DEPARTMENT OF COMMERCE

Bureau of the Census

Accuracy and Coverage Evaluation; Statement on the Feasibility of Using Statistical Methods To Improve the Accuracy of Census 2000

AGENCY: Bureau of the Census, Department of Commerce.

ACTION: Notice.

SUMMARY: The Director of the Census has issued Accuracy and Coverage Evaluation; Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of Census 2000, his statement on the feasibility of using modern statistical methods to correct Census 2000 counts. The document sets forth the rationale for the Census Bureau's preliminary determination that (1) statistically corrected census data can be produced within the time frame required by law and (2) that statistically corrected data will be more accurate. The Secretary has adopted the Director's analysis and conclusions in a written decision forwarded to the Director. For public information, set forth below is Accuracy and Coverage Evaluation; Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of Census 2000, as well as three related memoranda (the Director's memorandum transmitting the document to the Secretary, the Secretary's memorandum to the Director, and a supporting legal opinion of the Commerce Department's General Counsel).

Authority: 13 U.S.C. 141, 13 U.S.C. 195.

William G. Barron,

Deputy Director.

June 12, 2000.

MEMORANDUM FOR The Honorable William Daley, Secretary of Commerce

Through: Robert Shapiro, Under Secretary for Economic Affairs

From: Kenneth Prewitt, Director

Subject: Accuracy and Coverage Evaluation: Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of Census 2000

Attached is my statement on the feasibility of using modern statistical methods to correct Census 2000 counts as stipulated by the U.S. Supreme Court's decision in *Department of Commerce* v. *United States House of Representatives* (January 1999). This statement was prepared after extensive discussions with the U.S. Census Bureau's senior staff and review of all relevant documents.

The Census Bureau is committed to making its data as accurate as possible for all uses. This document sets forth the rationale for the Census Bureau's preliminary determination that (1) statistically corrected census data can be produced within the time frame required by law and (2) that statistically corrected data will be more accurate.

ACCURACY AND COVERAGE EVALUATION

STATEMENT ON THE FEASIBILITY OF USING STATISTICAL METHODS TO IMPROVE THE ACCURACY OF CENSUS 2000

Table of Contents

Executive Summary

This document sets forth the rationale for the Census Bureau's preliminary determination that (1) it is feasible to produce statistically corrected census data within the time frame required by law and (2) the statistically corrected data will be more accurate.

Data from the decennial census are used to produce the state population totals for congressional apportionment. Additionally, detailed state data are used for redistricting, federal funds distribution, and other public and private sector purposes. Section 141(b) of Title 13 requires the Secretary of Commerce to report state population totals from Census 2000 to the President by January 1, 2001. Section 141(c) requires the Census Bureau to report redistricting data directly to the states by April 1, 2001.

The Census Bureau is committed to making its data as accurate as possible for all uses. In accordance with a 1999 Supreme Court ruling, the Census Bureau will not use statistical sampling to produce the state population totals used for congressional apportionment. Because the Census Bureau expects it can produce more accurate data by supplementing traditional enumeration procedures with statistical sampling, it plans to use these statistical methods to produce the more detailed data required for redistricting and federal program purposes.

Prior to April 1, 2001, the Census Bureau will have completed an enumeration of the American population, including a coverage measurement survey, that is designed to improve the accuracy of the initial counts. The coverage measurement survey, called the Accuracy and Coverage Evaluation (A.C.E.), is based on the established statistical method known as Dual System Estimation (DSE) and is designed to correct for missed individuals or erroneous enumerations in the traditional enumeration. The method of Demographic Analysis will also be used to evaluate the completeness of population coverage in Census 2000 at the national level, and to assess changes from previous censuses.

The operations used to produce the apportionment counts are designed with the goal of counting and correctly locating every individual residing in the United States on April 1, 2000, and also to count federal employees and their dependents living overseas as of that date. This goal cannot be completely and accurately realized. Every decennial census, from 1790 to 1990, has included in the census counts some who should have been excluded, and has missed some who should have been included. The first source of error leads to an overcount; the second source to an undercount. Every census for which the effect of these errors has been systematically measured has shown a net undercount—that is, the number of residents who were missed was greater than the number of erroneous enumerations.

Furthermore, in studies going back to 1940, the Census Bureau has documented and measured not only an overall net undercount, but also a higher net differential undercount for the Black population than for the non-Black population. Studies from the 1990 census also indicate differentially higher net undercounts for the Hispanic population and American Indians on reservations, compared to the White population. This persistent problem of differential undercounts is the most significant error for the population totals obtained through the traditional enumeration. As part of the operations for Census 2000, the Census Bureau will conduct the A.C.E., which is designed to improve census accuracy by increasing overall coverage and reducing the differential undercount. The A.C.E. also corrects for the smaller, though not insignificant, overcount that occurs when erroneous enumerations are included in the census.

The Census Bureau has determined that the A.C.E. is operationally and technically feasible and expects, barring unforeseen operational difficulties that would have a significant effect on the quality of the data, that these corrected data will be more accurate than the uncorrected data for their intended purposes. This determination is based on more than 20 years of Census Bureau research and experience with coverage measurement surveys using DSE and is supported by external experts in statistical methodology. From these years of experience, Census Bureau statisticians have a comprehensive understanding of the technical underpinnings of DSE. This understanding has guided the design of the A.C.E., allowing the Census Bureau to focus on the completeness and quality of the estimates of the population corrected for estimated net census error.

It is possible, though very unlikely, that problems with census operations could lead the Census Bureau to conclude that the data are not of sufficient quality for their intended purposes. These problems could occur in the operations leading to production of the apportionment counts and/or in the operations leading to the production of the corrected counts. This document does not address factors that the Census Bureau will consider in its determination that the apportionment counts are of sufficient quality to be used for their intended purposes. Because this document does focus on the feasibility of using statistical methods to improve the accuracy of Census 2000 for purposes subsequent to the production of apportionment counts, it discusses the review process for the final decision on whether to release statistically corrected data. This review process will be based on a determination of whether the A.C.E. operations were conducted in a way that met expectations. In the fall of 2000, the Census Bureau will present this review process to the statistical community and other interested parties.

Background and Overview

Census data are critically important in achieving equitable political representation and fair allocation of resources. Finding and enumerating approximately 275 million individuals in the correct location is, of course, an extremely challenging task. The traditional decennial census misses certain identifiable population groups at greater rates than others and therefore contains inherent inaccuracies. The Census Bureau designed the Accuracy and Coverage Evaluation (A.C.E.) using proven statistical methodologies to correct for this differential undercount and thereby make the census more accurate.

Uses of Decennial Census Data

The Constitution requires that a census of the nation's population be taken every 10 years to reapportion seats in the House of Representatives,[1] but the information provides more than just state-by-state population totals. State and local governments use census data to draw legislative districts of equal population to comply with the constitutional ''one-person-one-vote'' mandate and the statutory requirements of the Voting Rights Act. The federal government distributes billions of dollars in grants according to population-based formulae that rely on census data. Federal, state, local and tribal officials study the patterns of detailed census data before constructing hospitals, highways, bridges, and schools. Businesses, large and small, have come to depend on the Census Bureau's population, income, education, and housing data to make informed decisions about locating new offices, shops, and factories, and finding markets for new products and services. Census data also serve as definitive benchmarks for many of the household surveys conducted by federal agencies.

As will be explained in more detail below, the Census Bureau has designed the A.C.E. so that it will produce statistically corrected census data down to the block level. Census blocks are the ''building blocks'' employed by users of census data. The Census Bureau does not define the aggregations employed by data users; it provides the data that users can tabulate as needed for their programmatic purposes. For example, an administrator distributing funds under the Elementary and Secondary Education Act might need to distribute funds tabulated to school districts, which can range in size from large counties and cities to small towns and districts, while a state official responsible for redistricting might need to aggregate and re-aggregate census blocks into many different configurations to satisfy the requirements of the Voting Rights Act, 42 U.S.C. § 1971 *et seq.*

[1] *Constitution*, Art. I, Sec. 2, cl. 3.

The A.C.E. was designed to accommodate the needs of data users by allowing them to aggregate census blocks as appropriate for their particular program purposes. The accuracy of aggregated census data is more important than the accuracy of any particular block because data users rely on aggregated data, not block-level data.[2] Different types of accuracy and how they can be assessed at various levels of aggregation are reviewed below.

This recitation of the uses of census data illustrates the importance of taking as accurate a census as possible by reducing the differential undercounts of geographic areas and demographic groups. The belief that the census should be as accurate as possible has motivated the Census Bureau for more than 20 years to develop techniques to reduce the differential undercount.

The Differential Undercount

The Census Bureau has documented and measured a substantial differential undercount since the 1940 census.[3] After the 1940 census, Census Bureau statisticians and academic researchers refined a statistical technique known as Demographic Analysis, a technique that measures coverage trends as well as differences in coverage by age, sex, and race. Demographic Analysis uses records and estimates of births, deaths, immigration, emigration, and Medicare enrollments to develop estimates of the population at the national level, independently from the census. Demographic Analysis, though not without its errors, reveals the persistence of the differential undercount that exists between the Black and the non-Black populations. The following table illustrates this differential:

DEMOGRAPHIC ANALYSIS ESTIMATES OF PERCENTAGE NET UNDERCOUNT, BY RACE: 1940–1990

	1940	1950	1960	1970	1980	1990
Percent:						
Total	5.4	4.1	3.1	2.7	1.2	1.8
Black	8.4	7.5	6.6	6.5	4.5	5.7
Non-Black	5	3.8	2.7	2.2	0.8	1.3
Percentage Point Difference:						
Black/Non-Black	3.4	3.6	3.9	4.3	3.7	4.4

Source: J.G. Robinson and others, ''Estimates of Population Coverage in the 1990 United States Census Based on Demographic Analysis,'' *Journal of the American Statistical Association* 88, (September 1993): 1065.

The 1990 census revealed that the Black population was not the only group undercounted differentially. Children were much more likely than adults to have been undercounted in the 1990 census. While children under the age of 18 represented 26 percent of the total national population that year, they accounted for 52 percent of the net estimated undercount as estimated by the 1990 Post-Enumeration Survey (PES).[4] Another characteristic that affected the likelihood of being missed in the census was tenure, whether one rents or owns. Renters were more likely to have been left out of the 1990 count. The 1990 PES found higher undercounts among renters than for owners.[5] As the chart below demonstrates, a substantial differential undercount also was estimated in 1990 for Hispanics and American Indians on reservations:

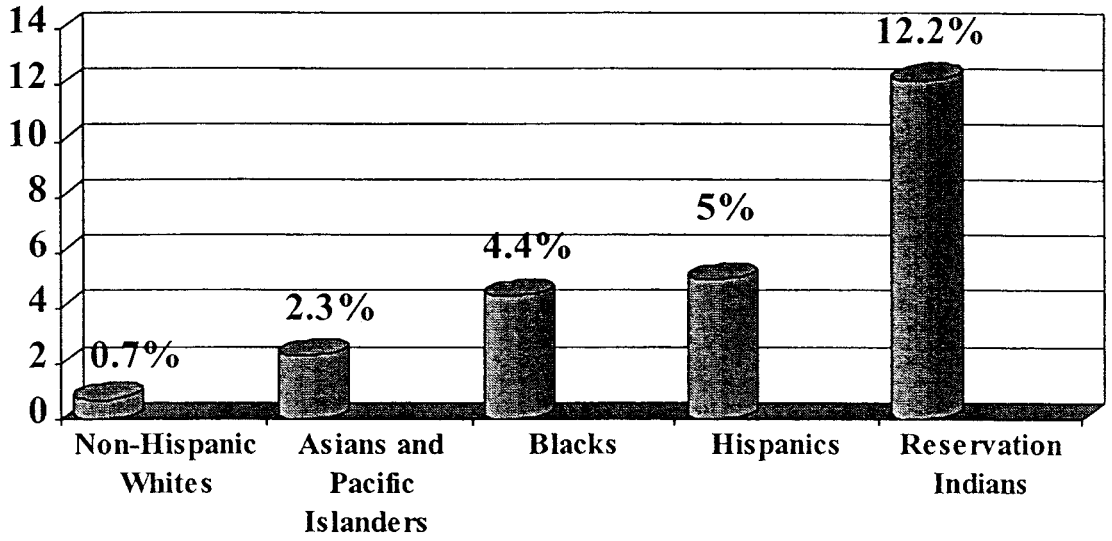
[2] The National Academy of Sciences agrees that accuracy at the block level is not an appropriate criterion of accuracy, that accuracy should be evaluated at aggregated levels. See Andrew A. White and Keith F. Rust, eds., *Preparing for the 2000 Census: Interim Report II* (Washington, D.C.: National Academy Press, 1997), 11–12.

[3] Bureau of the Census, ''Report to Congress—The Plan for Census 2000,'' originally issued July 1997, revised and reissued August 1997, 2–6.

[4] Ibid., 3.

[5] Howard Hogan, ''The 1990 Post-Enumeration Survey: Operations and Results,'' *Journal of the American Statistical Association* 88 (September 1993): 1054, Table 3.

Estimated 1990 Census Net Undercount Percent By Race[6]



Within each demographic group, the undercount for renters was considerably higher than for owners. For example, the estimated undercount was 6.5 percent for Black renters versus only 2.3 percent for Black owners.[7]

The differential undercount is a longstanding problem and one that the Census Bureau has not been able to solve despite increased efforts and resources. The National Academy of Sciences has calculated that the per housing unit cost of the census, in 1990 constant dollars, increased from less than $10 per housing unit in 1960, to $11 per housing unit in 1970, to $20 per housing unit in 1980, and to $25 per housing unit in 1990.[8] This steady increase in unit cost from 1960 to 1990, in large part due to increased efforts to reduce coverage errors, did not result in any appreciable reduction in the differential undercount.

The differential undercount clearly affects census accuracy. When identified areas and demographic groups are differentially undercounted, the relative population shares across states and sub-state areas are incorrect. Census data also provide the foundation for a large number of federal demographic statistics and household statistical surveys. These data are also extensively used by the private sector. Inaccuracies in the decennial census are carried over into these many other statistical series, and therefore, the persistent differential undercount has far-reaching consequences across public and private sector programs based on census data.

Summary of Census 2000 Operations

The Supreme Court determined in 1999 that Title 13 statutorily precludes the use of sampling to produce congressional apportionment counts.[9] Accordingly, the plan for Census 2000, as outlined in the Updated Summary: Census 2000 Operational Plan (February 1999), is to produce apportionment numbers without the use of statistical sampling by January 1, 2001. Rather than conducting the Integrated Coverage Measurement (ICM) survey [10] to produce statistically corrected numbers as part of the original Census 2000 plan, the plan now includes the A.C.E., which will produce statistically corrected numbers for non-apportionment uses of the data.

Within the constraint of the 1999 Supreme Court decision, the Census Bureau is committed to producing the most accurate data possible without the use of sampling for purposes of apportionment. The constraint does not apply to non-apportionment uses, and the Census Bureau also remains committed to producing the most accurate data possible for these other uses by implementing the A.C.E. As a prelude to the discussion of the A.C.E., this paper will briefly review basic census operations to be conducted prior to the A.C.E. A more extensive explanation of the operations for Census 2000 can be found in the Census 2000 Operational Plan.

The Census Bureau uses three basic data collection methods: mailout/mailback (where the Census Bureau mails questionnaires to housing units on the address list and the residents mail them back), update/leave (where Census Bureau workers deliver questionnaires at the same time they update the address list, and the residents mail them

[6] Bureau of the Census, ''Assessment of Accuracy of Adjusted Versus Unadjusted 1990 Census Base for Use in Intercensal Estimates,'' Report of the Committee on Adjustment of Postcensal Estimates,'' 7 August 1992, Attachment 3A, Table 2, later referred to as CAPE; and Bureau of the Census, ''Report to Congress—the Plan for Census 2000,'' 4.

[7] Hogan, ''The 1990 Post-Enumeration Survey: Operations and Results,'' 1054, Table 3.

[8] Barry Edmonston and Charles Schultze, eds., *Modernizing the U.S. Census* (Washington, D.C.: National Academy Press, 1995), 44. The cost for Census 2000 is currently estimated to be over $50 per housing unit in current dollars, indicating the increasingly greater cost of taking a census using traditional methods.

[9] *Department of Commerce* v. *House of Representatives*, 119 S.Ct. 765 (1999).

[10] The Census Bureau's original plan to use sampling was to conduct an Integrated Coverage Measurement (ICM) survey to produce a one-number census through the use of statistical sampling (''Report to Congress—The Plan for Census 2000,'' 29–32). The Census Bureau dropped its plans to conduct an ICM after the Supreme Court ruled that sampling could not be used to produce the apportionment counts.

back), and list/enumerate (where Census Bureau enumerators create the address list while canvassing their assignment areas and conducting interviews with respondents). Individuals can also respond to the census through the Internet or by telephone. The Be Counted program provides an additional means for people to be included in the census by allowing them to fill out a blank form made available in various public locations. Special enumeration procedures are followed for remote parts of Alaska, for locations containing a concentration of persons with a transient lifestyle (*e.g.*, trailer parks, marinas, and campgrounds), for group quarters (*e.g.*, prisons and long-term care facilities), and for people with no usual residence.

After allowing a reasonable amount of time for respondents to mail back their questionnaires, the Census Bureau conducts an operation called nonresponse followup (NRFU), which involves conducting a field followup of housing units that do not return their questionnaires by mail. A census enumerator will make up to six attempts to contact housing units that appear occupied to secure an interview. If an interview cannot be obtained, the enumerator attempts to interview a proxy respondent, that is, a neighbor, rental agent, building manager, or other knowledgeable individual.

