

John H. Thompson and Robert E. Fay*

John H. Thompson, U.S. Bureau of the Census, Washington, DC 20233-7000

Key words: Census design, decennial census, dual-system estimation, post-enumeration survey

1. Introduction

At the time this paper is presented, the Dress Rehearsal for Census 2000 is in process in Sacramento, CA; Columbia, SC and surrounding counties; and Menominee County, WI. Substantially, the designs in Sacramento and Menominee (primarily an American Indian Reservation) represent the Census Bureau's plan for Census 2000. The Dress Rehearsal in South Carolina, originally intended as an additional test site for the planned design, is instead being conducted as a "nonsampling census," without the sampling and estimation innovations in the other two sites. This change was required by the Budget Agreement for FY 1998 in November, 1997.

The title of this session is "The Controversy about Sampling in the Census." We conjecture that few, if any, U.S. statisticians are unaware that the 2000 Census is now controversial. The fielding of different designs in the Dress Rehearsal two years before Census Day, 2000, is a direct consequence of the controversy. Indeed, the Census Bureau's plans for 2000 may be set aside through either of two possible external actions. First, it is possible that the Supreme Court may forbid sampling to estimate the population count, an issue now under litigation. Although reviews of the issues have supported the position that our plans meet legal requirements (Edmonston and Schultze 1995, pp. 239-258), we will not revisit this issue here. Secondly, the Congress may decide to prohibit sampling in the spring of 1999. The Census Bureau expects to issue its detailed plans for the 2000 census in the fall of 1998 and its assessment of the Dress Rehearsal at the end of January, 1999.

Census Bureau staff are presenting numerous papers on specific technical matters and operational plans at these meetings, but we believe that most of the controversy originates from the broader issues. Consequently, beyond the two planned reports just mentioned, we believe that general reports and papers will have a critical role in presenting the Census Bureau's plans for 2000 to both the public and our professional peers. The Census Bureau (U.S. Bureau of the Census 1997) has issued an overview of its general plans to Congress and an operational plan (U.S. Bureau of the Census 1998). Recently, Wright (1998) summarized the basic sampling and estimation concepts for Census 2000 for a wide audience. A paper summarizing technical aspects of Dress Rehearsal

sampling and estimation has been submitted for publication (Farber, Fay, and Schindler 1998). Waite and Hogan (1998) will present a more detailed overview of the 2000 design, particularly as it is reflected in the Dress Rehearsal designs in Sacramento and Menominee.

Recently, Acting Director Holmes (1998) summarized the Census Bureau's overall situation and plans to the Monitoring Board, an eight-member board composed of four Presidential and four Congressional appointees. In this paper, we plan to address the same question used by Holmes to frame many of his remarks, "How did we get here?". We will summarize important milestones and evidence shaping the Census Bureau's plans. We will indicate the remaining questions that we expect to address with our Dress Rehearsal data both by Fall, 1998 and February, 1999.

2. The Context for the Census Bureau's Plans

Although now controversial, the Census Bureau's plans have a basis in virtually undisputed facts. The method of demographic analysis (Coale 1955, Fay, Passel, and Robinson 1988, Robinson, Ahmed, Das Gupta, and Woodrow 1993), using data sources essentially independent of the decennial censuses, has shown a consistently higher net undercount of Blacks than non-Blacks. Over the period 1940-1990, the estimated difference between the Black and non-Black undercount rates varied over the range 3.4-4.4 percentage points. Over this comparatively narrow range, the greatest difference was for 1990, a 4.4 percentage point difference between an estimated 5.7 percent undercount for Blacks and 1.3 percent for non-Blacks. There is no suggestion of a downward pattern to the difference over time (1). Demographic analysis also shows the 1990 census to have a higher total net undercount than the 1980 census, reversing a downward trend from 1940-1980. Nonetheless, the 1990 census has the second lowest estimated undercount rate during 1940-1990.

Demographic analysis of the U.S. population provides an informative assessment of the net coverage of the censuses nationally and across time. Because it does not provide satisfactory estimates for states and other geographic units (2), however, the Census Bureau has not regarded demographic analysis by itself as an adequate basis for undercount adjustment. It also does not provide usable evidence on the undercount of other important groups, including American Indians, Asian and Pacific Islanders, and persons of Hispanic origin (Robinson, Ahmed, Das Gupta, and Woodrow 1993, p. 1064).

A second undisputed fact is that the census has become increasingly expensive. In fixed 1990 dollars, the 1960 census cost \$10 per housing unit; the 1970, \$11; the 1980, \$20; and the 1990, \$25 (Edmonston and Schultze 1995, p. 44).

Controversy is not new to recent censuses. The notion of correcting the census counts for undercount dates back 20 or more years. Coverage measurement studies based on sample surveys have been regarded as the primary source for such data. The survey-based methods provide both the geographic detail and, with adequate sample size, are able to evaluate coverage for American Indians, Asian and Pacific Islanders, and persons of Hispanic origin. On the other hand, coverage studies have not fully replicated the findings from demographic analysis that Black males are missed at substantially higher rates than Black females, even though the results from demographic analysis are regarded as authoritative in that respect (3).

