it presumes that statistical techniques can be developed that apply equally well to persons, housing units, households, and characteristics. In the course of research and development, it is possible that only partial success is realized. For example, we may learn how to improve the count of persons through estimation, but not be able to develop techniques for units and/or households. Or, we may find we can do major but not total improvement by legal deadlines. We need a process to assess and gain consensus on actions to be taken to address such situations.
8. What is the role of demographic analysis (DA)? At minimum, demographic analysis will be used as an independent reading on the quality of the census. However, DA results have the potential to pinpoint coverage errors during the census taking process and identify areas of concern in the census or DA itself.

Proving in the One-Number Census

We have proposed that the 1995 Census Test be a one-number census. This means that we will use several improved design features, including estimation-based coverage improvement techniques, to produce the single set of official census results for the test sites. Given the overall objective to improve coverage--especially to reduce differential coverage relative to 1990--it will be important to be able to illustrate that this concept does in fact yield these advances. Determining how to do this combines some intriguing technical, perceptual, and operational issues.

One set of issues involves the estimation-based coverage improvement methods to be used. We have a great deal of experience with Post-Enumeration Survey methods, based on dual system estimation, especially from 1990. But the PES is a complex operation and it is not clear that it can be accomplished by legally mandated dates. One line of current investigation is to determine its potential to do so within the framework of our redesigned census. Concurrently, we are looking at new methodologies--called the SuperCensus and CensusPlus--to determine their potential to address the operational and timing difficulties of the PES while performing at least as well to improve coverage. We are currently investigating the technical and operational aspects of these methods. Sometime this fall, using the research on all of these methods, we expect to select between the SuperCensus and CensusPlus approaches, and determine if we have enough information to select between that choice and

PES. It is possible that we may want to use both methodologies in the 1995 Census Test.

Another set of issues has to do with determining if the new one-number approach, which will rely less on counting techniques and more on estimation than the 1990 census, yields the desired improvement. This set of issues encompasses two areas:

- What measures are needed to make a convincing case? Is it sufficient to show that we have improved the relative coverage of population groups; or is it important to also show that the "proportionate shares" of the total population across all areas are improved? In either case, what is the geographic level at which such improvement must be shown?
- How can we obtain these measures? Even if we can gain agreement on what they should be, designing the test to produce valid measures is not a trivial undertaking. Ideally we would "split" each of the test sites into at least two portions: one using 1990 procedures, another using the new design. But actually controlling such a configuration will be virtually impossible because the new design includes features--such as multiple response options and new address list compilation methods--that cannot easily be separated from 1990 procedures. Therefore, we would want to consider other approaches based on modelling 1990 data for comparisons to 1995 results; or using a series of indirect indicators--such as average household sizes, sex ratios, or tabulations of persons added by new design features--to draw conclusions about improvements. All of these approaches have drawbacks, including the fact that the new design uses innovative and changes methods that may not go smoothly the first time they are used.

We are currently discussing these issues within the Census Bureau and invite the input of the technical community on them.

Footnotes

1. It's possible that the effect of shorter length would have been slightly different had it been tested in combination with the 1990 style of form construction. A panel to independently test length apart from respondent-friendly construction was not included in the SQT test.

2. This assumption seems tenable inasmuch as response effects of individual elements are usually less when tested after other response inducing elements have been added. For example in the IT the effect of the reminder postcard was 8.3 when added before the other elements tested, 6.4 when added after the prenotice, and only 4.4 when added after both the stamped return and prenotice were used. The addition of other treatment groups would have obviated the need for making this assumption, but cost considerations prohibited the addition of the many additional group that might have been relevant to developing such estimates of effects.
3. Following such an alternative approach in design we might for example have taken the 1990 short form, and done a series of tests in which only one element at a time was tested to see how much response could be approved by each. Such a line of research would have allowed us to make conclusions about e.g. if only one element were to be added in an effort to improve response, which should it be: a telephone invitation? a prenotice letter? a benefits appeal?, etc.
4. Exact comparisons are very difficult to make because of somewhat different sample frames, with different numbers of vacant and nonexistent household addresses. Thus, any such comparisons can only be approximate.

References

- Armstrong, J. S. and E. J. Luske, (1975), "Return Postage in Mail Surveys: A Meta Analysis," Public Opinion Quarterly, 51 (1) 233-248.
- Bryant, Barbara Everitt, Miskura, Susan M., and Dinwiddie, James L., "Strategic Issues for 2000 Census Design," presented at the Joint Statistical Meetings, August 1992.
- Clark, J. R., D. A. Dillman, and K. K. West, "Influences of an Invitation to Answer by Telephone on Response to Census Questionnaires." Paper presented at 1993 Joint Statistical Meetings, San Francisco, CA.
- Dillman, D. A., M. Sinclair and J. R. Clark, (In Press), "Effects of Questionnaire Length, Respondent-Friendly Design and a Difficult Questionnaire Response Rates for Occupant-Address Census Mail Surveys." Public Opinion Quarterly.
- Dillman, D.A., J. R. Clark and M. Sinclair, (1993), "How Prenotice Letters, Stamped Return Envelopes and Reminder Postcards Affect Mailback Response," Annual Research Conference, U.S. Bureau of the Census, Washington, D.C.
- Dillman, D. A. (1991), The Design and Administration of Mail Surveys," Annual Review of Sociology, 17, 225-49.
- Dillman, D. A. (1978), Mail and Telephone Surveys: The Total Design Method, New York: Wiley-Interscience, John Wiley & Sons.
- Keeley, Catherine, Documentation of Alternative Designs Assessment Meeting (ADAM) Presentation, May 14, 1992, Internal Census Bureau Memorandum.
- Lucas, Sandy, Memorandum to Robert Tortora, Subject: Census Lottery/Sweepstakes Conclusion, April 26, 1993.
- Miskura, Susan M., "Forward from 1990: Designing the 2000 Census" Presented at the Joint Statistical Meetings, August 1992.
- Year 2000 Research and Development Staff, "2000 Census Research and Development Alternative Design Program, " June 1992, Internal Census Bureau Memorandum.
- Year 2000 Research and Development Staff. Design Alternative Recommendations," May 17, 1993 - Series of 14 documents.
- Year 2000 Research and Development Staff, "1995 Census Test Design Recommendation, " July 1993.