A number of other operations are being implemented to ensure as complete coverage as possible in the initial enumeration. Computer edits are performed on mail-return questionnaires to identify those that may contain missing persons and those that contain large households (more than six persons). Interviewers conduct telephone interviews with these households during the coverage edit followup operation in order to obtain accurate data about the persons residing there. Another operation, coverage improvement followup, is conducted after NRFU. This operation includes an interviewer recheck of housing units classified as vacant or nonexistent during NRFU to ensure that no units have been misclassified. Finally, all major operations of the Census 2000 plan are subjected to enhanced quality assurance (QA) activities designed to detect and correct errors before they affect accuracy or data quality.

The Census Bureau also designed and implemented an enhanced marketing and partnership program that provides an integrated communications effort to increase both awareness of the decennial census and public cooperation. The marketing program is designed around the first-ever paid advertising campaign, including a national media campaign aimed at increasing mail response, targeted advertising directed at raising mail response among historically undercounted populations, and special advertising messages and campaigns targeted to hard-to-enumerate populations. In the partnership program, the Census Bureau is working nationwide with state and local partners to encourage all individuals to respond to the census.

After the data collection efforts have been completed, the data are processed through a number of computer operations for unduplication of multiple responses for the same housing unit and for editing of inconsistent or missing responses. For items that are not reported by respondents, the Census Bureau uses the statistical process of imputation to determine a response. The data are then tabulated, and the tabulations and other statistical aggregations are released.

The Accuracy and Coverage Evaluation Methodology

Following the initial census, the Census Bureau will conduct the A.C.E. Key components of the A.C.E. include the sample design, the survey itself, and the Dual System Estimation (DSE) used to compute the estimates of the true population.[11]

The A.C.E. in Brief

The A.C.E. methodology planned for Census 2000 involves comparing (matching) the information from an independent sample survey to initial census records. In this process, the Census Bureau conducts field interviewing and computerized and clerical matching of the records. Using the results of this matching, the Census Bureau will apply the statistical methodology of DSE (described below) to develop coverage correction factors for various population groups. The results will then be applied to the census files to produce all required Census 2000 tabulations, other than apportionment. The A.C.E. can be summarized as follows:

- Select a stratified random sample of blocks for the A.C.E.
- Create an independent list of housing units in the sample of A.C.E. blocks.
- Begin conducting telephone interviews of mail return housing units on a subset of the independent list.
- After the initial census nonresponse followup, conduct a personal visit interview at every housing unit on the independent list not already interviewed by telephone.
- Match the results of the A.C.E. interview to the initial census.
- Resolve cases that may not match but that require additional information by conducting a personal visit followup interview.
- Use information from other similar people to impute missing information.
- Categorize the A.C.E. data by age, sex, tenure, and other appropriate predefined variables into groupings called post-strata.
- Calculate the coverage correction factors using DSE, that is, determine the extent to which people in each post-stratum have been over- or undercounted by the initial census.
- Apply the coverage correction factors to correct the initial census data.
- Tabulate the statistically corrected census results.

The Sample Design

For the 2000 A.C.E., the Census Bureau selected a stratified random sample of blocks designed to be representative of racial and ethnic composition; tenure (owner or renter); and other variables. The sample consists of approximately 11,800 block clusters with approximately 314,000 housing units. The sample is designed to provide sufficient precision to estimate the true population for groupings of the population known as post-strata. Each person belongs to one and only one post-stratum. Post-strata are constructed with the goal of grouping individuals who have a similar probability of having been included in the initial census. Census 2000 post-stratification variables include race, ethnicity, age, sex, tenure, mail return rate, and metropolitan status/census enumeration method. For example, one post-stratum would include non-Hispanic Black males, aged 18–29, in non-owner units, in mailout/mailback areas of metropolitan statistical areas with 500,000 or more population, in tracts with a low mail return rate in the census. By comparing the estimated true population based on the dual system estimate for each post-stratum to the number of individuals counted in the initial census enumeration for each post-stratum, the Census Bureau estimates over- and undercounts for each post-stratum.

[11] A more extensive description of the A.C.E. can be found in Howard Hogan's paper, ''Accuracy and Coverage Evaluation: Theory and Application'', prepared for the February 2–3, 2000, DSE Workshop of the National Academy of Sciences Panel to Review the 2000 Census; and Bureau of the Census, ''Accuracy and Coverage Evaluation: Overview of Design'', by Danny R. Childers and Deborah A. Fenstermaker, DSSD Census Procedures and Operations Memorandum Series S–DT–02, 11 January 2000.

Conducting the Survey

Essential to the proper conduct of the A.C.E. is the need to ensure that the A.C.E. and the initial census are operationally independent. Independence requires that the probability of a particular household or person being included in the A.C.E. is not affected by the initial census operations and that the probability of people being included in the initial census is not affected by A.C.E. operations. Such independence is a critical criterion for DSE.

The A.C.E. independent interview is conducted by separately hired and trained staff through the use of Computer Assisted Personal Interviewing (CAPI) either by telephone or in person. CAPI is a method of data collection using a laptop computer in which the questions to be asked are displayed on the screen and responses are entered directly into the computer. The Census Bureau expects that the use of CAPI will improve the accuracy of the A.C.E. interview. To get an early start for the A.C.E. interviewing, where possible, a telephone interview using CAPI may be conducted for households where the census questionnaire has been completed and for which a telephone number was obtained. This activity is carried out concurrently with the initial census followup of nonresponse households. The door-to-door interviewing with CAPI does not begin until the initial census nonresponse followup is nearly completed in a given block cluster. The A.C.E. enumerators will attempt to secure an in-person interview with a household member. If the interview cannot be obtained, the enumerator will interview a proxy respondent.

After the A.C.E. independent interviews have been completed, computer matching between the initial census and the A.C.E. person records is carried out, followed by a clerical matching operation using an automated review system. The matching process allows the Census Bureau to determine who may have been missed by the initial census or to determine erroneous enumerations. It should be noted that the census can miss either entire households or individuals within households. This is also the case for erroneous enumerations.

The Census Bureau has carefully designed the A.C.E. to minimize matching errors. Incorrect matching generally results either from errors caused by incomplete, inaccurate, or conflicting data, or from errors where a poor match decision was made even though the data were sufficient. It is critical that the matching be as accurate as possible. Accordingly, as necessary, the Census Bureau conducts a personal visit follow-up operation to obtain the additional information needed to accurately code A.C.E. and census nonmatches. After this followup, the Census Bureau conducts a final clerical matching operation.

Even after this intense effort, occasionally some information will still be missing, either person characteristics, status of enumeration in the initial census, or match status for A.C.E. cases that could not be resolved. Before any calculations can be made to determine the estimated true population, missing person characteristics, initial census enumeration status, and A.C.E. match status must be resolved. Missing person characteristics such as age, race, sex, and tenure are statistically imputed from data reported for other household members or from similar households in the geographic area. For unresolved cases, the Census Bureau uses statistical imputation methodology to impute probabilities of being correctly enumerated or matched. The Census Bureau then estimates the true population by using these results in Dual System Estimation.

Dual System Estimation

DSE is an established and accepted statistical technique that is also referred to as ''capture/recapture.'' [12] Because the Census Bureau has conducted years of research into the likelihood that people of varying characteristics will be included in the census enumeration (this likelihood is known as inclusion probability), it is able to divide the nation's population into post-strata. Each post-stratum is defined so as to contain people with a similar probability of being included in the initial census. At the conclusion of the A.C.E. processes described previously, data are available for each post-stratum to calculate a dual system estimate.[13]

The dual system estimate is an estimate of the true population total for each post-stratum. The dual system estimates are then used to calculate a coverage correction factor for each post-stratum. The coverage correction factor is a ratio of the dual system estimate (the estimate of the true population) to the initial census count. These factors are then applied to correct the initial census data files. For example, if the coverage correction factor for non-Hispanic Black males, aged 18–29, in non-owner units, in mailout/mailback areas of metropolitan statistical areas with 500,000 or more population, in a tract with a low mail return rate in the census, is 1.02, then for every 100 such person records counted in the census in those areas, two numerical records will be added. Once these factors are applied, the corrected population estimates are created and tabulated.

Assessment of Feasibility

Section 195 of the Census Act states that ''the Secretary shall, if he considers it feasible, authorize the use of sampling,'' but the term ''feasible'' is not defined. As discussed in a legal opinion from the Department of Commerce's General Counsel, the Census Bureau understands this term in accordance with its ordinary meaning and the overall purposes of Title 13. It is important to note that even if Title 13 were silent as to the obligation to use sampling if feasible, the Census Bureau would apply criteria similar to those described below to determine whether to correct the census through the use of statistical sampling. The Census Bureau is committed to using reliable statistical methods if those methods can be expected to improve the overall accuracy of the census.

The Definition of Feasibility

The Census Bureau's determination that sampling is ''feasible'' is based on whether its use is possible, that is, compatible with other aspects of the census plan and with any statutory, timing, and funding constraints. Equally important, this determination is based on whether the use of sampling is expected to improve the overall accuracy of census data by improving overall coverage and reducing the differential undercount. These two components of the feasibility determination represent operational feasibility and technical feasibility. Can the Census Bureau produce the statistically corrected block-level numbers by the April 1, 2001, statutory deadline? Can the statistically corrected counts be expected to improve the overall accuracy of census data?

More specifically, in the context of Census 2000, the use of statistical sampling is feasible to correct the census if the two components of feasibility, operational and technical,

[12] Michael L. Cohen, Andrew A. White, and Keith F. Rust, *Measuring a Changing Nation—Modern Methods for the 2000 Census* (Washington, D.C.: National Academy Press, 1999), 31; and Kirk M. Wolter, ''Some Coverage Error Models for Census Data,'' *Journal of the American Statistical Association* 81 (June 1986): 338.

[13] Production of these estimates is discussed in more detail in Bureau of the Census, ''Accuracy and Coverage Evaluation Survey: Dual System Estimation,'' by Donna Kostanich and Richard Griffin, DSSD Census 2000 Procedures and Operations Memorandum Series #Q–20, 12 January 2000.

are satisfied. Operational feasibility refers to the Census Bureau's ability to conduct the A.C.E. with available resources and within required deadlines or time frames. Technical feasibility refers to the Census Bureau's expectation that the A.C.E. statistical methodology, if carried out as planned, will improve the accuracy of the census for non-apportionment uses of the data. As discussed below, the Census Bureau's extensive experience with coverage measurement surveys, including its incorporation of improvements since 1990, confirms the conclusion that the A.C.E. is both operationally and technically feasible.

Operational Feasibility

Operational feasibility refers to the Census Bureau's ability to conduct each major component of the census within applicable deadlines and with available resources. The Census Bureau expects to conduct each major component of the census, including the A.C.E., in time to meet the April 1, 2001, deadline for producing the redistricting data.

Release of Data Products for Use in Redistricting

The Census Bureau's goal is to produce the most accurate numbers possible within the constraints imposed by the federal statute and available resources. Section 141(c) of Title 13 requires the Census Bureau to deliver redistricting numbers to the states by April 1, 2001.[14] In past decennial censuses, the Census Bureau has been able to release redistricting numbers to certain states prior to the federal deadline, enabling redistricting officials in those states to meet deadlines set by state statutes and constitutions. The Census Bureau will, as in the past, release the numbers from Census 2000 to the states as they are ready, giving priority to states that need to meet early deadlines.

Operational Considerations

The Census Bureau's detailed plan for carrying out the entirety of the census operation, including the A.C.E., is set forth in the ''Master Activity Schedule'' (MAS).[15] This plan has undergone thorough reviews and analyses and supports the Census Bureau's confidence that it can implement the A.C.E. methodology correctly and successfully. The Census Bureau introduced its original Census 2000 plan in 1995. Since that time, the plan has been refined to incorporate testing, analysis, expert and other public input, and policy and programmatic changes, including the Supreme Court's January 25, 1999, ruling. During the last five years, the Census Bureau has put into place a comprehensive project management framework based on a powerful project management tool used by some of the world's largest private organizations. The use of this and other project tools, such as an integrated cost model and function and process modeling software, led to the Census Bureau's determination that it could produce the statistically corrected numbers by April 1, 2001. A revised Census 2000 MAS, reflecting this determination, along with the Census 2000 Operational Plan, were presented to the U.S. House of Representatives Subcommittee on the Census, as well as the Census Bureau's other oversight and appropriations committees and subcommittees, in March 1999.

Resource Considerations

Resources are also relevant to a feasibility determination. Based on current FY 2000 appropriations and the anticipation that the Administration's FY 2001 budget request for Census 2000 will be appropriated, the Census Bureau should be able to hire sufficient staff and acquire the necessary equipment to complete Census 2000 and produce statistically corrected redistricting numbers by the April 1, 2001, statutory deadline.

The Census 2000 Dress Rehearsal

In preparing for Census 2000, the Census Bureau, as has been its practice for many decades, conducted a dress rehearsal, or full-scale census simulation, in several sites across the country.[16] The dress rehearsal demonstrated the operational feasibility of producing the statistically corrected block-level data by the statutory deadline. The Census Bureau was able to produce data without the use of statistical sampling within nine months (as it is required to do for apportionment) and statistically corrected data within 12 months (as it is required to do for redistricting).

Technical Feasibility

Technical feasibility refers to whether the statistical methodology used by the A.C.E. will improve accuracy. Measuring the accuracy of the census is not a simple task. There are two types of accuracy—numeric and distributive—central to census operations and the uses of census data. Starting with the planning for the 1980 census, the Census Bureau has developed, tested, refined, and implemented statistical methods to improve both the numeric and distributive accuracy of the census enumeration, culminating in the 2000 A.C.E. design. The Census Bureau expects the A.C.E. to improve both numeric and distributive accuracy.

Defining Numeric and Distributive Accuracy

In analyzing the effect of the A.C.E. on accuracy, this discussion focuses on the accuracy of population totals for geographic areas and demographic groups, and, though important in the overall understanding of the census, not on the accuracy of detailed characteristic data for people or housing units.

Numeric accuracy refers to how close the overall count of a particular geographic area or demographic group is to the ''truth,'' that is, to the actual number of people who reside in that area or belong to that group. Distributive accuracy refers to how close the relative proportion or share of a geographic area or demographic group is to its true share relative to other areas or groups. A census operation that increases numeric accuracy moves the overall count for any particular area or demographic group closer to the true total. For example, an operation that enumerates individuals in a particular state who would otherwise be missed, increases the numeric accuracy of that state. A census operation that increases distributive accuracy will improve the accuracy of the population share for a given area or demographic group compared to other areas or demographic groups '' in other words, improve the accuracy of the estimated proportions or shares of the total population for the areas or groups.

A perfect census—one in which every resident is counted once and only once and is correctly located—would be both numerically and distributively accurate. But, as noted above, the Census Bureau's experience leads it to expect that, absent statistical correction, Census 2000 will result in both a net national undercount and various differential undercounts. Such undercounts affect both numeric and distributive accuracy. Although much of the analysis of the 1990 census focused on distributive accuracy, both types of accuracy are important and must be considered in designing a census that

[14] The Census Bureau's FY 1998 Appropriations Bill (P.L. 105–119) requires the Census Bureau, when it releases redistricting numbers based on statistical methods, to also release data produced without the use of statistical methods at all levels of geography.

[15] Bureau of the Census, ''Master Activity Schedule for Census 2000.''

[16] Sacramento, California; Menominee County, Wisconsin; and Columbia, South Carolina and 11 surrounding counties.

will provide the most accurate count possible.

Importance of and Relationship Between the Two Types of Accuracy

The decennial census can be viewed as one of the nation's most important civic ceremonies. Viewed in this broad perspective, securing maximum participation must be a key Census Bureau goal. To the extent that the census has the obligation to fully reflect who Americans are and how they live, everyone should be counted. Census operations that improve numeric accuracy, irrespective of their impact on distributive accuracy, meet this most basic goal.

In contrast, census operations that improved distributive accuracy but left many residents out of the count would not meet this basic goal. For example, a census that counted 90 percent of every demographic group in every geographic area would be distributively accurate, but would fail the obligation of the census to include everyone.

Numeric accuracy of census data is particularly important when population thresholds determine eligibility for program funding. For example, in FY 1998, the U.S. Department of Housing and Urban Development obligated over $3 billion under its Community Development Block Grants Entitlement Program. For this program, the population thresholds used are central cities of metropolitan statistical areas; other cities over 50,000 in metropolitan statistical areas; and qualified urban counties of at least 200,000 (excluding the population in entitlement cities located within the boundaries of such counties). Central city and metropolitan statistical area designations themselves depend upon certain population thresholds.

Additional uses of census data for which numeric accuracy is critical are associated with the Census Bureau's intercensal population estimates and survey controls. Decennial census data are the base for the Census Bureau's intercensal population estimates and projections programs, programs that produce annual population estimates for all general purpose governments and projections for the nation and states, respectively. Specific uses of the population estimates that depend on numeric accuracy include use of the estimates as controls for many federal surveys, including the Current Population Survey (which provides monthly labor force and employment data), and as denominators for many critical federal data series, such as birth, mortality, and cancer rates, as well as per capita income.

For the purpose of reapportioning seats in the House of Representatives, distributive accuracy becomes a principal concern, because reapportionment is based on a proportionate allocation formula. Federal and state redistricting are based on criteria for dividing state populations into districts of equal size; thus both numeric and distributive accuracy are important. Distributive accuracy is also central to federal funding allocations that distribute funds based on relative percentage of the population.