The Census Bureau considered but rejected plans to incorporate results from the Post-Enumeration Program (PEP) into the population totals in the 1980 census. The official position of the Census Bureau was that the findings from the 1980 PEP were too uncertain to form the basis of an adjustment. Substantial missing data and other difficulties in interpreting the findings were represented by 12 alternative sets of estimates (shown in Fay, Passel, and Robinson, 1988), which varied widely. A lawsuit by the city of Detroit, New York and others for adjustment did not prevail.

The Census Bureau developed a research strategy to adjust the 1990 census, based on Post-Enumeration Survey (PES) with a sample size of 300,000 households. The Census Bureau developed a dual strategy to complete the traditional census by December 31, 1990, and, if the PES later proved successful, to revise the census totals accordingly. In 1987, the Secretary of Commerce decided against adjusting the census, and in effect reduced the PES to the 150,000 households budgeted as an evaluation. A 1988 lawsuit, again by New York, resulted in an agreement for the Census Bureau to proceed according to the dual strategy, reinstating the 150,000-household PES as the potential source of census adjustment. In June, 1991, Director Barbara Bryant, based on a 7-to-2 recommendation of a Census Bureau committee, in turn recommended adjustment to Secretary Robert Mosbacher. The Secretary, in consultation with an external panel of eight members and internal Commerce staff, issued a report (Mosbacher 1991) deciding against adjustment of the census figures on July 15, 1991. The Secretary's report, while recognizing that the PES adjustment would move the national total closer to the truth, argued that the adjustment would not necessarily improve the relative accuracy of the census at all levels.

Because of a general dissatisfaction with the outcome of the 1990 census (4), Congress enacted the Decennial Census Improvement Act of 1991, PL 102-135 (Appendix 1). The law established the Panel on Census Requirements in the Year 2000 and Beyond, under the Committee on National Statistics of the National Academy of Sciences, to study how the most accurate population count possible could be obtained and how, consistent with this goal, other demographic and housing data could be obtained. The committee was charged to study improving enumeration methods; alternative collection methodologies, including use of administrative records; appropriateness of sampling methods for collecting or refining population data for different levels of geography; how needs for data besides population data are assessed and whether alternatives can be used. For each alternative, the panel was directed to examine its advantages, disadvantages, and cost. If the method did not involve the direct collection of data from persons, the panel was to assess its effect on privacy, public confidence in the census, and the integrity of the census. The legislation gathered broad support in both houses (5). Thus, the Academy was given wide latitude to consider fundamental changes to the census design.

Simultaneously, the Census Bureau continued to study the undercount of the 1990 census. After the 1991 decision not to adjust the census, Secretary Mosbacher delegated to the Census Bureau the decision on whether to incorporate the results from the 1990 PES into postcensal estimates. (A similar approach is used by Canada, which issues its census numbers unadjusted but later adjusts the census base for postcensal estimates.) The Census Bureau examined the criticisms of the 1991 methodology, discovered and corrected significant errors in processing the PES, revised the methodology for distributing the undercount, and conducted further methodological research (Hogan 1993, pp. 1052-1054). Evaluation of the estimates produced a mixed result, however. Estimates for states appeared more accurate after adjustment, whereas the evidence was less clear with respect to substate accuracy. Director Bryant, faced with divided technical opinion on the merits of adjusting (6) ultimately decided against doing so for the postcensal estimates, primarily because adjustment could not be proven significantly better at many substate levels. Since the weight of evidence indicated a net improvement in the state figures, however, a specific official use of the 1990 PES was sanctioned: estimates of undercounted persons at the state level from the 1992 estimates could be incorporated into state controls used by demographic surveys, such as the Current Population Survey (CPS), by joint agreement with sponsoring agencies. In fact, this change was subsequently implemented in the CPS, the

Survey of Income and Program Participation, and a number of other surveys.

After Secretary Mosbacher's July 15, 1991 decision, New York and other plaintiffs reinitiated the suit. At the District level, the Judge ruled that the Secretary had both the authority and sufficient basis for his decision, even though the judge expressed the view that he would have acted differently. The Court of Appeals overturned this ruling, concluding that the District judge had not recognized the correct legal standard. The Supreme Court concurred with the District judge's interpretation of the law. The decision effectively blocked further consideration of compelling an adjustment of the 1990 census, while not appearing to block the possibility of a statistical adjustment for undercount in future censuses. On the other hand, the language of the decision raised the point that the proposed 1990 adjustment adjusted a state's population on the basis of data from other states (7).

The litigation over adjustment, along with public and Congressional reaction to the finding that the 1990 census had not improved on the net undercount of 1980, provided an impetus for the bureau staff planning the 2000 Census to undertake a fundamental redesign. An early and central concept was of the *One-Number Census*. Tortora, Miskura, and Dillman (1993, p. 124) provided the following 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.**

The post-enumeration survey, with results available after initial census results were released, would be redesigned as *Integrated Coverage Measurement (ICM)*, to become an integral part of the concept and timing of the census. In hindsight, the dual strategy used in 1990, producing 1) a census count meeting legal requirements by the mandatory dates, followed by 2) a consideration of the merits of adjustment based on coverage findings and related assessments available only months after release of the preliminary results, appeared to guarantee controversy. The 1990 PES estimates, once produced, in effect divided state and local governments into winners and losers. It was reasoned that, if results from the undercoverage survey were incorporated into the official population totals from the start as the one-number census, then perhaps there would be wide acceptance of the results. In other words, the one-number census would attempt to build agreement on the process beforehand, rather than trying to obtain agreement on the outcome (8). Of course, the plan requires establishing a broad political consensus on the process in advance.

In addition to the Congressionally-mandated panel, the Panel on Census Requirements in the Year 2000 and Beyond, the Census Bureau commissioned the National Academy of Sciences to form a second panel, the Panel to Evaluate Alternative Census Methods, to provide more specific technical advice. The primary findings of the two panels were reported by Edmonston and Schultze (1995) and Steffey and Bradburn (1994), respectively. As previously noted, the first panel was charged by Congress to consider a very wide range of options, including extensive use of the U.S. Postal Service to collect census data and use of administrative records as partial replacement for enumeration. The independent review and recommendation by the Requirements Panel was instrumental in bringing about legal changes allowing the Census Bureau to share address information with the post office and local governments to improve the accuracy. The panels also considered important issues that we will not elaborate here, such as the future of the long form and questions on race and ethnicity.

Both panels supported in principle the Census Bureau's one-number census concept, and we have widely cited their reports and considered their advice. At the same time, however, the Census Bureau's plans have developed on the basis of subsequent evidence and public reaction. For example, the Requirements Panel elaborated the notion of a truncated census with a period during which a reasonable effort was to be made to follow-up all nonrespondents. (We note in the next section why the Census Bureau selected direct sampling instead.) The design was also shaped by a Task Force for Planning the Year 2000 Census comprising internal staff, staff of other agencies, and outside users (U.S. Bureau of the Census 1995). In the next section of this paper, we shall try to summarize the milestones along the way, providing a response to "How did we get here?".

3. Milestones in the Research Program

In this section, we describe principal milestones in the research program shaping our design. Because some of these decisions have had strong implications on others, we have organized them in an approximate chronological order, even though some projects, such as research to improve response, continued over an extended period of time.

Strategies to Improve Response The mail-out/mail-back decennial census in the U.S. may be viewed as the largest mail survey in the country, but, until 1990, little had been done to connect this aspect of the census with appropriate research literatures. The considerable drop in mail response from 1980 to 1990 was largely unanticipated by the Census Bureau, but in hindsight may have simply shown that cooperation with the decennial census was affected by the same factors leading to

declines in survey response generally. The Census Bureau recruited a leading researcher in mail surveys, Don Dillman from Washington State University. Dillman served in a senior technical position during 1991-1995, with continual contractual ties thereafter, leading a program of experimental research in questionnaire design and mail strategies. Devices proven effective in other mail surveys also were shown to increase response significantly in experimental tests with the questionnaire (Dillman, Clark, and Treat 1994). Contributions to the design of the census forms have also been made by other designers under contract.

The Dress Rehearsal essentially is based on the planned strategy to improve response. A remaining question, however, is whether to mail a second questionnaire to all households, without regard to whether the original form had already been received. The Dress Rehearsal used this blanket second mailing strategy, unlike previous tests which had mailed the second questionnaire only to nonrespondents. Discussions with commercial printers led the Census Bureau to conclude that selective mailing on the scale of the decennial census is not yet operationally feasible within the narrow time frame available. Dress Rehearsal experience will help to guide the decision on whether to perform a blanket mailing in 2000.

A second major change in outreach design to improve response was to contract advertising rather than to accept pro bono work through the Advertising Council. The budgeted \$100,000,000 is intended to result in more effective and targeted advertising than for the previous census. In addition, the Census Bureau is continuing a proactive program to form partnerships with governments and other organizations.

Largely, the planned changes in design, mailing strategy, and advertising are important components in making the Census Bureau's 2000 plan, because they will encourage many to respond directly by mail. Generally, these elements would also be included in a nonsampling census as well (9).

Service-Based Enumeration In 1990, the Census Bureau attempted to enumerate segments of the "homeless" population based on a night-time count. Evaluation studies had suggested the alternative of enumerating homeless through their use of services, including both soup kitchens and shelters (Martin 1992, Wellens and Gerber 1995). These procedures are included in the Dress Rehearsal tests (10).

Direct Sampling for Nonresponse Follow-Up (NRFU) The proposals regarded as viable for 2000 by the Census Bureau and the Academy panels incorporate an opportunity for all households to respond and be directly included in the census. For example, the Requirements Panel expressed the opinion that extending an opportunity

to respond to all households was probably required by the Constitution (Edmonston and Schultze 1995, pp. 240-241). Recent decennial censuses using mail-back procedures have followed up all nonresponding housing units. A number of alternatives to this traditional followup were considered before the 1995 test. For example, the report by the Requirements Panel (Edmonston and Schultze 1995, pp. 79) proposed dividing response into three phases:

- response to the initial mail census;
- unrestricted followup of nonrespondents by census enumerators completing the most accessible cases;
- sampling of the remaining cases, completing the sample, and estimating the nonsample cases.