The goal of the Census Bureau is to conduct a census that is both numerically and distributively accurate. This said, it is numeric accuracy that drives the process for designing Census 2000 operations other than the A.C.E. When it designs a decennial census, the Census Bureau has available a very large number of possible operations. It assesses these operations against such criteria as cost, statutory deadlines, whether the staff necessary to implement these operations can be recruited and adequately trained, and how well the operations fit with other operations under consideration. In this extensive process of evaluating individual operations and then assembling them in the final design, there is one paramount criterion: what census design has the highest probability of correctly enumerating the population? That is, can an operation considered separately, and when combined with other operations, be expected to help the Census Bureau correctly count as many people as possible, given funding, timing, and other constraints?

Obviously, if perfect numeric accuracy were achieved for all geographic areas and demographic groups, then perfect distributive accuracy would also result. However, because it is difficult and perhaps impossible to know a priori the effects of a particular census operation on distributive accuracy, assessing an operation's effect on distributive accuracy can rarely be part of the planning process. The difficulty of designing operations for distributive accuracy is compounded if it is to be achieved across geographic areas and multiple demographic groups and then simultaneously across many levels of geography. For example, the Local Update of Census Addresses program, being voluntary, may have benefitted communities with strong local planning departments more than other communities. This program, then, had an unpredictable effect on distributive accuracy.

In principle, any given census operation designed to increase numeric accuracy can increase distributive accuracy, leave it the same, or make it worse. But in assembling a census design, the Census Bureau does not reject operations that would improve numeric accuracy (and meet other criteria for inclusion) even though such operations might affect distributive accuracy negatively, or indeterminately. For example, the Census Bureau has developed for Census 2000 an extensive partnership program to assist local jurisdictions and community organizations in promoting participation in the census. But increasing the counts for these participating localities will not necessarily translate into improvements in distributive accuracy. If one state promotes the census more effectively than another state, the state with the better promotion program may earn a higher share of the total national population than would otherwise be the case.

Although the Census Bureau has largely targeted its coverage improvement programs in the areas that have been the most difficult to count, it has not rejected census operations that might disproportionately improve the count for groups that are already well counted. An example of the latter in Census 2000 is the ''New Construction'' program.[17] Moreover, the Census Bureau has supported the efforts of neighborhoods, cities, and states to increase the accuracy of their census counts, irrespective of the effect on distributive accuracy. The Census Bureau views these increases in numeric accuracy, even for well counted groups, as important to the most basic goal of the census—counting everyone.

Finally, although different uses of census data depend to varying degrees on each type of accuracy, the two concepts are related. When the census falls short of overall numeric accuracy, states and localities with large populations that are differentially undercounted will suffer a diminution in proportionate shares. For example, the differential undercount in the 1990 census caused states and localities with large minority populations to suffer a diminution in share. The Census Bureau can and does try to improve both

[17] In this program, local and tribal governments liaisons in mailout/mailback areas review the Census Bureau's address list for their areas and provide the agency with the addresses of all newly constructed housing units as of April 1, 2000. The Census Bureau matches these addresses to its address list, updated with United States Postal Service files, and verifies and enumerates those addresses that are not on its address list.

numeric and distributive accuracy by bringing the total count for each area or demographic group closer to its true count.

Impact of the A.C.E. on Accuracy

The preceding discussion of accuracy included a discussion of the design of the census for apportionment purposes, but did not consider the effects of the A.C.E. on numeric and distributive accuracy. As discussed below, the A.C.E. measures and corrects for the deficiencies in the initial census, and consequently the Census Bureau expects that the A.C.E. will improve both distributive and numeric accuracy.

Based on decades of research identifying and measuring the undercount, as well as the 1990 census evaluations (discussed below), the Census Bureau expects the differential undercount to persist in Census 2000, with properties similar to those measured in 1990.[18] This extensive research into measuring and correcting the differential undercount, augmented by enhancements to prior coverage measurement surveys, leads the Census Bureau to expect that the A.C.E. will improve accuracy. The A.C.E. is expected to improve numeric accuracy by moving total counts closer to the true count and to improve distributive accuracy by more accurately counting areas that contain significant populations of historically undercounted groups.

It is important to consider the contribution of the A.C.E. to numeric and distributive accuracy at different levels of geography. The Census Bureau expects that the A.C.E. will, on average, improve numeric accuracy for geographic areas down to and including census tracts.[19] ''On average'' means that, while some tracts will be more numerically accurate using uncorrected numbers and others more accurate using corrected numbers, the average effect over all tracts is greater accuracy with than without the A.C.E. The Census Bureau also expects that improvement will be greatest for those areas that contain groups that have been historically undercounted.

Regarding distributive accuracy, the Census Bureau's extensive evaluations following the 1990 census led it to conclude that the 1990 PES would have, on average, increased distributive accuracy for larger geographic areas, including states and cities and counties with more than 100,000 people. These evaluations did not determine whether the 1990 PES would have improved the distributive accuracy of smaller geographic areas.[20] In addition, these evaluations did not address whether the unadjusted counts were more accurate for these areas. The research on these issues conducted by the Census Bureau's Committee on Adjustment of Postcensal Estimates (CAPE) is discussed more fully below. Based on this research, the Census Bureau expects the incorporation of the A.C.E. results in the Census 2000 counts to have a similar effect, that is, to improve distributive accuracy for larger geographic areas, as in 1990.

Historical Experience With Coverage Measurement Surveys Demonstrates Feasibility

The Census Bureau has a longstanding practice of employing scientific sampling techniques in the decennial census whenever sampling has the potential to lower costs without negatively affecting quality.[21] It has devoted substantial resources for over two decades to the development of coverage measurement programs employing high quality sampling methodologies that enable the production of more accurate data. The feasibility assessment discussed in this document is one more logical step along that continuum.

The Census Bureau and leading professional statistical organizations have concluded that the best way to address the persistent problems of the undercount and the differential undercount is to complement traditional enumeration procedures with scientific sampling, using DSE. Extensive research, testing, and refinement of the tools of statistical adjustment have led the Census Bureau to determine that the A.C.E. will improve the overall accuracy of the census.

The Census Bureau also has used Demographic Analysis to evaluate coverage in decennial censuses and broadly validate the coverage measurement survey results.[22] Since independent Demographic Analysis estimates are not available below the national level, nor have estimates been available for detailed demographic groups (for example, tenure or detailed racial groups), the Demographic Analysis method has not been used to adjust the census for undercoverage.

The 1980 Census Experience

Development of the modern coverage measurement survey began with the 1980 Post Enumeration Program, or PEP.[23] The PEP was a coverage measurement survey, based on DSE methodologies, designed to evaluate the accuracy of the 1980 census. Over 50 lawsuits were filed regarding the 1980 census, most contending that the results of the PEP should have been used to adjust the census. However, the PEP had been designed primarily as a coverage evaluation tool, rather than an adjustment mechanism, making its use to correct the census results problematic. The Director of the Census decided not to adopt the numbers produced from this first attempt at statistical correction using DSE, judging the estimates to be flawed by missing and inaccurate data.[24]

Significantly, however, the PEP operation provided a wealth of information on measuring coverage in a census using DSE. The PEP illustrated the potential use of coverage measurement surveys as a coverage evaluation tool for U.S. censuses. It was clear in principle that coverage measurement surveys could be used to correct the census. In the two subsequent decades, the Census Bureau built upon the knowledge and experience gained in the 1980 census.

Early Research and Development for the 1990 Census

After the 1980 experience, the Census Bureau began an extensive review of its coverage measurement program to enhance the methods that had been used in 1980 and to determine the feasibility of a statistical adjustment in 1990. Adjustment of the census was a topic of lively debate in the statistical

[18] This conclusion is based on the assumption that the coverage improvement programs used for Census 2000 will have similar results as those used in 1990.

[19] Bureau of the Census, ''Report to Congress—The Plan for Census 2000,'' 44–46. Census tracts are small, homogeneous, relatively permanent statistical subdivisions of counties formed for the purpose of collecting and tabulating decennial census data. Tracts typically contain between 1,000 and 8,000 people.

[20] When apportionment is calculated based on Census 2000 counts, the average Congressional District is expected to be over 600,000 people.

[21] The Census Bureau first used sampling in a decennial census in 1940, in the program now known as ''long form'' enumeration, which is used to obtain detailed demographic information. The Census Bureau has used sampling to conduct federal surveys to collect key information, including unemployment and labor force data, etc., for many decades.

[22] J.G. Robinson and others, ''Estimates of Population Coverage in the 1990 United States Census Based on Demographic Analysis,'' *Journal of the American Statistical Association* 88 (September 1993): 1061–77.

[23] For a detailed discussion of the 1980 Census Post Enumeration Program, see Robert E. Fay and others, *The Coverage of Population in the 1980 Census* (Washington, D.C.: Government Printing Office, 1988), 37–92.

[24] Department of Commerce, ''Position on Adjustment of the 1980 Census Counts for Underenumeration,'' **Federal Register** (16 December 1980) vol. 45, no. 243, p. 82872.

community during the 1980s. Census Bureau professionals and outside statisticians published more than 100 papers on coverage measurement issues.[25] In 1983, the Census Bureau formed the Undercount Research Staff, a staff of agency professionals charged with addressing coverage measurement issues and with assessing the potential correction of the 1990 census. This group conducted research over the decade leading up to the 1990 census.

Planning for the 1990 census progressed with a two-track approach—preparing to take the best traditional enumeration possible, while simultaneously developing a Post-Enumeration Survey (PES), a coverage measurement survey the results of which could be used to statistically correct the census. The Census Bureau's position was that it would proceed with correction if it could determine, prior to the spring of 1987, that implementation of a PES-based correction was feasible. As part of its research effort, the Census Bureau carried out the Test of Adjustment Related Operations (TARO) in 1986. Based on the results of the TARO, as well as various theoretical and empirical studies conducted since 1980, senior statisticians at the Census Bureau concluded that statistical methods existed that could produce census counts with a reduced differential undercount, and that if funded and successfully completed, the program incorporating these methods could be used to statistically correct the 1990 census.[26]

As discussed below, the Department of Commerce overruled the Census Bureau and decided not to allow adjustment of the 1990 census. The Census Bureau's research on the PES as a coverage measurement tool continued, including the conduct of the 1988 dress rehearsal Post-Enumeration Survey. The 1988 dress rehearsal demonstrated significantly improved operations and once again demonstrated DSE's consistent ability to measure the undercount and the differential undercount.[27]

Litigation Challenging Decision to Halt 1990 Adjustment-Related Planning Activities

The Department of Commerce, in the fall of 1987, directed the Census Bureau not to proceed with its plans to produce adjusted census figures, prompting the filing of a lawsuit against the Department and the Census Bureau. As part of that lawsuit, on July 17, 1989, the Department of Commerce entered into a stipulation, vacating the Department's 1987 decision against adjustment and requiring the Secretary to consider *de novo,* after the completion of the census, whether adjustment was warranted. The Census Bureau would conduct a PES and certain other adjustment-related planning operations,[28] and the Secretary was to announce his decision on the adjustment issue by July 15, 1991. Pursuant to the stipulation, the Department of Commerce agreed to develop and adopt promptly ''guidelines articulating what defendants believe are the relevant technical and nontechnical statistical and policy grounds for the decision on whether to adjust the 1990 Decennial Census population counts.'' An adjustment would be made if the Secretary of Commerce, in his judgment, determined that doing so would satisfy the guidelines. The stipulation also set up a Special Advisory Panel composed of four experts chosen by the plaintiffs and four experts chosen by the defendants; the Panel's role was to advise the Secretary regarding adjustment. At this time, the Census Bureau convened the Undercount Steering Committee, a group of senior career agency employees, and charged the committee with evaluating the conduct of the 1990 PES and assessing the accuracy of the adjusted versus the unadjusted census counts.

The Department of Commerce published its final guidelines on March 15, 1990.[29] The guidelines established, among other things, the principle that the unadjusted census counts would be presumed more accurate unless it could be shown that the adjusted counts were more accurate at the national, state, and local levels. This presumption and the guidelines in general will be discussed in greater detail below.

Conducting the 1990 Census and Deciding Against Adjustment

The Census Bureau applied the DSE methodology in the 1990 PES to produce a second set of population counts for every block in the nation.[30] Under the direction of the Undercount Steering Committee the Census Bureau analyzed the PES results extensively, producing 33 separate and detailed technical reports analyzing various aspects of the survey and its results. The Census Bureau's extensive analysis was complemented by a large volume of outside expert analysis of the PES results.[31]

Based on the Census Bureau's analyses, then Census Bureau Director Barbara Bryant and the majority of the Undercount Steering Committee recommended that the 1990 census be statistically adjusted. The Special Advisory Panel, convened as part of the stipulation, was divided in its recommendations regarding adjustment. The panel members selected by defendants all recommended against statistical adjustment, and the panel members selected by the plaintiffs all recommended in favor of adjustment. On July 15, 1991, Secretary Mosbacher announced that the 1990 decennial census would not be statistically adjusted.[32]

After the Secretary announced his decision, the plaintiffs returned to court, seeking an order compelling the Department to adjust the 1990 census. On April 13, 1993, Judge McLaughlin of the U.S. District Court upheld Secretary Mosbacher's decision, determining that the decision was not arbitrary or capricious, although he stated that ''were this Court called upon to decide this issue *de novo,* I would probably have ordered the adjustment.''[33] Judge McLaughlin noted also that ''light of recent improvement in statistical tools and the practical benefits that the 1990 PES has provided, the use of adjustment in the next census is probably inevitable.''[34]

[25] See Tommy Wright and Joyce Farmer, ''A Bibliography of Selected Statistical Methods and Development Related to Census 2000,'' 3rd ed., 1 May 2000, for a list of many of the most significant of these papers.

[26] Dan Childers and others, ''The Technical Feasibility of Correcting the 1990 Census,'' in *Proceedings of the Social Statistics Section of the American Statistical Association Held in San Francisco, California, 17–20 August 1987.*

[27] Dan Childers and Howard Hogan, ''The 1988 Post Enumeration Survey Methods and Preliminary Results,'' in *Proceedings of the Survey Research Methods Section of the American Statistical Association Held in New Orleans, Louisiana, 22–25 August 1988.*

[28] Prior to the Department's 1987 decision halting the Census Bureau's adjustment-related planning activities for the 1990 census, the agency had planned to conduct a PES of 300,000 housing units. Under the terms of the stipulation, the Census Bureau agreed to conduct a PES of approximately 165,000 housing units, the results of which could be used to adjust the census.

[29] Department of Commerce, ''Final Guidelines for Considering Whether or Not Statistical Adjustments of the 1990 Decennial Census of Population and Housing Should be Made for Coverage Deficiencies of the Population,'' **Federal Register** (15 March 1990) vol. 55, p. 9838.

[30] Hogan, ''1990 Post-Enumeration Survey,'' 1054.

[31] Wright and Farmer, ''A Bibliography of Selected Statistical Methods Related to Census 2000.''

[32] Department of Commerce, ''Adjustment of the 1990 Census for Overcounts and Undercounts of Population and Housing: Notice of Final Decision,'' **Federal Register** (22 July 1991) vol. 56, p. 33583.

[33] *City of New York* v. *U.S. Dept. of Commerce,* 822 F. Supp. 906, 928 (E.D.N.Y. 1993).

[34] 822 F. Supp. 906 at 918, fn. 27.

Postcensal Estimates and Survey Controls Decision

Although Secretary Mosbacher determined not to adjust the 1990 census for estimated net census undercount, he deemed it appropriate that the Census Bureau consider using the adjusted counts as the basis for producing postcensal estimates: ''I am today requesting that the Census Bureau incorporate, as appropriate, information gleaned from the Post-Enumeration Survey into its intercensal estimates of the population.'' [35]

Census Bureau Director Bryant convened the Committee on Adjustment of Postcensal Estimates (CAPE) to study this issue and make recommendations to her. CAPE was a group of senior statisticians, demographers, and other Census Bureau professionals assembled to conduct additional analyses of the adjusted counts. The Committee's work extended over a 15-month period. The Committee issued a report on August 8, 1992, and an Addendum on November 24, 1992. [36] The Addendum was the result of continuing and more focused analysis by the team. Taken together, the initial CAPE report and the Addendum found that the adjusted numbers were overall more accurate in terms of distributive accuracy at the state level and for areas with greater than 100,000 population. For areas with populations of less than 100,000, the CAPE could not identify any improvement in distributive accuracy for the adjusted data. [37]

In January 1993, Dr. Bryant announced that the Census Bureau would not use the 1990 adjusted counts as the basis for producing postcensal estimates of the population. [38] Director Bryant's Census decision was made in light of, though not explicitly governed by, the litigation guidelines that stated that adjustment was not warranted unless improvement could be clearly demonstrated down to small levels of geography, such as places and counties.

Recognizing the improvements in accuracy for certain uses of census data, Dr. Bryant decided to offer sponsors of federal sample surveys the option of having their surveys calibrated to population estimates benchmarked to adjusted census results. [39] Accordingly, in December of 1993, the Bureau of Labor Statistics (BLS) requested that the Census Bureau convert the Current Population Survey controls to ones based on estimates incorporating the results of the 1990 PES. The BLS stated its conviction ''that the undercount-adjusted estimates provide a more accurate reflection of the level and distribution of the national population and that of most States than the estimates based on the raw Census counts.'' [40] The BLS also requested that the population controls for the Consumer Expenditure Survey be adjusted in a similar fashion. [41] Subsequent to the BLS decision, all other major national household surveys conducted by the Census Bureau for other agencies of the federal statistical system were converted to an adjusted population basis. Thus, corrected data from the 1990 census are already incorporated in many federal statistical series.