In this design, the followup of the traditional census would be truncated, and the last and typically most difficult cases completed through sampling and estimation. In fact, after initial consideration of this design, the Census Bureau committed early to direct sampling, which eliminates the middle step by defining the entire followup work load through sampling. The 1995 Census Test employed direct sampling. Basically, the Census Bureau reached an internal consensus that the three-step process, particularly in the context of the 2000 census, would prove operationally unworkable and not have significant advantages over direct sampling in terms of cost, acceptability, and accuracy (11).

Reengineering of the PES The one-number census goal required incorporating the ICM results into the census in time to meet statutory requirements, the first of which is the delivery of state population totals to the President by December 31, 2000. Completing the smaller 1990 PES in time for the July 15, 1991 decision (which included producing the estimates by May and a report by June) had required extensive effort. The PES operations have now been redesigned as the ICM to meet the much tighter time demands of the one-number census. In particular, independent relisting of addresses in ICM sample blocks begins early, prior to the census, rather than waiting until the end of followup, as in previous studies. This is also followed by matching of housing units in April, substantially earlier than in previous studies. Some ICM interviews by telephone of census mail respondents also start after their forms are received. (Nonetheless, ICM personal interviews generally wait until the end of NRFU.) Technological improvements in matching and processing have been implemented.

The plan is to collect the initial interview for ICM through Computer-Assisted Personal Interviewing (CAPI) in order to obtain the benefits of instantaneous editing during collection and direct electronic transmittal of the data. These features were selected to improve the quality and timeliness of the ICM.

Some methodological changes to 1990 PES procedures have been driven by necessity, and we will reexamine their effect based on the Dress Rehearsal (12).

Reengineering the Census The Requirements Panel (Edmonston and Schultze 1995) advocated a conceptual shift from the 1990 dual strategy of a traditional census followed by a PES to an initial phase based on enumeration, assignment, and sampling followed by a tightly integrated ICM. Expensive coverage improvement operations incorporated in the previous censuses were to be eliminated in favor of cost-effective measurement in the ICM, wherever appropriate.

In addition to the use of sampling for NRFU and the use of sampling and a single followup of UAA/vacants, described in the following paragraphs, the initial phase has been simplified in other ways as well. As a consequence, the ICM might be expected to make a significantly larger correction to the initial phase estimate in 2000 than the 1.6% result obtained in 1990. Further description of operational aspects of the reengineered census have been provided elsewhere (U.S. Bureau of the Census 1997, 1998).

Unit Sampling for NRFU Early research on estimation for NRFU tended to be based on an assumption that the block would be the unit of sampling. In other words, the early candidate sample design for NRFU was to draw a sample of blocks in which all nonresponding units would be followed up and used to estimate the characteristics of nonsample blocks. The block sampling design was used in two of the test sites in 1995, but the Oakland, CA, site was divided into two experimental panels estimated through block and unit sampling. In unit sampling, a sample of nonresponding units would be sampled generally without regard to block.

Although there is arguably some cost advantage to block sampling, the variance advantages of unit sampling were far more substantial. Equally important, the Oakland results also indicated that the two methods appeared to have the same expected value. Consequently, the Census Bureau selected unit sampling for NRFU for 2000 and implemented it in Sacramento. There are the following favorable consequences:

- Unit sampling substantially reduces sampling variances at the block and tract level relative to block sampling, thus addressing concerns that the sampling census would adversely affect the quality of small area data.
- Unit sampling enables relatively simple estimation methodology to be implemented. In short, nonsample housing units will be imputed from sampled housing units in the nonresponse sample. This methodology is far easier to explain than the estimation alternatives under consideration for block sampling.

In the majority of cases, nonsample housing units will be imputed from a sample unit in the same block.

Unit sampling does impose significant requirements, however:

- To simplify matching and estimation, the ICM sample design employs block clusters as the unit of sampling, and NRFU operations will be carried out for all mail nonresponse units in ICM blocks. For comparability, therefore, block totals under the block sampling for ICM must have the same expected value under unit sampling.
- Unit sampling makes some coverage improvement operations more complex or inappropriate. For example, previous censuses included procedures to add housing units during NRFU. The actual adds in 1990 appeared to have been low. This avenue has to be curtailed to avoid creating a difference in expectation between complete nonresponse followup in the ICM blocks and unit sampling in the non-ICM blocks.

While the findings from Oakland currently represent the strongest evidence that we have to date on the effective equivalence of block and unit sampling, the Dress Rehearsal outcomes in Sacramento will represent a second, although more limited, opportunity. We plan to compare the estimates from NRFU and related initial phase operations (before consideration of ICM findings) for the ICM blocks and the remainder of the Sacramento site (13).

UAA Vacant Followup As part of the primary delivery of census forms, postal carriers are to designate regular housing units that they know to be vacant as "undeliverable as addressed, vacant." (There are also other categories of undeliverable forms, not discussed here.) If the postal determination were definitive, then no further action would be required. In previous censuses, however, census enumerators have often found these units to be occupied. In the 1990 census, UAA/vacant units were checked by two enumerator visits before being classified as vacant for the census. In the 1995 test, this issue was reexamined, but the findings were similar, with approximately 28% of units initially classified as UAA/vacant yielding occupied households (Green and Vazquez 1996). Consequently, a finding of the 1995 test was that followup by census enumerators, at least once on a sample basis, was critical.

A sampling rate of 10% was initially proposed, but the rate has been increased to 30%, used in the Dress Rehearsal in Sacramento, to avoid substantial variance coming from this one source.

Use of Dual-System Estimation A key research objective in both the 1995 Census Test and a smaller

1996 test in seven tracts in Chicago was to compare dual-system estimator (DSE) with a new strategy, called CensusPlus (14). DSE, which had been the basis for several survey-based coverage studies of the decennial censuses, including the 1990 PES, is based on an assumption of independence between the census and an independent sample of the population or *P sample*. An *E sample* is also selected from the census to estimate duplications and other forms of erroneous enumerations included in the census count. (Hogan (1993) describes the DSE employed in the 1990 PES and Waite and Hogan (1998) discuss details for Census 2000.) The DSE incorporates an estimate for persons not included in either the census or *P sample*. CensusPlus approaches the issue somewhat differently, attempting to obtain an improved list of the true Census Day composition of the household. After the census establishes the first list of members of the household, CensusPlus obtains a second list for the household through an independent interview and then forms the improved list through reconciliation of the census roster and the second list with the CensusPlus respondent. An initial attempt to implement CensusPlus in 1995 produced erratic results compared to more plausible results based on dual-system estimation. Because significant operational problems in the 1995 implementation of CensusPlus were discovered, a second test was repeated in 1996. By the time the test results showed that CensusPlus again appeared to fail, the Census Bureau had committed to the DSE in 2000.

We make two observations to account for the attraction of CensusPlus for the bureau, even though hindsight investigation of this alternative required considerable resources. The first of these is that CensusPlus attempted to eliminate the need for the estimate of persons never observed incorporated in DSE. If CensusPlus could be as successful, or almost as successful, as DSE, the Census Bureau saw this conceptual simplification as an advantage. Secondly, the CensusPlus may have provided a basis for constructing statistical corrections to the census files, so that the statistical corrections implied by the CensusPlus findings could be incorporated into detailed census data at all levels in a consistent manner. (Because the DSE includes estimates for persons not observed, construction of a file showing changes in household composition and number of new households from persons added by ICM is problematic. Our planned solution for 2000 is to categorize persons estimated to be present by ICM into a special adjustment cell.)

Sample Size for NRFU In presenting the evolving plans for 2000 to legislators and a wide public, Census Bureau staff observed that many judged the reasonableness of the sampling aspects of the plan not on the basis of technical authority but instead on intuitive notions about how

samples behaved. With further evidence from focus groups, the Census Bureau decided in 1995 to commit to a 90% completion rate for each tract. For example, in a tract with 50% initial response, a 4-in-5 sample of nonrespondents will be selected. The sampling rates have now also been set at 1-in-3 for tracts with initial response above 85% (15).

Obviously, this decision was consequential, substantially reducing, although not eliminating, the cost and time savings originally envisioned from NRFU sampling.

Dress Rehearsal The Dress Rehearsal in Sacramento incorporates all of the preceding elements of the 2000 plan. The Dress Rehearsal in South Carolina will establish the operational feasibility of conducting a census without the new sampling and estimation techniques. The South Carolina site does employ essentially the same questionnaire and response improvements as the other sites and a modified version of service-based enumeration.

All sites now use the multiple response option for race mandated by the Office of Management and Budget for Census 2000 and eventual adoption of the Federal statistical system. Although this question has been tested on national samples, this will be the first opportunity to observe its properties under census-like conditions.

We have also used the Dress Rehearsal to test pay scales for census enumerators hired on a short-term basis to complete NRFU. We return to this aspect in the concluding section.

A number of evaluations are planned for the Dress Rehearsal, including studies of the quality of the ICM interview and an error profile for the test censuses.

4. Issues Now Under Consideration

The major issues facing the Census Bureau's plans for Census 2000 fall broadly into three categories: legal, political, and technical. As noted in the introduction, the constitutionality of the plan is currently under litigation. We will simply note that the Census Bureau has proceeded on the basis of a legal opinion from the Department of Justice that the plan is both constitutional and in every other sense legal, but we will not add here to our previous brief remarks on this subject.

In general, it is inappropriate for us to comment on the complex political situation facing the plan. We will make two exceptions, however.

The potential for manipulation of the census totals for political reasons has on occasion entered the public discussion. There are two important reasons that manipulation of the census totals will not occur. The Census Bureau is staffed almost exclusively by career civil servants, who are committed to maintaining the Census Bureau's reputation for integrity. To avoid even

the appearance of manipulation, however, we recognize the necessity for communicating the specification for 2000 in advance, including those aspects for statistical sampling and estimation. The general specification is already available (e.g., Waite and Hogan 1998), although we expect to add refinements in the next months based on our Dress Rehearsal experience.

The second defense against manipulation is the openness with which we plan to conduct the process. We expect the Monitoring Board will actively review our work from an independent perspective. The National Academy of Sciences is moving forward on the Census Bureau's request to set up a technical review panel. We expect continued review from our Congressional oversight subcommittee, chaired by Congressman Dan Miller of Florida, the Office of the Inspector General, and the GAO.