Early Census 2000 Planning

The results of the 1990 census led the Census Bureau, other professional statisticians, and Congress to conclude that significant changes were required for the next census. A comprehensive re-examination of census methodology was needed to identify a census design that would improve the accuracy of the census. To this end, in November 1990, the Census Bureau established the ''Task Force for Planning for the Year 2000 Census and Census-Related Activities for 2000–2009.'' The Task Force was responsible for defining a census design for Census 2000, considering both policy and technical issues, and a demographic measurement system for related activities for 2000 through 2009.

In June 1992, the General Accounting Office (GAO) released a comprehensive evaluation of the 1990 census, discussing lessons learned and identifying opportunities for fundamental, effective reforms. The GAO concluded that reduced data quality (including failure to make reductions in the net and differential undercounts)'' * * * is a cost of the current approach to taking the census * * *'' and that ''[t]he results from 1990 demonstrate that adding more resources [while employing traditional census-taking methods] is unlikely to allow the Bureau to enumerate that last remaining segment of the population.'' [42]

Also at the beginning of the decade, two panels of the National Academy of Sciences' (NAS) National Research Council were convened to study ways to improve the census for 2000. The Decennial Census Improvement Act of 1991, signed into law by President Bush, required the Census Bureau to contract with the National Academy of Sciences to study * * * the means by which the Government could achieve the most accurate population count possible * * *'' specifically considering, among other things, ''. . . the appropriateness of using sampling methods in combination with basic data-collection techniques or otherwise, in the acquisition or refinement of population data, including a review of the accuracy of data for different levels of geography * * *.'' [43] The Panel on Census Requirements in the Year 2000 and Beyond was established pursuant to this statutory requirement, supplementing the work already being performed by the NAS Panel to Evaluate Alternative Census Methods. This latter panel was established to provide an independent review of the technical and operational feasibility of the design alternatives and of the tests to be conducted by the Census Bureau. The Methods Panel's recommendations on testing and design alternatives informed the final design of the original plan for Census 2000. The Panel issued its final report in 1994, recommending that the agency use sampling as an essential part of census-taking in Census 2000. [44]

In June 1995, the Task Force convened at the beginning of the decade issued final recommendations in its ''Global Report,'' suggesting a number of

[35] Department of Commerce, ''Adjustment of the 1990 Census,'' 33582.

[36] CAPE; ''Additional Research on Accuracy of Adjusted Versus Unadjusted Census Base for Use in Intercensal Estimates,'' Addendum to Report of the Committee on Adjustment of Postcensal Estimates, 25 November 1992, referred to later as CAPE Addendum.

[37] As will be discussed more fully below, more recent research has confirmed that the Census Bureau similarly cannot determine that the uncorrected 1990 data were more distributively accurate. For aggregations below 100,000, the evidence as to accuracy is indeterminate, that is, neither favoring the unadjusted nor the adjusted counts. See Bureau of the Census, ''Analysis of CAPE Findings on PES Accuracy at Various Geographic Levels,'' by Sally M. Obenski and Robert E. Fay, 9 June 2000.

[38] Department of Commerce, ''Decision of the Director of the Bureau of the Census on Whether to Use Information from the 1990 Post Enumeration Survey (PES) to Adjust the Base for the Intercensal Population Estimates Produced by the Bureau of the Census,'' **Federal Register** (4 January 1993) vol. 58, no. 1, 69.

[39] Ibid., 70.

[40] Katharine G. Abraham, U. S. Bureau of Labor Statistics, to Harry A. Scarr, U. S. Bureau of the Census, 1 December 1993.

[41] Ibid.

[42] General Accounting Office, *Decennial Census: 1990 Results Show Need for Fundamental Reform,* Report to Congressional Requesters, 9 June 1992, 62, 49 GAO/GGD–92–94.

[43] Congress, House, *Decennial Census Improvement Act of 1991,* 102nd Cong., 2nd sess., H.R. 3280, *Congressional Record,* daily ed. (9 October 1991), H7694 became Public Law 102–135 on October 24, 1991. It was set forth in the commentary to Title 13, U.S. Code, sec. 141.

[44] Duane L. Steffey and Norman A. Bradburn, *Counting People in the Information Age* (Washington, D.C.: National Academy Press, 1994), 4.

improvements for Census 2000.[45] The Task Force endorsed the Census Bureau's basic plan to conduct an Integrated Coverage Measurement (ICM) survey and suggested that the Census Bureau pursue an ICM design that would incorporate the best features of alternative methodologies, including DSE and CensusPlus.[46] The Census Bureau tested these alternate methodologies in the 1995 Census Test, concluding along with the NAS Panel to Evaluate Alternative Census Methodologies that DSE offered the best opportunity to produce high quality statistical correction.[47]

Also in that year, the Panel on Census Requirements in the Year 2000 and Beyond issued its final report. The Panel recommended the use of sampling and estimation techniques in Census 2000, concluding that:

> It is fruitless to continue trying to count every last person with traditional census methods of physical enumeration * * *. It is possible to improve the accuracy of the census count with respect to its most important attributes by supplementing a reduced intensity of traditional enumeration with statistical estimates of the number and characteristics of those not directly enumerated.[48]

The Census 2000 Dress Rehearsal

The Census Bureau conducted a dress rehearsal in 1998 in several sites across the country, an important opportunity to test the DSE methodology in as near a census-like environment as possible. The Census Bureau concluded from the dress rehearsal results that ''[t]he data showed across-the-board that the undercount, which has been measured in every census since 1940, persists today, but that scientific methods used at two of the three test sites corrected for it.''[49] The dress rehearsal data also displayed the persistence of the differential undercount.[50] In Sacramento, the estimated undercount rates that would have resulted without the use of Integrated Coverage Measurement were 4.7 percent for non-Hispanic Whites, compared to 8.7 percent for African Americans, 8.3 percent for Hispanics, and 6.0 percent for Asians. In Menominee County, Wisconsin, which is largely composed of the Menominee American Indian Reservation, the estimated undercount rate for non-Hispanic American Indians that would have resulted without the use of Integrated Coverage Measurement was 4.1 percent. In the South Carolina site, the estimated undercount rate for non-Hispanic Whites was 6.3 percent and 13.2 percent for all others (Hispanic, Black, American Indian, Hawaiian, and Asian).

It is clear from these results that, based on traditional census-taking methods alone, there was a substantial net undercount in all three sites, as well as a differential undercount of racial and ethnic minorities in those jurisdictions. The dress rehearsal demonstrated the operational feasibility of the A.C.E. and enhanced the Census Bureau's knowledge of the properties of statistical correction.

External Review

The Census Bureau's confidence that the application of the DSE methodology will result in a more accurate census is shared by many other entities that have critically examined this issue. Four different NAS panels over the decade have clearly endorsed the concept that a properly designed and executed coverage measurement survey has the potential to produce a more accurate census.[51] In 1999, the NAS Panel to Evaluate Alternative Census Methodologies concluded that:

> The only cost-effective methodology available for measuring the degree of differential undercoverage for subnational areas is a large-scale post-enumeration survey coupled with dual-system estimation * * *. If the Supreme Court prohibits use of integrated coverage measurement for apportionment, the panel still strongly supports a post-enumeration survey * * * for purposes other than apportionment.[52]

This recent conclusion is in line with those of the other three NAS panels. For example, in 1995, the Panel on Census Requirements in the Year 2000 and Beyond concluded that use of a high-quality survey in conjunction with the 2000 census will result in ''* * * improved accuracy with respect to the count and differential undercount for the nation as a whole as well as large areas and groups.''[53]

Numerous other organizations agree that the use of a properly conducted scientific survey in conjunction with the enumeration has the potential to produce a more accurate census in 2000.[54] These include, among others, the American Statistical Association, the American Sociological Association, the General Accounting Office, the Inspector General of the Department of Commerce, the Secretary of Commerce's Census 2000 Advisory Committee, the Census Bureau's Advisory Committee of Professional Associations, and the Census Bureau's Race and Ethnic Advisory Committees.[55]

A.C.E. Implementation Issues

The 1990 census coverage measurement survey was one of the most thoroughly evaluated programs conducted by the Census Bureau. The Census Bureau and other interested parties have analyzed volumes of data on the survey's effects on accuracy and how its results compared to the 1990 unadjusted census. Some of this analysis was performed in conjunction with Secretary Mosbacher's 1991 decision and the 1992 Committee on

---

[45] Bureau of the Census, ''Reinventing the Census,'' Global Report of the Task Force for Planning the Year 2000 Census, June 1995.

[46] For a description of these methodologies and the differences between them, see White and Rust, *Preparing for the 2000 Census: Interim Report II,* 48–51.

[47] A third National Academy of Sciences panel, the Panel to Evaluate Alternative Census Methodologies, was convened to study ways to improve the census for 2000. In its earlier report (Andrew A. White and Keith F. Rust, eds., *Sampling in the 2000 Census: Interim Report I* (Washington, D.C.: National Academy Press, 1996) following the 1995 Census Test, but before all the analyses from that test had been completed, the Panel concluded that ''* * * nothing in the [1995] census test, nor any other development, suggests that a decennial census that * * * reduces differential undercoverage can be conducted without the use of some form . . . of sampling for integrated coverage measurement'' (pp. 2–3). Based on the performance of DSE versus CensusPlus in the 1995 Census Test, the Census Bureau selected the former methodology for Census 2000, and the Panel supported that decision (White and Rust, *Preparing for the 2000 Census: Interim Report II,* 51–59).

[48] Edmonston and Schultze, *Modernizing the U.S. Census,* 3.

[49] Department of Commerce, ''Census 2000 Dress Rehearsal Shows Undercount Persists; Scientific Methods Correct Race and Ethnic Differential,'' *Commerce News,* 20 April 1999, CB99–CN.16 (revised).

[50] Bureau of the Census, ''Some Results from the Census 2000 Dress Rehearsal,'' by Rajendra Singh, DSSD Census 2000 Dress Rehearsal Memorandum Series A–76, 26 February 1999, 6.

[51] A fourth NAS panel was convened in June 1998 to review the Census Bureau's plans, procedures, and operations in connection with the Dress Rehearal and Census 2000. Experts from this panel are examining, among other things, the statistical methodology and procedures for the A.C.E.

[52] Cohen, White, and Rust, *Measuring a Changing Nation,* 4.

[53] Edmonston and Schultze, *Modernizing the U.S. Census,* 100.

[54] While support is widespread, the Census Bureau does not mean to imply that there is unanimous support on the issue. See, for example, Lawrence D. Brown and others, ''Statistical Controversies in Census 2000,'' *Jurimetrics* 39 (Summer 1999).

[55] Bureau of the Census, ''Report to Congress—The Plan for Census 2000,'' 24–25; Joint Census Advisory Committees on the Racial and Ethnic Populations, ''Recommendations Agreed Upon by the Four Census Advisory Committees on the African American, American Indian and Alaska Native, Asian and Pacific Islander and Hispanic Populations Made at the Meeting Held on May 22–23, 1997,'' Recommendation 3; the Secretary of Commerce's 2000 Census Advisory Committee, ''Final Report, Recommendation 3B, Post Enumeration Survey with a Traditional Census,'' 22 January 1999; and Census Advisory Committee of Professional Associations, ''Recommendations Made as a Result of the Meeting on April 22–23, 1999,'' Recommendation 1.

Adjustment of Postcensal Estimates (CAPE) report, but the Census Bureau has continued to examine the adjusted and unadjusted census data from 1990. These analyses have further clarified the relationship between the adjusted and the unadjusted 1990 census counts.

The extensive study of the 1990 coverage measurement survey identified a number of issues. The Census Bureau has considered these and other issues in assessing the feasibility of statistically correcting the Census 2000 counts. The following discussion presents many of these issues and addresses why the Census Bureau expects the A.C.E. to improve the overall accuracy of the census. In addition, changes in the A.C.E. design and their impact on accuracy are discussed.

Measuring Accuracy

Measuring accuracy in both the enumeration and the coverage measurement survey involves examining two types of error. One type, sampling error or variance, arises from the use of a sample to represent a population. Sampling error will occur only in the A.C.E. The other type, often termed nonsampling error, represents all other sources of error. Of particular concern in nonsampling errors are systematic errors or biases. Nonsampling errors will occur in both the initial census and the A.C.E. The most serious source of bias in the initial census is coverage error resulting from people missed or erroneous enumerations. The most notable example of bias in the enumeration is the historical phenomenon of the net undercount, including the differential undercount. Bias can also occur in the A.C.E., including errors due to false matches or nonmatches, inaccurately accounting for missing information, and other systematic collection or processing errors.

In designing coverage measurement surveys, the Census Bureau must strike a balance between sampling variance and bias. In comparing the accuracy of the 1990 coverage measurement survey to the accuracy of the unadjusted census, the Census Bureau concluded that the combined error in the coverage measurement survey was lower than the large bias in the census enumeration and therefore recommended adjustment. Secretary Mosbacher did not accept this recommendation and explained his reasons for not adjusting in his 1991 decision paper.

Assessment of Issues Emerging from 1990

The scrutiny and analysis of the 1990 census adjustment decision extended and sharpened discussions in the statistical community regarding the use of a coverage measurement survey to correct for census undercounts. Many of these issues were the subject of extensive discussion in Secretary Mosbacher's July 1991 decision document and in the 1992 CAPE report. Over the past decade, issues regarding the use of sampling to correct the census have been debated frequently in the technical literature.[56]

Some of these issues primarily address the basic principles and theories that must be considered in determining the proper application of a coverage measurement survey and DSE. For these issues reasoned judgment has to be invoked, and it is difficult to resolve these issues definitively by quantitative measurements. For example, what is the proper standard for deciding whether the coverage measurement survey should be used to correct the census? What priority should be given to numeric versus distributive accuracy? What are plausible assumptions about the distribution of individuals who are missed by both the initial census and the coverage measurement survey?

Other issues focus more on how well the Census Bureau can implement the coverage measurement survey, including the estimation processes. Is it operationally feasible to conduct the A.C.E. and produce the corrected results within the decennial time frame? Are the levels of sampling variance associated with the A.C.E. estimates reasonable? Can the levels of matching or other processing errors that occur in A.C.E. operations be kept to a minimum? These issues, while still subject to some degree of technical judgement, can often be evaluated by an examination of quantitative data.

As part of its comprehensive assessment of the A.C.E. design, senior Census Bureau officials requested a careful analysis of the technical issues identified in both the Mosbacher document and the CAPE report in order to ensure that cited concerns about accuracy had been adequately addressed. The Census Bureau's analysis of the Mosbacher document focused on the Secretary's guidelines and on supporting evidence for his decision.[57] The Census Bureau's analyses of the CAPE report focused on the accuracy of the unadjusted versus the adjusted census counts for different levels of geography and the status of the technical issues introduced.[58]

In addition to the discussion of technical issues, Secretary Mosbacher's analysis (and other reports critical of sampling) introduced a number of non-technical considerations. Secretary Mosbacher, for example, opined that ''adjustment would open the door to political tampering with the census in the future'' [59]—a theme frequently repeated in political, though not in scientific, discussions of sampling. No evidence has been presented that the Census Bureau has the competence to assess how its selection or implementation of census operations, including the many technical components of the A.C.E., might predetermine partisan outcomes. Furthermore, the highly pre-specified A.C.E. procedures make Census 2000 highly resistant to any form of manipulation. Although there are a number of agencies and groups—including the congressional committees charged with oversight of Census 2000, the General Accounting Office, the Census Monitoring Board, the Inspector General of the Department of Commerce, numerous advisory committees and other watchdog efforts—scrutinizing the planning and conduct of Census 2000, no evidence has been presented suggesting that the Census Bureau has any intention to affect political outcomes, or, if it did, that it has the technical ability to do so. The Census Bureau disputes any and all accusations that it would act out of political motives, and in this document restricts its discussion of concerns about the A.C.E. to those with technical and scientific content.

The Proper Standard To Use in Deciding Whether to Statistically Correct the Counts for Non-Apportionment Purposes

As was discussed earlier, Secretary Mosbacher's adjustment decision regarding the 1990 census was controlled by eight guidelines promulgated in connection with pending litigation. Secretary Mosbacher's decision not to adjust the 1990 census was based in large part on the standard articulated in the first guideline—that the unadjusted census would be ''* * * considered the most accurate count of the population of the United States, at the national, state, and

[56] Wright and Farmer, ''A Bibliography of Selected Statistical Methods Related to Census 2000.''

[57] Bureau of the Census, ''An Analysis of the Consistency of the 1990 Mosbacher Guidelines to U.S. Census Bureau Standards,'' by Sally M. Obenski and Robert E. Fay, 16 May 2000.

[58] Obenski and Fay, ''Analysis of CAPE Findings on PES Accuracy''; and Bureau of the Census, ''Analysis of CAPE Findings on 1990 PES Technical Issues,'' by Sally M. Obenski, 9 June 2000.

[59] Department of Commerce, ''Adjustment of the 1990 Census,'' 33583.

local level, unless an adjusted count is shown to be more accurate.''