A related issue is political in a very general sense. PL102-135 included the phrase, "public confidence in the census," which we see as an essential goal of the redesigned census. In addition to respect for the Census Bureau's professional integrity, public understanding and acceptance of the soundness of the census plan is critical. We return to this issue at the conclusion of this section.

The majority of this section focuses on remaining technical issues. The list of issues is not novel, since most have been raised previously by the Census Bureau and outside commentators, including the GAO (U.S. Government Accounting Office 1998). We review the current status for each issue and what evidence we expect in the next few months, as we integrate the findings from the Dress Rehearsal into a final plan for 2000.

Complexity By almost any standard, a decennial census is a complex operation. Nonetheless, the Census Bureau has largely succeeded at these efforts previously. Both the initial phase of the reengineered Census 2000 and the smaller but more complex ICM involve numerous operations, some of which have not been tested until the Dress Rehearsal. For example, the data capture system for the census is just being tested in the Dress Rehearsal.

Except for its shorter schedule, the initial phase for 2000 is not operationally more complex than previous censuses and in fact is in many respects simplified by the removal of some resource intensive coverage improvement operations.

Although smaller than the initial phase, the ICM is quite large by any other standard of survey research, and its size imposes constraints. As one example, the use of computers in ICM interviewing will require considerable technical support. Both the size of the ICM and its timing may constrain application of some procedures to assure high quality, such as the use of intense followup to reduce ICM nonresponse.

The combination of: 1) sampling for NRFU and UAA/vacant followup, 2) acceptance of late mail returns and Be Counted Forms after the cutoff date defining the samples, and 3) the ICM, generate a number of complex estimation issues. Waite and Hogan (1998) and Farber, Fay, and Schindler (1998) review the most important of these.

In the next months, the Census Bureau will reassess many of the operational and statistical plans in light of the Dress Rehearsal findings.

Sampling Variance For the most part, measures of reliability such as sampling variance are only used comfortably by statisticians and those with formal training in statistics. Thus, public discussions of the adjustment issue are rarely stated in these terms. Sampling variances are nonetheless important, however, because, if the ICM adjustments are subject to large sampling variances, they are likely to lack face validity once they become available.

Most of our projections of sampling variance for the ICM design have been based on the 1990 PES. For example, Schindler (1998) provides a proposal for allocating the 750,000 housing units of the ICM sample by state, achieving coefficients of variation of 0.5% or less for each state. His analysis of the 1990 data indicates that this level of reliability is adequate to provide a more accurate apportionment of the House of Representatives than the unadjusted 1990 census.

In these and similar calculations, estimates of expected reliability shape the Census Bureau's plans. Consequently, even before we discuss the accuracy of data at different levels, we comment here the current and future state of evidence on the ICM reliability.

Because we have redesigned both the census and the ICM, there is the possibility that the variance properties will be affected relative to the 1990 PES experience. For one, the reengineered census curtails some coverage improvement programs, with the consequence noted previously that the ICM correction is likely to be somewhat larger than the 1.6% of the 1990 PES.

Until the Dress Rehearsal, the redesigned ICM restricted searching to the sample block, unlike the 1990 PES and its predecessors. When only the sample block is searched, misassignment (geocoding errors) of housing units into neighboring blocks in the initial phase contributes to the gross error and, in turn, to the variance of the DSE. Indeed, because such errors can affect large numbers of units in a block simultaneously, this sort of error can have a substantially detrimental effect on the reliability. The Dress Rehearsal includes a test of a procedure to extend the searching for the most discrepant of these blocks. In turn, this experience will provide some evidence on how such procedures might be expanded for

2000 and what yield they may have in terms of variance reductions.

The Census Bureau has projected coefficients of variation of about 1.5% for both Sacramento and South Carolina. We will not have the results of direct variance estimation for the Dress Rehearsal until the end of this year, however. Whether or not the stated reliability goals of the Dress Rehearsal are met, it will be important to establish the necessary assumptions required to reconcile the observed reliability with our 2000 targets (16).

Substantial Reduction of the Total and Differential Undercount The Census Bureau plan represents an opportunity to obtain results for Census 2000 substantially reducing the total and differential undercount of previous censuses. The Census Bureau concurs with the conclusion of the Requirements Panel that there is no prospect of substantial reduction in the Black/non-Black differential without application of statistical methods. The 1990 PES sample size of approximately 150,000 households was adequate to measure this differential. Similarly, both the 1990 PES and the ICM sample size for measuring the national undercount of Hispanics will be adequate. For both groups, however, there will be an issue of how far the ICM data can be disaggregated geographically. For example, in a strictly state-based design, there will be too little data to support a separate Hispanic estimate within many states, and the proposed approach of combining races, such as Blacks with Hispanics, provides a less direct estimate than the separation of these groups in the 1990 PES.

There will be additional difficulty in measuring undercount for other racial groups. For example, in the 1990 PES, a separate poststrata were defined nationally for American Indians on reservations. To do so for 2000 requires departing from a strictly state-based design. Other groups, such as American Indians not on reservations, Asians, and Pacific Islanders are likely to lack adequate sample size for direct estimation in almost states.

Along the same lines, the Dress Rehearsal and Census 2000 will follow the new OMB policy permitting the reporting of multiple races. Tabulation of 63 combinations of races is possible and under consideration. A distinct analysis in the ICM of each of these possible groups is infeasible, and groups will be combined for purposes of the ICM (Waite and Hogan 1998). For example, a person reporting both white and Black for race may be treated in the ICM in the same way as if they reported only Black. The Dress Rehearsal will provide some useful data on some aspects of the problem, such as the consistency of reporting race between the initial enumeration and the ICM interview.

State Estimation As we noted in a previous collaboration (Fay and Thompson 1993), the 1990 PES sample design provided limited direct estimates of undercount by state. Consequently, it was not possible from the 1990 PES data alone to obtain an assessment of the assumptions underlying the estimates. We remarked:

Although models have an assured role in any departure from conventional enumeration, the 1990 experience suffered from too great a reliance on models. Even estimates for the highest and most important units, states, rested on assumptions that could not be satisfactorily verified from the PES data themselves. We sympathize to a degree with those who objected to the possibility that their state might be adjusted by 1990 PES data primarily collected in states other than their own. Consequently, the experience points to designing for adequate direct estimates for states, and potentially some substate units, with a satisfactory precision so that the state estimates could stand alone or be adequate to evaluate indirect estimates. This will probably require a precision on the order of a c.v. of about 0.5 percent or less for states.

These remarks were in the context of a paper including an empirical study the validity of the final (1992) PES model, based on 357 poststrata, in predicting proxy variables, such as mail return rate and unemployment. The findings suggested that the 357-poststrata model was moderately successful at state prediction, but might account for only 50% of the total variability between states, although results varied considerably depending on the choice of proxy variable (17). These data and their analysis were further discussed by Freedman and Wachter (1994) and Belin and Rolf (1994). The text of the Supreme Court ruling on the 1990 Census, cited previously, has considerably influenced the Census Bureau's position that the 2000 ICM should be based on direct state estimation (18).

As we just noted, the emphasis on direct state estimation is in partial conflict with the objective of measuring and correcting for the differential undercount by race and ethnicity. We will be examining the Dress Rehearsal data to evaluate the extent to which we can achieve the stated reliability targets for state estimates.

A further consideration in favor of direct state estimation is the possibility of completing processing for states separately. The design permits ICM estimation to be completed on a state basis, allowing some states to be completed to produce the final state count without waiting until all ICM processing is completed.

Substate Estimates As noted earlier, the Census Bureau's 1992 decision not to adjust the postcensal estimates with the 1990 PES was driven by the lack of statistically significant evidence of improvement at the substate level. Specifically, the Census Bureau primarily

considered the importance of relative shares within the state, rather than the accuracy of the overall totals (19). The formal tests employed represented extensions of the loss function analysis described by Mulry and Spencer (1993). In large part, the finding could have been driven by two important factors: 1) the specific 357-poststratum design, whose geographic stratification (urbanized area of 250,000 or more, other urbanized area, balance) probably increased the variances of the relative population estimates within state; and 2) the eventual 1990 PES sample size, which was only approximately half the size as originally planned. Selection of an estimator designed to reduce the variance of population shares combined with the larger ICM sample size should improve both aspects of this situation.

For the Dress Rehearsal, we are applying a raking of the adjustment factors (discussed in Farber, Fay, and Schindler 1998) in order to define more poststrata within state but to limit the resulting sampling variability of individual factors. The Census Bureau plans to investigate this approach further before committing to use it in 2000. Although relatively straightforward to a statistician, this procedure is more complex than the 357-poststrata approach used in 1990.

This issue will require further attention in the coming months. The details of ICM estimation for 2000 are still under review, and we have already noted the current uncertainty in projecting the ICM sampling variances. Detailed proposals for ICM estimation should be accompanied by calculations of expected reliability, to see if estimates for some substate areas will be subject to excessive variability from the ICM.

At the state level, it is clear that the dominant source of sampling variance will be from the ICM rather than NRFU and UAA/vacant estimation.

At the very lowest level of published geography, the census block, variance from NRFU and UAA/vacant estimation will dominate. The Dress Rehearsal will provide us with live data to estimate sampling variance at the block level combining the effects of our sample design and estimator, which had not been specified in detail when the Bureau issued its overall plan (U.S. Bureau of the Census 1997). Fay and Town (1998) suggest perhaps a 8% c.v. for a typical block of size 30 and 5% for a block of size 90, and they detail how the sampling variances will be estimated from the Dress Rehearsal data. In past censuses, geocoding error, such as misplacing an apartment building across the street, has been the dominant source for the occasional large census error at the block level. We consequently believe that control of error at the block level will depend much more on the success of our MAF reengineering work than the

contribution of sampling variance from NRFU and UAA/vacant sampling and estimation.

Adjustment for Net Overcounts Results from the 1990 PES and previous studies confirm that there are significant census overcounts from duplication and other sources affecting the census totals. Consequently, it is logically possible for the initial phase to overcount some poststrata in the population. Many of adjustment factors for the 357-design for the 1990 PES were less than 1.0, consistent with net overcounts.