Analysis and Response

This guideline assumed a priori that the unadjusted census counts were superior and required proof that the adjusted counts were better in terms of distributive accuracy at all three levels. This decision guideline required the adjusted counts to satisfy criteria that no other census operation could meet—in effect, the 1990 census coverage measurement survey was subjected to a higher standard than all other census operations.

If the Census Bureau had historically applied a similar presumption that a change to the census operation must demonstrate increased accuracy with convincing evidence for small levels of geography, it would not have made many important changes in census-taking methodology. For example, such a standard would not have permitted the Census Bureau to replace 100-percent in-home ''personal'' visits with mail questionnaires in the 1970 census. The Census Bureau did not know whether this fundamental change to the census operation would increase accuracy at all levels. Nor, in 2000, could the Census Bureau determine a priori that extensive promotion and paid advertising would increase accuracy at all levels, or for that matter, would be effective in all areas or for all demographic groups. If applied to all proposals to improve the initial census counts, this standard would effectively halt the Census Bureau's long tradition of scientific and technical innovation.

For Census 2000, the Census Bureau will make the determination on whether to use the A.C.E. to correct Census 2000 after evaluating (1) the conduct of key operations, (2) the consistency of the A.C.E. results with historical measures of undercount, and (3) measures of quality. As described previously, the Census Bureau's comprehensive ongoing analyses and experience with conducting coverage measurement surveys have led it to expect that the A.C.E. will improve overall numeric and distributive accuracy and that it will reduce the differential undercount. Therefore, statistical correction is appropriate absent strong evidence that it will degrade the overall quality of the final census data. However, the Census Bureau will conduct an objective review before making a final determination to release the statistically corrected data. The process that the Census Bureau will follow in making this determination is described in more detail at the end of this document. The Census Bureau will be documenting and discussing both this process and the criteria on which the determination will be made in a public setting in the fall of 2000.

Numeric v. Distributive Accuracy

The 1990 census adjustment decision (and the closely related decision on the adjustment of the postcensal estimates) was unequivocal in giving priority to distributive over numeric accuracy. Secretary Mosbacher interpreted the Constitutional and legal purposes of the census to require that:

> * * * accuracy should be defined predominately in terms of getting the proportional distribution of the population right among geographic and political units. This argues for putting aside the judgment of accuracy based on getting absolute numbers right (numeric accuracy) and instead focusing on the question of whether there is convincing evidence that the accuracy of population distribution in the adjusted numbers (distributive accuracy) is superior to the distributive accuracy of the actual enumeration.[60]

This injunction, when joined with the standard in the first guideline, requires not only that the adjusted counts be demonstrably more accurate at very low levels of geography but that they be more distributively accurate at those levels. This emphasis was reflected in many of the technical papers that have been written on the 1990 census. Comparatively less attention has been directed to the importance of numeric accuracy, despite the importance that the Census Bureau attaches to it. In fact, Secretary Mosbacher critiqued the Census Bureau for its interpretation ''of accuracy as concerned with getting the number of people closer to the truth rather than getting the allocation of the population for the purposes of political representation and funding closer to the truth.'' [61]

Analysis and Response

The Census Bureau believes that the adjustment decision in 1990 did not adequately consider the improvements to numeric accuracy that can result from statistical correction. Numeric and distributive accuracy are discussed more fully above. The issue here is the relative importance that should be assigned to numeric and distributive accuracy in assessing the results of the coverage measurement survey. Judgments can differ on this issue. It is the strong judgment of the Census Bureau that in deciding whether to use a coverage measurement survey to improve the census, both numeric and distributive accuracy should be taken into account.

The analysis and decision in 1990 focused almost exclusively on distributive accuracy. Although Secretary Mosbacher stated that the Census Bureau had provided substantial evidence (although ''not necessarily convincing'') that the adjusted counts were more numerically accurate, he based his conclusion not to adjust partially on the fact that improvements to distributive accuracy could not be demonstrated by convincing evidence at national, state, and local levels.[62] Given the decision criteria introduced by Secretary Mosbacher, the CAPE also focused on distributive accuracy.

The interaction between numeric and distributive accuracy is quite complicated, but must be considered in the analysis of the two types of accuracy. Clearly, there are situations where gains in numeric accuracy are expected without improvement in distributive accuracy. For areas or groups that have similar undercount rates, improvements to numeric accuracy are expected from the A.C.E. corrections. However, the distributive accuracy of these areas will be unchanged by the correction, because they will experience similar corrections. This outcome is expected, because gains in distributive accuracy are realized when areas corrected for significant undercounts are compared with areas that have little undercount. Because the A.C.E. is designed to improve the numeric accuracy of areas with significant undercounts, the Census Bureau expects that the A.C.E. will improve both numeric and distributive accuracy and thus result in a more accurate census overall.

Correlation Bias

Correlation bias is the result of either lack of independence between the initial census and the coverage measurement survey, or of variable inclusion probabilities within a post-stratum.[63] Frequently, the term is used to refer to error caused by individuals systematically missed in both the initial census and the coverage measurement survey. Important assumptions for DSE are that everyone in a given post-stratum has a similar inclusion probability and that the census and the coverage measurement survey are independent. Technically, these assumptions are referred to as homogeneity and causal independence, respectively. Correlation bias occurs when these assumptions are not fully satisfied. Although it is theoretically possible for correlation bias to result in

[60] Ibid., 33593.

[61] Ibid., 33592.

[62] Ibid., 33584.

[63] CAPE, 21–23.

either underestimation or overestimation by DSE, it is generally expected that correlation bias leads to underestimation. This will be the case, for example, when there are individuals who have little or no chance of being included in either the initial census or the coverage measurement survey. Some critics of the 1990 coverage measurement survey were concerned that correlation bias was so large as to preclude an improvement in distributive accuracy from adjustment.[64]

Analysis and Response

Correlation bias exists and will affect all dual system estimates. Post-stratification is used to minimize correlation bias. However, post-stratification is not a perfect solution, and it is reasonable to presume that some heterogeneity or causal dependence will persist, leading to some correlation bias. Comparisons with Demographic Analysis, though subject to limitations, have been used to obtain indications of possible correlation bias at the national level by age-sex-race groups. These comparisons in 1990 suggested correlation bias for adult Black males, and gave much less or no evidence of correlation bias for other groups. These analyses were restricted to the national level, and gave no indication of how any persons reflected in correlation bias may have been distributed geographically. In fact, there are no empirical data that can be used to definitely measure correlation bias below the national level. As a result, different hypotheses have been set forth regarding whether the A.C.E. will improve accuracy, particularly distributive accuracy. In the absence of quantitative data, the issues regarding the effects of correlation bias can only be resolved by a review of the assumptions underlying the various hypotheses, and by making judgments regarding which assumptions are more plausible.

The uncertainty about the geographic distribution of persons reflected in correlation bias relates to a concern of Secretary Mosbacher—the concern that because the distribution of those people missed by both the census and the coverage survey was not known, it could not be demonstrated that a statistical correction would improve distributive accuracy.[65] However, such a concern implicitly assumes that the distribution of correlation bias in dual system estimates differs from the distribution of undercount, as estimated in A.C.E. While recognizing the inherent limitations of its knowledge about the distribution of correlation bias, the Census Bureau believes it is more plausible to assume that correlation bias will tend to be distributed in a positive relation to the distribution of estimated undercount rates. A range of models reflecting plausible assumptions for the distribution of correlation bias have been analyzed.[66] This analysis of correlation bias, based on plausible assumptions, leads the Census Bureau to expect that improvements in distributive accuracy will be achieved by a properly designed and conducted coverage measurement survey.

Potential effects of correlation bias on numeric accuracy can also be addressed. Correlation bias, when present, is generally expected to lead to underestimation by dual system estimates. Therefore, when the DSE estimates an undercount in the initial census, by implication the initial census counts are even more severely undercounted. So the statistical corrections based on DSE are moving the census counts in the right direction, though not far enough. Thus, the statistical correction improves numeric accuracy when the groups subject to correlation bias are also undercounted by the census. In fact, the group identified by Demographic Analysis as probably subject to significant correlation bias in 1990 '' adult Black males '' also had a high estimated undercount rate from the 1990 PES.[67]

The Census Bureau expects that a properly designed and conducted coverage measurement survey should improve both numeric and distributive accuracy, even accepting that correlation bias cannot be eliminated. The Census Bureau will continue to use Demographic Analysis to assess the possibility of correlation bias at the national level.

Accuracy at Different Geographic Levels

When Secretary Mosbacher decided not to use the adjusted data in 1991, he indicated that the adjusted data could not be shown by convincing evidence to be more distributively accurate at the national, state, and local levels. The June 1991 Undercount Steering Committee report and later the August 1992 CAPE report concluded that adjustment, on average, improved distributive accuracy for states and areas with populations of more than 100,000.[68] The CAPE report, however, left the erroneous impression that the unadjusted census was more accurate at small geographic areas, generally, areas with a population of fewer than 100,000.

Analysis and Response

The CAPE report, issued on August 7, 1992, was followed by a November 25, 1992, Addendum. Because the CAPE work was conducted at the request of Secretary Mosbacher, the committee implicitly adopted the framework of the Mosbacher adjustment decision process in reaching its conclusions. That is, the adjusted census counts had to be shown to be more accurate at state and local levels in order to be adopted. The committee determined that it was unable to show that the adjusted census counts were more distributively accurate than the unadjusted counts for areas with fewer than 100,000 in population. Accordingly, the CAPE concluded that the unadjusted counts should be used in the postcensal estimates program. Unfortunately, the initial CAPE report could be interpreted as indicating that there was a problem with the accuracy of the adjusted census numbers for areas with a population of fewer than 100,000.

It is important to understand, however, that the Census Bureau did not stop its research into small area accuracy with the initial CAPE report. The initial CAPE analysis reported the Census Bureau's results from its first comparisons, comparisons of similar areas. For example, areas with populations of fewer than 25,000 were compared to each other, and major metropolitan areas were compared to each other. But the Census Bureau conducted additional research, comparing large cities and counties to each other, to the balance of the nation, and to the balance of their respective states. This additional research reported in the Addendum documented additional evidence of improvements in distributive accuracy at sub-state levels.[69]

The correct interpretation of the CAPE report and the Addendum is that the Census Bureau could distinguish no improvement in distributive sub-state accuracy if the corrected numbers had been used to produce estimates for areas

[64] Department of Commerce, ''Adjustment of the 1990 Census,'' 33591–92.

[65] Ibid.

[66] William R. Bell, ''Using Information from Demographic Analysis in Post-Enumeration Survey Estimation,'' *Journal of the American Statistical Association* 88 (September 1993): 1106–1118; and Bureau of the Census, ''Report of the Working Group on the Use of Demographic Analysis in Census 2000,'' by William R. Bell and others, 6 May 1996.

[67] CAPE, Table 2, Attachment 3A.

[68] Bureau of the Census, ''Technical Assessment of the Accuracy of Unadjusted Versus Adjusted 1990 Census Counts,'' Report of the Undercount Steering Committee, 21 June 1991, p. 2; and CAPE, 1.

[69] CAPE Addendum.

with populations of less than 100,000. It is incorrect to infer that the unadjusted census produced more distributively accurate sub-state data. That question was not tested in the CAPE research.

More recently, the Census Bureau has re-examined the CAPE data and determined that, based on available data, there is no basis for concluding that the unadjusted census was more distributively accurate than the adjusted counts for small areas.[70] That is, in general, no differences in the distributive accuracy of these two sets of counts have been demonstrated for geographic areas with less than 100,000 population.

Based on the CAPE and subsequent research and the expectation that the error structures of the initial census and A.C.E. operations for Census 2000 will be similar to 1990, the Census Bureau expects that the A.C.E. will, on average, increase distributive accuracy for areas with 100,000 or more residents. For areas with fewer than 100,000 people, the predicted effect of the A.C.E. on distributive accuracy is indeterminate—neither favoring the initial census nor the corrected counts.

With respect to numeric accuracy, as noted above, the Census Bureau expects that the A.C.E. will, on average, improve accuracy for geographic areas down to and including census tracts. Furthermore, the Census Bureau expects that improvement will be greatest for those areas that contain groups that have been historically undercounted.

Consistency with Demographic Analysis

The analysis of the 1990 coverage measurement survey included a comparison of the adjusted census with estimates based on Demographic Analysis (DA).[71] Discrepancies between the adjusted census and DA estimates led Secretary Mosbacher and others to question the accuracy of the 1990 adjusted census counts.

Analysis and Response

Demographic Analysis uses records and estimates of births, deaths, immigration, Medicare enrollments and estimates of emigration and undocumented immigration to estimate the national population, separately from the census. These demographic benchmarks are compared to the census counts, and the differences are used to create an estimate of the net census undercount. These estimates are produced for age groups (single years of age), sex, and broad race groups (Black, Non-Black). DA estimates can be used as independent benchmarks to validate the accuracy of coverage measurement survey estimates for corresponding demographic categories.

It is important to note that DA, like coverage measurement surveys, has an associated level of uncertainty. The Census Bureau developed quantitative measures of uncertainty for the 1990 DA estimates, but these measures are based in part on professional judgment about the range of error in each of the underlying demographic components.

How much uncertainty to assign to a DA estimate is therefore a matter of judgment. Different conclusions will be reached depending on basic assumptions about the accuracy of vital statistics and other records used in DA. In 1990, the Undercount Steering Committee concluded that the uncertainty in the DA estimates was of a magnitude that meant that many of the differences with the coverage measurement survey estimates resulted from random variation. However, Secretary Mosbacher reached another conclusion, citing several ''important and puzzling differences'' between the survey estimates and the DA estimates.[72] The Census Bureau, based on previous work in this area, concluded that some noted differences were expected, but these differences did not call into question the results of the coverage measurement survey.[73] Indeed, the difference between the DA and 1990 PES estimates for adult Black males was beyond the bounds of uncertainty, demonstrating the utility of Demographic Analysis for assessing correlation bias at the national level. Other differences fell within acceptable bounds of uncertainty associated with both sets of estimates. The Census Bureau considered all differences between the DA estimates and coverage measurement survey estimates in its determination that the coverage measurement survey did improve the accuracy of the census counts. For Census 2000, the Census Bureau will continue to compare both the uncorrected and corrected census counts with DA estimates.

Timing

In 1990, the adjusted data were not available for release until July 1991. This raises a concern about whether the Census Bureau can produce the statistically corrected data within the statutory deadline of April 1, 2001, for redistricting, without sacrificing the quality of the initial census or the A.C.E.

Analysis and Response

The timing and quality of the initial census and the A.C.E. are related. The Census Bureau has developed a schedule for the initial census and for the A.C.E. operations that allows adequate time to produce uncorrected data for apportionment and corrected data prior to the statutory deadline. Barring some major, unanticipated operational difficulty,[74] the Census Bureau expects to complete all data collection and processing functions for the initial census and the A.C.E. in time to deliver quality, statistically corrected redistricting numbers to the states prior to April 1, 2001.

Critical differences between the 1990 census and plans for Census 2000 should allow production of the corrected numbers within the required period. First, the 1990 plan was not premised on producing the adjusted numbers by the April 1 deadline. In fact, the 1990 litigation established a deadline of July 15, 1991, for delivery of the adjusted data.

Second, there are improvements to the census that will make the initial Census 2000 operations more timely. While these improvements are directed at allowing enumeration data collection to occur closer to Census Day and therefore to be more accurate, they will also allow for an earlier start for the A.C.E. With respect to the key issue of staffing nonresponse followup so as to finish on schedule, which is crucial to the progress of both the census and the A.C.E., the Census Bureau has developed strategies to avoid the recruitment and retention problems that extended the 1990 census NRFU operation. The Census Bureau has conducted extensive research on how to ensure the recruitment and retention of well-qualified temporary employees. These strategies, successfully employed during the Census 2000 dress rehearsal, included the targeting of wage rates to local areas and a technique called frontloading. Frontloading is directed at

[70] Obenski and Fay, ''Analysis of CAPE Findings on PES Accuracy.''

[71] Bell, ''Using Information from Demographic Analysis in Post-Enumeration Survey Estimation,'' 1106–1118; and Robinson and others, ''Estimates of Population Coverage in the 1990 United States Census,'' 1061–77.

[72] Department of Commerce, ''Adjustment of the 1990 Census,'' 33587.

[73] Bureau of the Census, ''Technical Assessment of the Accuracy of Unadjusted Versus Adjusted 1990 Census Counts,'' 4.

[74] Census Bureau Director Dr. Prewitt provided examples of such operational difficulties in his February 14, 2000, letter to Chairman Dan Miller of the House Subcommittee on the Census. These examples include: (1) Problems with the payroll system that prevent the Census Bureau from paying its employees on a timely basis; (2) widespread problems filling enumerator positions, despite the agency's extensive pool of qualified applicants; and (3) problems with the Census 2000 address file that prevent Census Bureau employees from being able to fulfill their responsibilities.

reducing the effects of early turnover of employees by hiring two employees for every position. As a result of these and other changes, nonresponse followup will take place in a shorter time period in Census 2000. This shortening of nonresponse followup is in accord with the observations of the Census Bureau and the General Accounting Office that NRFU results decrease in accuracy as the time from Census Day increases.[75] In addition, Census 2000 will not be repeating certain ineffective coverage improvement programs that delayed processing of the initial census in the 1990 coverage measurement survey.[76]

Third, several important changes will improve the timeliness of the A.C.E. operation. For example, the A.C.E. interviewers will have received more extensive training than in 1990. Additionally, the Census Bureau has developed a Computer Assisted Person Interviewing (CAPI) system for the A.C.E. that will allow enumerators to collect the data more quickly and accurately, and to transmit it electronically in a more expeditious manner by using laptop computers.

In the unlikely event of an unanticipated, major operational difficulty, the Census Bureau will not curtail important operations key to the quality of the entire census to stay on schedule. For example, the Census Bureau will not curtail nonresponse followup in difficult-to-enumerate neighborhoods to stay on the A.C.E. schedule. Likewise, the Census Bureau will not curtail the A.C.E. data collection activities. The Bureau is committed to achieving high quality in all census operations, and Census Bureau statisticians will be monitoring key A.C.E. performance information, such as response rates, for early warning about areas warranting corrective actions.

Level of Sampling Variance/Smoothing

The levels of sampling variance and bias in the 1990 coverage measurement survey were important topics in the adjustment debate. Sampling variance is discussed in this section; bias will be discussed in the following section.

Analysis and Response

One issue in 1990 was the use of a statistical technique called smoothing, a complex, model-based method designed to control sampling variance. The use of smoothing led to an extensive discussion regarding the robustness of the 1990 methodology. For Census 2000, the Census Bureau has developed the A.C.E. sample design so that smoothing will not be necessary. There were also concerns about the overall level of sampling variance in the 1990 coverage measurement survey.[77] In developing the A.C.E. design, the Census Bureau thoroughly examined 1990 variance issues and made important design decisions to reduce sampling variance levels. These include:

- The A.C.E. sample size is almost double that of 1990, increased from approximately 165,000 to 314,000 housing units. Because sampling variance is inversely proportional to sample size, this increase will reduce the level of sampling variance in 2000.
- The A.C.E. sample was designed to minimize the range in size of the sampling weights. Weights are assigned to categories of blocks (that is, small and large) that have different probabilities of being selected in the sample. When there is a wide range of weights, variance increases because blocks with large weights have a disproportionate effect on the variance of the estimates. The Census Bureau has designed its sampling procedures for Census 2000 specifically to limit how much these weights will vary. This design will result in reduced sampling variance.[78]

The Census Bureau has used the 1990 experience to develop an enhanced A.C.E. sampling design, and does not anticipate that variance-related issues will be a serious source of concern for the Census 2000 coverage measurement survey.

Level of Nonsampling Error/Bias

One concern was that the level of nonsampling error or bias in the 1990 coverage measurement survey was so large that statistical correction would not result in an improvement in distributive accuracy.[79] Critics of the A.C.E. have expressed similar concern about the anticipated level of bias in the Census 2000 DSE.[80]

Analysis and Response

The Census Bureau conducted extensive evaluations of nonsampling error in the 1990 coverage measurement survey.[81] These evaluations have given the Census Bureau a detailed understanding of nonsampling error. Based on this extensive work, the Census Bureau has concluded that the levels of nonsampling error in the 1990 PES did not prevent the statistical correction based on the coverage measurement survey from improving the accuracy of the census counts.[82] The A.C.E. design includes enhancements to the 1990 coverage measurement survey that will even further control nonsampling error.

It is important to note that some amount of bias in both the initial census and the A.C.E. is inevitable. However, the Census Bureau's analysis of bias, grounded in sound statistical principles, leads to the expectation that the improvements described in the following sections will control the levels of nonsampling error in the A.C.E. so that a statistical correction based on the A.C.E. will improve the uncorrected counts.

Enhancements to the Matching Process

Matching refers to the determination of whether an individual enumerated in a coverage measurement survey is the same person as an individual enumerated in the initial census operation. Because errors in matching can significantly affect undercount estimates, highly accurate matching is an important component of the A.C.E. methodology. Although neither Secretary Mosbacher nor CAPE identified matching error as a significant problem with the 1990 coverage measurement survey, the Census Bureau has made significant improvements to the matching process in the 2000 A.C.E. design, and matching error is expected to be even lower in Census 2000 than in 1990:

- A fully automated system supports computer and clerical matching, an advance over 1990 procedures that required handling and control of paper documents. This improvement provides for a number of built-in edits and quality checks to control matching error.

[75] Bureau of the Census, ''Characteristics of Census Errors,'' by Deborah Griffin and Christopher Moriarity, 1990 Decennial Census Preliminary Research and Evaluation Memorandum No. 179, 15 September 1992; and General Accounting Office, *Decennial Census—1990 Results, 47.*

[76] These coverage improvement programs are discussed briefly in Cohen, White, and Rust, *Measuring a Changing Nation, 32–33.*

[77] Bureau of the Census, ''Adjustment of the 1990 Census,'' passim.

[78] The NAS agrees that design changes in the A.C.E. will reduce the variance in block sampling weights, a ''key improvement in comparison to the 1990 design,'' May 3, 1999, letter from Janet L. Norwood, Chair, NAS Panel to Review the 2000 Census, to Kenneth Prewitt, Director, U.S. Bureau of the Census.

[79] Bureau of the Census, ''Adjustment of the 1990 Census,'' and Leo Breiman, ''The 1991 Census Adjustment: Undercount or Bad Data?'' *Statistical Science* 9, no. 4 (1994): 458–537.

[80] David A. Freedman and Kenneth Wachter, University of California, letter to Rep. Miller, Chairman, House Subcommittee on the Census, 17 May 2000; and Brown and others, ''Statistical Controversies in Census 2000.''

[81] Bureau of the Census, ''Technical Assessment of the Accuracy of 1990 Census Counts,''4; and Mary H. Mulry and Bruce D. Spencer, ''Accuracy of the 1990 Census and Undercount Adjustments,'' *Journal of the American Statistical Association* 88 (September 1993): 1080.

[82] Bureau of the Census, ''Technical Assessment of the Accuracy of 1990 Census Counts,'' 1.

The automated matching system is the culmination of Census Bureau analyses and refinements over the last 20 years and will make searching and matching easier and more reliable.

• The matching processes have been centralized in one site, rather than decentralized as in 1990, allowing for more effective control—a well-trained staff will perform all matching at a single location.

• As discussed below, the change in the treatment of people who have moved since Census Day will simplify matching for these movers. Unlike in 1990, it will only be necessary to match people who resided in the sample blocks on April 1.

Enhancements to Computer Processing

After the initial release of the adjusted numbers in July 1991, the Census Bureau discovered a computer processing error that resulted in a 0.4 percent decrease in the estimated undercount for the 1990 census. The CAPE report reduced the Census Bureau's official undercount estimate from 2.1 percent to 1.6 percent, with 0.4 percentage points attributable to the computer processing error, and 0.1 percent attributable to additional processing corrections. Concerns have been raised relating to the Census Bureau's late discovery of the computer processing error. These concerns have been cited as evidence that the complexity of the computer operations associated with incorporating the results of a coverage measurement survey—like the A.C.E.—in the census counts makes the final numbers vulnerable to significant processing errors.[83]

The Census Bureau has adopted a number of methods to improve the quality of the A.C.E. software to guard against a similar error in Census 2000:

• To ensure reliability, the Census Bureau has included software validation and verification strategies, such as independent software development of key computer programs (double programming).

• To reduce ambiguity and increase communication, the Census Bureau has enforced standardized nomenclature and adopted an improved documentation approach for technical issues.

• The Census Bureau has developed a Sample Design Control System. This system provided the necessary data to control, monitor, and validate the different phases of sampling. It also ensured that the software used to select the A.C.E. sample functioned correctly.

• The software programs supporting the A.C.E. estimation process will be further validated by an Integrated Review System. This system will provide data on all phases of the estimation process that will allow timely validation that the software is performing as specified.

These and other initiatives should result in a controlled, robust, and reliable A.C.E. computer processing environment. Therefore, the Census Bureau expects the processing for the Census 2000 A.C.E. to be not only more streamlined but also more reliable than it was for the 1990 PES.

Enhancements to Minimize Missing Data

Missing data cases involve the following situations where complete information cannot be obtained: missing characteristic data (race, age, or other characteristic information), complete non-interviews, or cases with insufficient information to determine an individual's enumeration or match status. In 1980, missing data in the coverage measurement survey was a serious problem and factored into senior statisticians' conclusion that the estimates were not sufficiently reliable to use for statistical adjustment of the census counts. The Census Bureau took steps to minimize missing data in the 1990 coverage measurement survey, and missing data in 1990 did not significantly affect the accuracy of the estimates.[84] Nonetheless, concerns remain regarding the potential for high levels of missing data in the A.C.E.

Building on its experience from the 1990 census, the Census Bureau has designed its field operations to minimize missing data. After the initial A.C.E. interview attempt, the Census Bureau will allow up to two additional weeks for attempts to revisit any nonresponding households. This two-week period of intense followup of nonresponding households will be conducted by the Census Bureau's best and most experienced available A.C.E. interviewers.[85] Finally, Census Bureau staff will be monitoring missing data rates closely throughout the conduct of the A.C.E.

The Census Bureau has developed additional extensive procedures to deal with missing data. One method the Census Bureau uses to handle missing data in both the initial census and the A.C.E. is imputation. Imputation is an established statistical methodology that completes missing respondent information by incorporating information provided by other similar respondents. The imputation process for Census 2000 draws on lessons learned in the 1990 census. Additionally, the imputation process for Census 2000 has been simplified, which should result in the production of more easily validated data.

While missing data were not a significant issue for the 1990 census,[86] some concerns have been expressed regarding the accuracy and robustness of the Census Bureau's imputation model for the 1990 coverage measure survey.[87] However, Census Bureau statisticians and others have conducted multiple evaluations using different methodologies to independently validate the imputation model used in the 1990 census.[88] These evaluations and the improvements to missing data procedures discussed earlier lead the Census Bureau to expect that missing data will not be a substantial problem in the A.C.E.

Homogeneity and the Synthetic Assumption

Generally speaking, homogeneity refers to the principle that individuals grouped in a post-stratum have similar probabilities of being included in the census, that is, similar coverage probabilities. If homogeneity holds, conclusions can be drawn from a sample about population groups or geographic areas and the initial enumeration for these population groups or areas can be corrected with a coverage measurement survey. The synthetic assumption states that the people in a particular post-stratum are

[83] Congress, House, Committee on Government Reform and Oversight, *Prepared Testimony of K.W. Wachter and D.A. Freeman Before the House Committee on Government Reform and Oversight*, 104th Cong., 1st sess., 29 February 1996.

[84] Bureau of the Census, ''Non-Response in Census Coverage Measurement Surveys and Its Impact—An Historical Review,'' by Ruth Ann Killion, DSSD Briefs, Information, and Topics Memorandum Series No. 44C, 17 September 1998, pp. 6–7.

[85] This intensive field operation designed to minimize missing data is described in Childers and Fenstermaker, ''Accuracy and Coverage Evaluation: Overview of Design,'' 7.

[86] Secretary Mosbacher stated that levels of missing data were sufficiently low so that variation in the Census Bureau's missing data models made no difference in the outcome of the survey, and he concurred with the Undercount Steering Committee's judgment that the outcome was robust (Bureau of the Census, ''Adjustment of the 1990 Census,'' 33600). The CAPE, accordingly, did not examine this issue.

[87] Brown and others, ''Statistical Controversies in Census 2000.''

[88] T.R. Belin and others, ''Hierarchical Logistic-Regression Models for Imputation of Unresolved Enumeration Status in Undercount Estimation,'' *Journal of the American Statistical Association* 88 (September 1993): 1149–66; and Bureau of the Census, ''Documentation of Handling Unresolved Enumeration Status in 1990 Census/Post-Enumeration Survey,'' by Greg Diffendal and Tom Belin, STSD Decennial Census Memorandum Series V–98, 15 January 1991.

relatively homogeneous and will generally share the same coverage factor. There are concerns, however, that a lack of homogeneity could lead to inaccuracies being introduced into the data for areas or population groups within the post-stratum.[89]

Analysis and Response

At issue is not whether there is perfect homogeneity; at issue is whether heterogeneity is too great to prevent an improvement from using the A.C.E. While the degree to which the homogeneity assumption holds is a continuing issue, the Census Bureau has made design improvements to the A.C.E. to control heterogeneity and believes that heterogeneity will not preclude the production of useful small area data in Census 2000.

The statistical correction that results from the A.C.E. is carried down to census blocks by applying the coverage correction factors within each A.C.E. post-stratum. The goal in constructing post-strata is to form groupings of the population that capture differences in the probabilities of being included in the census and the A.C.E.[90] In effect, the inclusion probabilities are more similar for individuals within the same post-stratum than for individuals in different post-strata. The coverage correction factors are calculated for each post-stratum, based on a representative sample of the post-stratum, and thus reflect the net coverage of all people within the post-stratum. This is the underlying basis for applying this factor to the data records within the corresponding post-stratum to produce statistically corrected block totals which serve as the basis for Census 2000 tabulations.

The more homogeneity within a post-strata and the more differences among post-strata, the greater the improvement from statistical correction. In designing post-strata, it is not necessary for each individual to have the same probability of inclusion. Since no two individuals are perfectly alike with respect to their chances of being included in either the initial census or the A.C.E., the goal for defining post-strata is to form groupings of the population with similar inclusion probabilities. That is, the goal is to form post-strata that differentiate between groups of the population with respect to inclusion probabilities, and with respect to net coverage in the initial census. Some have suggested that an improvement will result from applying Demographic Analysis-based corrections within national post-strata consistent with DA.[91] However, the Census Bureau expects to achieve greater improvements by having defined post-strata that take advantage of more local data.[92]

The accuracy of the estimates that result from the application of the coverage correction factors depends on the degree to which the net coverage for areas or groups within a post-stratum is similar to the coverage correction factor that was developed for that post-stratum. The coverage correction factor is measured for the post-stratum based on a representative sample, and thus represents the net coverage for the post-stratum. Clearly, within the post-stratum, some degree of variation is expected from the measured coverage correction factor, and this variation will most likely be relatively greater for small areas. Thus, it is inevitable that the A.C.E. will result in the population in some blocks being overestimated and the population in other blocks being underestimated. The A.C.E. statistical correction was never intended nor expected to produce unqualified improvement in the smallest geographic areas, like blocks. That the A.C.E. does not produce improvement for every single block, however, is no reason to forego the benefits that will flow from the use of corrected census population counts at geographic levels of significance to data users. The Census Bureau expects that the A.C.E. estimates will produce better data for aggregations—such as states, congressional districts, counties, and cities—that are the basic areas for which census data are used.

The Census 2000 A.C.E. incorporates improvements from the design used for the 1990 coverage measurement survey that are expected to improve the homogeneity within post-strata for Census 2000.[93] The Census Bureau analyzed heterogeneity as part of the 1990 CAPE process, and has continued research for the A.C.E. post-strata.[94] Building on the lessons learned from 1990, the Census Bureau has developed enhanced post-strata for Census 2000. For example, the A.C.E. post-strata definitions include mail return rate and type of enumeration variables.[95]

Some have cited the CAPE report as evidence that the Census Bureau had serious concerns about heterogeneity. A reading of the entire CAPE report, including the more technical Addendum, puts these concerns in proper perspective. That is, the full analysis of the CAPE report (including the Addendum) supports the expectation of the Census Bureau that the use of the A.C.E. results will lead to improvements in the accuracy of the Census 2000 data.

Additional Design Changes From 1990

In addition to the specific improvements discussed previously, the Census Bureau has implemented other changes to the 1990 coverage measurement survey design. These changes, which will improve operational efficiency, include the use of the telephone in the A.C.E., and changes in the treatment of movers and the search area for matching. The Census Bureau will also collect data on race and ethnicity differently in Census 2000. The Census Bureau continues to consider and examine issues relating to these changes.

Use of the Telephone in A.C.E. Interviewing

To gain efficiencies in the interviewing phase of the A.C.E., enumerators will conduct telephone interviews using CAPI laptop computers for households that have returned their census questionnaires by mail. By design, this interview will take place before or concurrent with the initial census nonresponse followup. The interviews will be conducted from the homes of the A.C.E. enumerators and will be conducted only for households that mail back a questionnaire that includes a telephone number. Furthermore, the households must be in areas where there is negligible risk of mail delivery problems—generally, single family housing units or large multi-unit structures in areas with city-style mail delivery.

The Census Bureau implemented this process to enhance the efficiency and

[89] K. Wachter and D. Freedman, ''Local Heterogeneity and Census Adjustment for the Intercensal Base,'' Technical Report No. 381 (Berkeley, CA: University of California, Department of Statistics, 1993); and Brown and others, ''Statistical Controversies in Census 2000,'' 13.

[90] This section presents a general discussion of the basis for synthetic or indirect estimation. There are more complex, but less stringent, requirements involving the relationship between census omissions and erroneous enumerations as discussed in Howard Hogan's paper, ''Accuracy and Coverage Evaluation: Theory and Application.''

[91] David Freedman and Kenneth Wachter, ''Planning for the Census in the Year 2000,'' *Evaluation Review*, 20, (August 1996): 355–77.

[92] Bureau of the Census, ''Accuracy and Coverage Evaluation Survey: Final Post-stratification Plan for Dual System Estimation,'' by Richard Griffin and Dawn Haines, DSSD Census 2000 Procedures and Operations Memorandum Series Chapter Q–24, 19 April 2000.

[93] *Ibid.*

[94] Robert E. Fay and John Thompson, ''The 1990 Post Enumeration Survey: Statistical Lessons in Hindsight,'' in *Proceedings of the 1993 Annual Research Conference, 21–24 March 1993.*

[95] Griffin and Haines, ''Accuracy and Coverage Evaluation Survey: Final Post-stratification Plan for Dual System Estimation.''

quality of the A.C.E. interview. Shortening the elapsed time from Census Day to the A.C.E. enumeration should improve data quality. Also, starting early in an environment that is more easily controlled should allow the A.C.E. enumerators to gain valuable experience in conducting interviews and in operating their laptop computers. The Census Bureau designed this process in a fashion that should maintain the independence between the A.C.E. and the other Census 2000 operations.