When a computer file was prepared in 1991 to incorporate the adjustment, imputed persons representing estimated overcounts were placed in a cell with a weight of -1, although the original record of the person remained on the file as well, still as a member of the appropriate household. We have become aware of a potential public concern with the possibility that a person could be enumerated in the census and then subtracted out by such an action. To address the potential public perception that the Census Bureau might be removing directly counted persons from the total, we are now modifying our procedures. We now plan to implement all overcount adjustments by imputing persons to receive negative weights from other imputations. In other words, directly enumerated persons will not be offset by imputations with negative weights, but some imputations may offset other imputations. Most of the imputations offset by negative weights will be in nonsampled NRFU units. Staff are currently working through the details of this approach for incorporation into Dress Rehearsal.

Although some erroneous enumerations occur among mail respondents, previous studies have suggested a disproportionate share arising from nonresponse followup. Although more complicated, the new procedure may be arguably better than our previous treatment of this problem.

Presentation of Estimates Adjustment factors based on the ICM will be applied to estimated totals by poststrata down to the block level. The resulting added or subtracted people, when rounded to integers by a procedure to control the effect of rounding both at the block level and for poststrata within state, supplied by imputation. The imputations will be assigned and tabulated in a separate ICM adjustment category. This is similar to the device used for the 1990 census in preparation for the potential adjustment. Research to incorporate appropriate proportions of these imputed persons into households or forming distinct households from them was undertaken but is not sufficiently developed for implementation.

Correlation Bias in Dual-System Estimation As previously noted, demographic analysis has provided evidence of a pronounced undercount of adult Black males in both the census and the dual-system estimates

based on coverage studies. A similar but less pronounced bias is present for non-Black males. This problem appears as persistent as the differential undercount. Proposals to address this issue statistically, such as Alho, Mulry, Wurdeman, and Kim (1993), Fienberg, Glonek, and Junker (1993), or Zaslavsky and Wolfgang (1993) rely on untested assumptions or require significant operational development. As previously noted, the Census Bureau has decided against statistical models to combine the ICM in 2000 with demographic analysis, along the lines of the models described by Bell (1993), on the grounds of complexity and the lack of independent data to confirm the assumptions.

The Census Bureau's approach to this problem is to avoid letting perfection be the enemy of the good. Basically, all past experience indicates that ICM underestimation from correlation bias will affect groups, specifically Blacks, likely to receive large corrections through ICM for other reasons. Thus, the expected effect of correlation is to have the ICM undercorrect historically undercounted groups.

Missing Data in the ICM As noted previously, the high level of missing data in the 1980 PEP was a primary factor cited by the Census Bureau in deciding not to use the results for adjustment in any form. Accordingly, considerable effort was directed in 1990 at controlling missing data (Hogan 1993), with appreciable success. In part, permanent staff of the Census Bureau assisted in the effort to reduce nonresponse.

The Dress Rehearsal may help us to evaluate whether it is possible to maintain the 1990 PES levels of completeness. The larger 2000 ICM sample size sharply limits the degree that permanent interviewers will be able to complete the work and imposes timing constraints.

Movers In both 1980 and 1990, P-sample respondents were asked their Census Day address. Movers, with Census Day addresses different from the sampled addresses, were matched back to their reported Census Day address. Because of NRFU and UAA/vacant sampling, however, the 2000 ICM design was based on reconstructing Census Day households. Besides serving to limit matching requirements to sampled ICM blocks in which NRFU had been completed on a 100%-basis, this approach had the additional advantage of avoiding the complexities of geocoding the Census Day addresses for movers. On the other hand, the accuracy of response is an issue, since Census Day residents must be obtained by a proxy interview in almost all cases. (Waite and Hogan (1994) and Farber, Fay, and Schindler (1994) provide additional details.)

Other Sources of Nonsampling Error Mulry and Spencer (1993) account for other major sources of error affecting the 1990 PES, including response error in the P-

sample and E-sample interviews, matching error, and other operational error including estimation of the number of duplicates. In general, analogous sources will arise in the 2000 ICM. Two factors may help to reduce these sources: the tighter time schedule for the initial phase and ICM will move many interviews closer to Census Day; and automation of some activities, such as searching for duplicates. On the other hand, the large sample size will pose some challenge to maintain the quality achieved in 1990.

Conflicting Goals We conclude this section by noting the tension among the following three goals:

- Substantial reduction or elimination of the differential undercount by race and ethnicity from the census,
- Provision of direct state estimates,
- Use of methods that are readily understood by a wide public.

Emphasis placed on the first objective favors the creation of poststrata corresponding to those groups already separated in the 1990 PES estimates, Black, Hispanic, Indians on reservations, and Asian and Pacific Islanders. In fact, the proposed approach now distinguishes between Asians and Pacific Islanders (Waite and Hogan 1998). The Census Bureau has committed to direct state estimates. The Requirements Panel (Steffey and Bradburn 1994, p. 126) previously noted how more complex estimators might be designed to satisfy both goals simultaneously, similar in approach to the raking estimation now applied only within state. Most such alternatives increase the complexity of the estimator and conflict with the third goal.

The next months will be critical in the resolution of these issues. We are looking forward to the completion of the Dress Rehearsal and the reporting of our results over the next months.

* The authors are Associate Director for Decennial Census and Senior Mathematical Statistician, respectively. 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. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Census Bureau. It is released to inform interested parties of current research and to encourage discussion.

Because of limitations of space, both notes and references have been entirely omitted. A full version of the paper is available from the authors.