New Treatment to Account for Movers

The Census Bureau has changed its treatment of individuals whose residence changes after Census Day. In the 1990 coverage measurement survey, movers were sampled where they lived at the time of the PES interview. The Census Bureau then searched the census records at the movers' April 1 usual residence to determine if they had been correctly enumerated in the census.[96]

In the modified procedure employed by the A.C.E., the Census Bureau will combine information on movers from two sources to produce an estimate of movers who are missed in Census 2000. First, an estimate of the total number of movers will be calculated based on people who moved into the A.C.E. sample blocks between April 1 and the time of the A.C.E. interview. Second, the rate at which movers match to Census 2000 will be based on reconstructing the Census Day residents of the A.C.E. sample housing units and matching these residents to the initial census records. Reconstructing the Census Day residents will be based on proxy interviews with the new residents or neighbors. These two estimates will be combined to form an estimate of the movers who are missed in Census 2000. These results are then used in the Dual System Estimation. The Census Bureau tested the modified procedure in the dress rehearsal and has judged this procedure to be the best blend of operational feasibility and accuracy.[97]

Search Area for Matching

The Census Bureau's search operation in the 1990 coverage measurement survey used an extended search area in blocks adjacent to the sample blocks.[98] The extended search area included one ring of adjacent blocks, or two rings of adjacent blocks in most rural areas. A person located in either the sample or an adjacent block was labeled a correct enumeration or match. Defining the search area in this fashion provided significant gains in reducing sampling variance. For Census 2000, the A.C.E. search area has been designed to achieve the gains in controlling sampling variance, while providing operational efficiencies.

The Census 2000 search operation uses a sampling procedure that selects A.C.E. block clusters for an extended search. All block clusters are selected where there is evidence that an extended search will provide substantial information needed for the A.C.E. matching. Additionally, a random subsample of all other clusters is selected for the extended search.[99] This decision was based on an analysis of the results of the 1990 census coverage measurement survey matching that indicated that this strategy would provide virtually the same gains in sampling variance reduction as compared to the 1990 results.[100]

Reporting More Than One Race

In accordance with direction from the Office of Management and Budget,[101] Census 2000 will for the first time allow individuals to report more than one racial category. This guidance from the OMB necessitates that the A.C.E. post-strata be defined taking into account people that report more than one race. The Census Bureau, therefore, has defined and documented the A.C.E. post-strata to include individuals that report more than one race.[102] The Census Bureau will conduct a study of the effects of multiple race reporting after completion of the census.

Making the Final Decision

The Census Bureau expects that the A.C.E., if properly conducted, will make the census more accurate by improving coverage and reducing differential undercounts. The Census Bureau will not, however, release corrected redistricting data until it has brought its technical judgment to bear in assessing the available data to verify that its expectations have been met. The Census Bureau will consider operational data to validate the successful conduct of the A.C.E., assess whether the A.C.E. measurements of undercount are consistent with historical patterns of undercount and independent Demographic Analysis benchmarks, and review measures of quality.

In preparing for this determination, the criteria and the process that will be followed for the assessment of the A.C.E. results will be shared and discussed with outside statistical experts and other interested parties in the fall of 2000. This plan is consistent with the principle of pre-specification adopted by the Census Bureau for the Census 2000 A.C.E. and with its open and transparent planning and decision processes. The extent of pre-specification already publicly provided is very extensive.

It should be noted that all major census operations are vulnerable to unanticipated difficulties. Such difficulties could affect production of the apportionment counts. If, for example, a major natural disaster were to occur in a region of the country during census nonresponse followup, and this operation were seriously disrupted, the Census Bureau might conclude that the apportionment count so misrepresented the ''true'' state-by-state population distribution that it should not be used until corrective action was taken, possibly delaying delivery of the apportionment counts past January 1, 2001. Unanticipated difficulties could also affect the A.C.E. The Census Bureau would respond to any major unanticipated operational difficulty by taking steps to conduct and complete (or repeat, as necessary) all planned operations necessary to ensure that an accurate A.C.E. had taken place before releasing the statistically corrected data. If the Census Bureau determines that incorporating the results of the survey would not improve the accuracy of the initial census counts, then the uncorrected data would be denominated as the P.L. 94–171 file.

Secretary Mosbacher's 1991 decision document raised the specter of ''political tampering'' in any use of statistically corrected census data. To avoid even the appearance of political manipulation, the Census Bureau has proposed a process for verifying the agency's expectations regarding the improvements in accuracy from the A.C.E. Under that proposal, a committee

[96] Hogan, ''Accuracy and Coverage Evaluation Theory and Application,'' 22.

[97] Bureau of the Census, ''Minutes of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy (ESCAP),'' 5 January 2000.

[98] Hogan, ''1990 Post-Enumeration Survey,'' 1054.

[99] Childers and Fenstermaker, ''Accuracy and Coverage Evaluation: Overview,'' 8–9; and Bureau of the Census, ''Accuracy and Coverage Evaluation: The Design Document,'' by Danny R. Childers, DSSD Census 2000 Procedures and Operations Memorandum Series Chapter S–DT–01.

[100] Bureau of the Census, ''Accuracy and Coverage Evaluation Survey: Targeted Extended Search Plans,'' by Alfredo Navarro, DSSD Census 2000 Procedures and Operations Memorandum Series #Q–18, 12 January 2000.

[101] President, Executive Office, Office of Management and Budget, ''Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity,'' Federal Register, (30 October 1997), vol. 62, no. 210, pp. 58782–90.

[102] The Census Bureau's plan for including individuals in the A.C.E. post-strata who report more than one race is described in Griffin and Haines, ''Accuracy and Coverage Evaluation Survey: Final Post-stratification Plan for Dual System Estimation,'' 1–2.

of senior Census Bureau officials responsible for resolving policy and technical issues regarding the A.C.E. and assessing the technical effectiveness of its operations would make a recommendation to the Census Bureau Director regarding the use of the statistically corrected census data. The Director would make a determination regarding the use of the statistically corrected data, taking into consideration the recommendation of the committee. This committee, known as the Executive Steering Committee for A.C.E. Policy (ESCAP), was formed in late 1999 and normally meets every two weeks to discuss technical and policy issues associated with the A.C.E. and to advise the Director on these issues. The ESCAP is chaired by the Associate Director for Decennial Census and includes the following other senior career staff: Deputy Director; Principal Associate Director and Chief Financial Officer; Principal Associate Director for Programs; Associate Director for Methodology and Standards; Associate Director for Demographic Programs; Assistant Director for Decennial Census; Chief, Decennial Statistical Studies Division; Chief, Planning, Research and Evaluation Division; Chief, Population Division; Chief, Decennial Management Division; and Senior Mathematical Statistician. The committee will document its discussions and decisions and will make this documentation available along with its recommendation to the Director.

Following the release of census data, the Census Bureau will continue its research and evaluation, budget permitting. The census is an ongoing process, and the Census Bureau implements refinements to the data over a 10-year period. These ongoing efforts are consistent with good science and are fundamental to the Census Bureau's work. The fact that further research will provide more information about the success of census operations, including the production of the apportionment counts and the A.C.E., does not alter the requirement to release the statistically corrected block-level numbers by the April 1, 2001, statutory deadline, if these data meet the Census Bureau's expectations with regard to improvements in accuracy. Evaluations of many Census 2000 operations and results, including the A.C.E., will continue after the release of the data; and program evaluation results will be available for planning the 2010 census and informing the scientific and public discourse over the intervening years.

Conclusions

The Census Bureau's mission is to produce the most accurate data possible, taking into account the intended uses of the data. The extensive body of research that the Census Bureau has conducted on census undercount, including the 1990 census evaluations, has conclusively demonstrated that traditional census methodologies will not effectively reduce the differential undercount. The Census Bureau has concluded that based on current state-of-the-art science, the best method or procedure that has the potential to reduce the differential undercount and thereby increase accuracy is the application of scientific sampling to improve traditional census methods. This view is widespread, though not unanimous, in the professional statistical community.

At the present time, the Census Bureau has also concluded that it is operationally feasible to complete the A.C.E. and produce statistically corrected census data prior to April 1, 2001, and expects that the corrected data will be the most accurate data available. The Census Bureau's final decision on what data to release as the most accurate data will not be made, however, until the Census Bureau has had an opportunity to review the conduct of the census and the A.C.E.

Bibliography

Bell, William R. ''Using Information from Demographic Analysis in Post-Enumeration Survey Estimation.'' *Journal of the American Statistical Association* 88 (September 1993): 1106–18.

Belin, T.R., G. J. Diffendal, S. Mack, D.B. Rubin, J.L. Schafer, and A.M. Zaslavsky. ''Hierachical Logistic-Regression Models for Imputation of Unresolved Enumeration Status in Undercount Estimation.'' *Journal of the American Statistical Association* 88 (September 1993): 1149–66.

Breiman, Leo. ''The 1991 Census Adjustment: Undercount or Bad Data?'' *Statistical Science* 9, no. 4 (1994): 458–537.

Brown, Lawrence D., Morris L. Eaton, David A. Freedman, Stephen P. Klein, Richard A. Olshen, Kenneth W.Wachter, Martin T.Wells, and Donald Ylvisaker. ''Statistical Controversies in Census 2000.'' *Jurimetrics* 39 (Summer 1999): 347.

Census Advisory Committee of Professional Associations. ''Recommendations Made as a Result of the Meeting on April 22–23, 1999, Recommendation 1.''

Childers, Dan, Gregg Diffendal, Howard Hogan, Nathaniel Schenker, and Kirk Wolter. ''The Technical Feasibility of Correcting the 1990 Census.'' In *Proceedings of the Social Statistics Section of the American Statistical Association Held in San Francisco, California, 17–20 August 1987,* by the American Statistical Association, 1987.

Childers, Dan and Howard Hogan. ''The 1988 Post Enumeration Survey Methods and Preliminary Results.'' In *Proceedings of the Survey Research Methods Section of the American Statistical Association, New Orleans, Louisiana, 22–25 August 1987,* by the American Statistical Association, 1988.

City of New York v. *U.S. Dept. of Commerce.* 822 F. Supp. 906, 928 (E.D.N.Y. 1993).

Cohen, Michael L., Andrew A. White, and Keith F. Rust, eds. *Measuring a Changing Nation—Modern Methods for the 2000 Census.* Washington, D.C.: National Academy Press, 1999.

Commerce Secretary's 2000 Census Advisory Committee. ''Final Report, Recommendation 3B, Post Enumeration Survey with a Traditional Census,'' 22 January 1999.

Department of Commerce v. *House of Representatives,* 119 S. Ct. 765 (1999).

Edmonston, Barry, and Charles Schultze, eds. *Modernizing the U.S. Census* Washington, D.C.: National Academy Press, 1995.

Fay, Robert E., Jeffrey S. Passel, J. Gregory Robinson, and Charles D. Cowan. *The Coverage of Population in the 1980 Census.* Washington, D.C.: Government Printing Office, 1988.

Fay, Robert E., and John Thompson. ''The 1990 Post Enumeration Survey: Statistical Lessons, in Hindsight.'' In *Proceedings of the 1993 Annual Research Conference, 21–24 March 1993,* by the Research Conference, 1993.

Freedman, David A., and Kenneth W. Wachter, University of California. Letter to Rep. Dan Miller, Chairman, House Subcommittee on the Census, U.S. House of Representatives. 17 May 2000.

____. ''Planning for the Census in the Year 2000.'' *Evaluation Review* 20 (August 1996): 355–77.

Hogan, Howard. ''The 1990 Post-Enumeration Survey: Operations and Results.'' *Journal of the American Statistical Association* 88 (September 1993): 1047–67.

____. ''Accuracy and Coverage Evaluation: Theory and Application.'' Dual System Estimation Workshop of the National Academy of Sciences Panel to Review the 2000 Census. 2 February 2000.

Joint Census Advisory Committees on Race and Ethnic Populations. ''Recommendations Agreed Upon by the Four Census Advisory Committees on the African American, American Indian and Alaska Native, Asian and Pacific Islander and Hispanic Populations Made at the Meeting Held on May 22–23 1997, Recommendation 3.''

Mulry, Mary, and Bruce D. Spencer. ''Accuracy of the 1990 Census and Undercount Adjustments.'' *Journal of the American Statistical Association* 88 (September 1993): 1080–91.

Norwood, Janet L., Chair, National Academy of Sciences Panel to Review the 2000 Census. Letter to Kenneth Prewitt, U.S. Department of Commerce, Bureau of the Census, 3 May 1999.

Robinson, J.G., B. Ahmed, P. Das Gupta, and K. A. Woodrow. ''Estimates of Population Coverage in the 1990 United States Census Based on Demographic Analysis.''

Journal of the American Statistical Association 88 (September 1993): 1061–77.

Steffey, Duane L., and Norman M. Bradburn. *Counting People in the Information Age.* Washington, D.C.: National Academy Press, 1994.

U.S. Bureau of the Census. ''Accuracy and Coverage Evaluation: The Design Document.'' By Danny R. Childers. DSSD Census 2000 Procedures and Operations Memorandum Series Chapter S–DT–01.

____. ''Accuracy and Coverage Evaluation: Overview of Design.'' By Danny R. Childers and Deborah A. Fenstermaker, 11 January 2000.

____. ''Accuracy and Coverage Evaluation Survey: Dual System Estimation.'' By Donna Kostanich and Richard Griffin, 12 January 2000.

____. ''Accuracy and Coverage Evaluation Survey: Final Post-stratification Plan for Dual System Estimation.'' By Richard Griffin and Dawn Haines. DSSD Census 2000 Procedures and Operations Memorandum Series Chapter Q–24, 19 April 2000.

____. ''Accuracy and Coverage Evaluation Survey: Targeted Extended Search Plans.'' By Alfredo Navarro. DSSD Census 2000 Procedures and Operations Memorandum Series Chapter Q–18, 12 January 2000.

____. ''Additional Research on Accuracy of Adjusted Versus Unadjusted 1990 Census Base for Use in Intercensal Estimates.'' Addendum to Report of the Committee on Adjustment of Post Censal Estimates, 25 November 1992.

____. ''Analysis of CAPE Findings on 1990 PES Technical Issues.'' By Sally M. Obenski, 9 June 2000.

____. ''Analysis of CAPE Findings on PES Accuracy at Various Geographic Levels.'' By Sally M. Obenski and Robert E. Fay, 9 June 2000.

____. ''An Analysis of the Consistency of the 1990 Mosbacher Guidelines to U.S. Census Bureau Standards.'' By Sally M. Obenski and Robert E. Fay, 16 May 2000.

____. ''Assessment of Accuracy of Adjusted Versus Unadjusted 1990 Census Base for Use in Intercensal Estimates,'' Report of the Committee on Adjustment of Postcensal Estimates, 7 August 1992.

____. ''Characteristics of Census Errors.'' By Deborah Griffin and Christopher Moriarity. 1990 Decennial Census Preliminary Research and Evaluation Memorandum No. 179, 15 September 1992.

____. ''Master Activity Schedule for Census 2000.''

____. ''Minutes of the Executive Steering Committee on Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP),'' 5 January 2000.

____. ''Non-Response in Census Coverage Measurement Surveys and Its Impact ''An Historical Review.'' By Ruth Ann Killion. DSSD Briefs, Information, and Topics Memorandum Series No. 44C, 17 September 1998.

____. ''Reinventing the Decennial Census.'' Global Report of the Task Force for Planning the Year 2000 Census, June 1995.

____. ''Report of the Working Group on the Use of Demographic Analysis in Census 2000.'' By William R. Bell, Campbell J. Gibson, Prithwis Das Gupta, Gregory K. Spencer, J. Gregory Robinson, Mary H. Mulry, Elizabeth A. Vacca, Robert E. Fay, and Charlene A. Leggieri, 6 May 1996.

____. ''Report to Congress—The Plan for Census 2000.'' Originally issued July 1997, revised and reissued August 1997.

____. ''Some Results from the Census 2000 Dress Rehearsal.'' By Rajendra Singh. DSSD Census 2000 Dress Rehearsal Memorandum Series A–76, 26 February 1999

____. ''Technical Assessment of the Accuracy of Unadjusted Versus Adjusted 1990 Census Counts.'' Report of the Undercount Steering Committee, 21 June 1991.

____. ''Updated Summary: Census 2000 Operational Plan,'' February 1999.

U.S. Bureau of Labor Statistics. Abraham, Katharine, Letter to Harry A. Scarr, U.S. Department of Commerce, Bureau of the Census, 1 December 1993.

U.S. Congress. House. *Decennial Census Improvement Act of 1991,* 102nd Cong., 2nd sess., H.R. 3280. *Congressional Record.* Daily ed. (9 October 1991), H7694.

____. Committee on Government Reform and Oversight. *Prepared Testimony of K.W. Wachter and D.A. Freeman Before the House Committee on Government Reform and Oversight,* 104th Cong., 1st sess., 29 February 1996.

U.S. Department of Commerce. ''Adjustment of the 1990 Census for Overcounts and Undercounts of Population and Housing: Notice of Final Decision.'' **Federal Register** (22 July 1991) vol. 56 pp. 33582–642.

____. ''Census 2000 Dress Rehearsal Shows Undercount Persist; Scientific Methods Correct Race and Ethnic Differential.'' *Commerce News,* 20 April 1999, CB 99–CN. 16 (revised).

____. ''Decision of the Director of the Bureau of the Census on Whether to Use Information from the 1990 Post Enumeration Survey (PES) to Adjust the Base for the Intercensal Population Estimates Produced by the Bureau of the Census.'' **Federal Register** (4 January 1993) vol. 58, no. 1, pp. 69–78.

____. ''Final Guidelines for Considering Whether or not Statistical Adjustments of the 1990 Decennial Census of Population and Housing Should be Made for Coverage Deficiencies of the Population.'' **Federal Register** (15 March 1990) vol. 55, p. 9838.

____. ''Position on Adjustment of the 1980 Census Counts for Underenumeration.'' **Federal Register** (16 December 1980) vol. 45, no. 243, pp. 82872–85.

U.S. General Accounting Office. *Decennial Census: 1990 Results Show Need for Fundamental Reform.* Report to Congressional Requesters, 9 June 1992, 49 GAO/GGD–92–94.

U.S. President, Executive Office. Office of Management and Budget. ''Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity.'' **Federal Register** (30 October 1997) vol. 62, no. 210, pp. 58782–90.

Wachter, Kenneth, and David Freedman. ''Local Heterogeneity and Census Adjustment for the Intercensal Base.'' Technical Report No. 381. Berkeley, CA: University of California, Department of Statistics, 1993.

White, Andrew A., and Keith F. Rust, eds. *Sampling in the 2000 Census: Interim Report I.* Washington, D.C.: National Academy Press, 1996.

____. *Preparing for the 2000 Census: Interim Report II.* Washington, D.C.: National Academy Press, 1997.

Wolter, Kirk M. ''Some Coverage Error Models for Census Data.'' Journal of the American Statistical Association 81 (June 1986): 338–46.

Wright, Tommy and Joyce Farmer. ''A Bibliography of Selected Statistical Methods and Development Related to Census 2000.'' 3rd ed., 1 May 2000.

June 13, 2000.

MEMORANDUM FOR KENNETH PREWITT

From: William M. Daley

Attached is my decision adopting the analysis and conclusions set forth in ''Accuracy and Coverage Evaluation: Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of Census 2000.''

The Department of Commerce and the Census Bureau are committed to making certain that the decennial census, the largest peacetime mobilization in our country's history, produces the most accurate count possible of the individuals in our Nation. The census is an important civic undertaking designed to find out who we are and how we live. We owe it to the American people to use all of the tools at our disposal to make the census as accurate as possible.

For decades, the experts at the Census Bureau and within the statistical community have recognized that the methodology used in the past fails to count many Americans. This phenomenon '' called the undercount '' has been measured since the 1940s. More disturbing, however, is the established fact that the undercount operates differently for different population groups, creating an inequity called the differential undercount. Despite the Census Bureau's best efforts, the differential undercount has persisted and, for at least the last 50 years, has meant that some groups in the population are undercounted and therefore underrepresented in political, resource-allocation, and other decisions.

The choice we face is whether to use modern statistical methods to produce a more accurate census, or whether we do nothing. Under the law, statistical methods may not be used in tabulating the population for purposes of apportioning seats in the House of Representatives, but I am required to authorize the use of modern statistical methods, if ''feasible,'' for all other releases of census data.

The Director of the Census, with guidance from the Department concerning the relevant legal standard, has provided an analysis of the feasibility of using statistical sampling to correct the persistent errors in the census and to improve its accuracy—''Accuracy and Coverage Evaluation: Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of Census 2000.'' As explained in that document, absent the use of statistical methods there is no way to correct the persistent differential undercount in Census 2000. With established statistical methods, however, the Census Bureau believes that it will be able to correct these errors and improve the overall accuracy of the census by increasing coverage and reducing the differential undercount.

I hereby adopt the analysis and conclusions of the Director of the Census set forth in ''Accuracy and Coverage Evaluation: Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of Census 2000.'' As explained in that document, the expert staff at the Bureau believe that the Accuracy and Coverage Evaluation Survey, which was designed to measure and correct for the overall undercount and the differential undercount, should make the census counts more accurate. As is appropriate, however, no final decision about whether to correct the census counts can be made until the operations have been completed and considered by the Bureau. The Director will make a final decision before April 1, 2001, the deadline by which the Bureau must provide data to the States for redistricting.

I am also proposing today a regulation that will insulate the final decision on whether to correct the census counts from even the appearance of political tampering and will make the decision-making process as transparent as possible. Because the final decision on whether to correct the census is a technical decision, the proposed regulation would delegate my authority over that decision to the Director of the Census. His decision would, in turn, be informed by a public recommendation made by a group of career experts at the Census Bureau. Through this process, we will be able to ensure public confidence in the final decision.

Dated: June 13, 2000.

William M. Daley,

Secretary of Commerce.

June 12, 2000.

MEMORANDUM TO: The Secretary, The Director of the Census

FROM: Andrew J. Pincus

SUBJECT: Legal Obligation to Produce Statistically-Corrected Non-Apportionment Census Numbers

As you know, the Department of Commerce and the Census Bureau have been reviewing what process to use in determining whether to statistically correct census data for purposes other than apportionment of the House of Representatives. As part of this review, we have examined the legal requirements of the Census Act. After careful analysis, we have concluded that Section 195 of the Census Act requires the Census Bureau, if feasible, to produce statistically-corrected numbers from the decennial census for all non-apportionment purposes.

The feasibility determination is a technical decision that should be made by the Director, to whom the Secretary delegated his Title 13 responsibilities in Departmental Organizational Order 35–2A (July 22, 1987). To this end, we also believe it appropriate to propose a regulation that would make certain that the Director has final authority over the feasibility determination.

I. Background

The Constitution requires Congress to apportion seats in the House of Representatives among the States every ten years based on the results of the decennial census, providing that ''[t]he actual Enumeration shall be made within three Years after the first Meeting of the Congress of the United States, and within every subsequent Term of ten Years, in such Manner as they [Congress] shall by Law direct.'' [1] Through the Census Act, which is codified in title 13 of the United States Code, Congress has delegated its broad authority over the census to the Secretary of Commerce.[2] In particular, 13 U.S.C. 141(a) provides that the Secretary of Commerce shall take ''a decennial census of [the] population * * * in such form and content as he may determine, including the use of sampling procedures and special surveys.'' As the Supreme Court recognized in *Wisconsin* v. *City of New York,* the Secretary's determination as to how to conduct the Census, pursuant to the delegation of authority provided to him by Congress, need only be reasonable, so long as it is also ''consistent with the constitutional language and the constitutional goal of equal representation.'' *Id.* at 19. The Court further recognized, in the context of the Secretary's decision in 1990 not to adjust the census, that the ''Constitution itself provides no real instruction'' on what methods the Secretary should use in performing the Census. *Id.* at 18.

II. Section 195 of the Census Act Requires the Census Bureau To Use Sampling When ''Feasible'' For Calculating the Population For Purposes Other Than Apportionment of Seats in the House of Representatives Among the States

Section 195 of the Census Act states:

Except for the determination of population for purposes of apportionment of Representatives in Congress among the several States, the Secretary shall, if he considers it feasible, authorize the use of the statistical method known as ''sampling'' in carrying out the provisions of this title.

13 U.S.C. 195. Section 195 refers specifically to only one of the many uses of census data. Decennial census data are used not only by the U.S. Congress for apportioning seats in the House of Representatives among the States, but also by the States in drawing the lines for congressional and state and local legislative districts, and by federal and state agencies in allocating funds.

In *Department of Commerce* v. *House of Representatives,* 119 S. Ct. 765 (1999), the Supreme Court held that Section 195 does not permit the use of sampling to produce population counts for the purpose of apportioning seats in the House of Representatives among the States. *Id.* at 777 (''there is only one plausible reading of the amended § 195: It prohibits the use of sampling in calculating the population for purposes of apportionment.''). Here, the question is what standard Section 195 applies with respect to the calculation of population by the Census Bureau for purposes *other than* ''apportionment of Representatives in Congress among the several States.'' The plain language of the provision supplies the answer: Section 195 states that the Secretary ''shall'' authorize the use of statistical sampling for all other purposes ''if he considers it feasible.'' Thus, when calculating population or other information for a purpose other than apportionment, the Secretary (or his designee, the Census Bureau) must first determine whether it is ''feasible'' to use sampling, and—if the use of sampling is feasible—its use must be authorized.

This interpretation of Section 195's plain language is confirmed by Congress's amendment of the provision in 1976. Prior to that amendment, Section 195 stated:

Except for the determination of population for apportionment purposes, the Secretary may, where he deems it appropriate, authorize the use of the statistical method known as 'sampling' in carrying out the provisions of this title.

The pre-1976 wording (''may, where he deems it appropriate'') gave the Secretary the option of using sampling. The 1976 amendment eliminated the Secretary's discretion, transforming Section 195 into a mandatory directive —the Secretary ''shall * * * authorize the use of'' sampling for all other purposes ''if he considers it feasible.'' The Census Act therefore unambiguously requires, with respect to non-apportionment calculations, that when sampling is feasible, it must be used.

The Supreme Court's recent decision in *Department of Commerce* v. *House of Representatives* confirms this conclusion. In explaining the purpose of the 1976 amendments, the Court stated, ''[t]hey changed a provision that permitted the use of sampling for purposes other than apportionment into one that required that sampling be used for such purposes if 'feasible.' 119 S.Ct. at 778. The Court explained that

[1] Constitution, Art. I, Sec. 2, Cl. 3.

[2] *Wisconsin* v. *City of New York,* 517 U.S. 1 (1996).

''section [195] now requires the Secretary to use statistical sampling in assembling the myriad demographic data that are collected in connection with the decennial census. But the section maintains its prohibition on the use of statistical sampling in calculating population for purposes of apportionment.'' 119 S.Ct. at 777.

III. The Census Bureau's Calculation of Population for the Purpose of Redistricting is Subject to Section 195's ''Feasib[ility]'' Standard

Section 141(c) of the Census Act permits the ''officers or public bodies having initial responsibility for the legislative apportionment or districting of each State'' to submit to the Secretary ''a plan identifying the geographic areas for which specific tabulations of population are desired.'' The same provision directs the Secretary to report such ''[t]abulations of population,'' as well as the ''basic tabulations of population'' for States that have not submitted a plan, within one year of the decennial census date. It is clear that these population tabulations are not ''the determination of population for purposes of apportionment of Representatives in Congress among the several States'' (Section 195), and therefore are subject to Section 195's directive that the use of sampling ''shall'' be authorized if ''feasible.''

To begin with, the population tabulations supplied to the States pursuant to Section 141(c) simply are *not* made or used for purposes of apportioning seats in the House of Representatives among the States. Section 141(c) makes clear that it relates to tabulations for ''legislative apportionment or districting of each State.'' And a separate subsection of Section 141—subsection (b)—governs the ''tabulation of total population by States * * * as required for the apportionment of Representatives in Congress among the several States.'' Indeed, the distinction between these two groups of calculations is confirmed by their different due dates: the latter set of numbers must be completed three months earlier than the redistricting information required by Section 141(c). See also Section 141(e)(2) (distinguishing between use of census data for ''apportionment of Representatives in Congress among the several States'' and for ''prescribing congressional districts'').

Some commentators have suggested that the term ''apportionment'' within Section 195's ''[e]xcept'' clause encompasses population calculation for the purposes of redistricting as well as for the purpose of allocating seats in the House of Representatives among the States. That position is inconsistent with the plain language of the statute. First, it ignores the clear distinction in Section 141 between these two categories of calculations. Second, Congress in 1976 revised the ''[e]xcept'' clause, replacing the word ''apportionment'' with the phrase ''apportionment of Representatives in Congress among the Several States.'' It is difficult to imagine how Congress could have more clearly evidenced its intent to limit Section 195's prohibition against the use of sampling to the calculation of population used to allocate among the States seats in the House of Representatives. And because Section 141(c) specifically refers to tabulations for redistricting purposes, but that reference does not appear in the ''[e]xcept'' clause of Section 195, it is plain that redistricting tabulations are not encompassed within the Section 195 prohibition. [3]

Finally, some commentators have suggested that as a practical matter these two sets of numbers are inextricably linked, asserting—for example—that it would be a plainly improper result if the Section 141(c) population tabulation of a State for redistricting purposes did not equal the Section 141(b) apportionment population tabulation for that State. Nothing in the Census Act requires that result and, moreover, the two totals have not been equal in the past. For example, government personnel stationed overseas are included in a State's Section 141(b) tabulation, but are not included in the data provided to that State under Section 141(c). Congress could have required such equality in either Section 141 or Section 195, but it did not do so. Rather, Congress in Section 141 expressly distinguished between the two categories of calculations.

The Census Act thus clearly directs that statistical sampling ''shall'' be used in tabulating population for the purposes set forth in Section 141(c) if the Secretary considers it ''feasible'' to do so. Even if the plain language of the Act were not clear on this point, we believe that this interpretation is most consistent with the purposes of the Census Act and that adopting such an interpretation is within your discretion. In *Wisconsin* v. *City of New York*, 517 U.S. 1 (1996), the Supreme Court unanimously concluded ''the wide discretion bestowed by the Constitution upon Congress, and by Congress upon the Secretary,'' mandates substantial judicial deference to the Secretary's determinations with respect to the decennial census (517 U.S. at 19). Given the long history of the use of sampling by the Census Bureau, and the importance of obtaining the most accurate population tabulations possible—because of the constitutional significance of the ''one person, one vote'' principle and of the equal protection principles reflected in the Voting Rights Act—interpreting the statute to permit the use of sampling when feasible is the most appropriate approach. The alternative interpretation would bar the use of statistical sampling even if the use of sampling would lead to more accurate results, a construction that conflicts with the basic goal of the decennial census—to obtain an accurate count of the persons within the United States.

IV. The Standard For the Feasibility Determination

Section 195 does not contain a definition of the term ''feasible.'' The dictionary definition of the term ranges from the most common ''capable of being done or carried out'' to ''capable of being used or dealt with successfully, suitable'' or ''reasonable, likely.'' Webster's Ninth New Collegiate Dictionary (1990). The Supreme Court has considered the word ''feasible'' in other contexts and found that the plain meaning of the term generally denotes the first and broadest definition—''capable of being done.'' In *American Textile Mfrs. Institute, Inc.* v. *Donovan,* 452 U.S. 490, 509 (1981), the Court interpreted the term ''to the extent feasible'' to preclude the Secretary of Labor from engaging in a cost-benefit analysis of a public health standard; as the Court explained, Congress itself, by requiring a standard ''to the extent feasible'' had made the policy choice for the Secretary. *See also Citizens to Preserve Overton Park, Inc.* v. *Volpe,* 401 U.S. 402, 411 (1971) (''the requirement that there be no 'feasible

[3] Some commentators have argued that the Supreme Court reached a different conclusion in Department of Commerce because it found standing ''on the basis of the expected effects of the use of sampling in the 2000 census on intrastate redistricting'' (119 S. Ct. at 774). The Court's standing decision, however, simply reflects a conclusion that an individual claiming injury by the use of that data for redistricting had alleged sufficient Article III injury in fact to challenge the plan. But the Census plan before the Court provided for the collection and production of a single set of sampling-adjusted data for use in both the apportionment tabulation and the redistricting tabulation. Because the Court invalidated the plan, there was no need for the Court to apply Section 195 to the use of sampling for redistricting purposes in order to redress these plaintiffs' purported injury. This conclusion is confirmed by the Court's careful limitation of its holding: ''The District Court below * * * concluded that the proposed use of statistical sampling to determine population for purposes of apportioning congressional seats among the States violates the Act. We agree.'' 119 S. Ct. at 765.

alternative' route admits of little administrative discretion.'').

We understand the term ''feasible'' in accordance with its ordinary meaning and the overall purposes of the Census Act. It also must be understood in terms of the uses to which non-apportionment census data are put, including, among other things, redistricting and allocation of federal funds. While in other contexts it might be appropriate to understand ''feasible'' to mean ''possible,'' given the obvious importance of obtaining the most accurate population (and other) tabulations possible, it would seem most appropriate to construe that term in a manner that focuses upon promoting accurate census results. [4] Thus, with respect to the proposed use of statistical sampling for data to be released to the States under Section 141(c), such use is ''feasible'' within the meaning of Section 195 if (1) the proposed use of sampling is compatible with the other aspects of the census plan, and with any statutory, timing, and funding constraints; and (2) the proposed use of statistical sampling would improve the overall accuracy of the census data.

The two components of ''feasibility '' can be termed ''operational feasibility'' and ''technical feasibility.'' These are matters that are properly within the expert judgment of the Census Bureau. The Census Bureau's extensive experience in the conduct of the census, the use of statistical sampling techniques, and the measurement of accuracy should be the basis for these essentially technical judgments.

V. The Decisionmaking Process

The determination whether the use of sampling is ''feasible'' under Section 195 should be based upon the information before the decisionmaker at the time the determination is made. Public Law No. 94–171 requires the Census Bureau to deliver official census data to the states for redistricting purposes by April 1, 2001. 13 U.S.C. 141(c). As with every decennial census, the Census Bureau will conduct extensive analyses on the census data in the ensuing years. In order to make a final decision on whether to deliver statistically corrected data for redistricting purposes, the Census Bureau need only consider the evidence available to it at the time of its decision to determine whether the statistically corrected numbers are more accurate and therefore that the use of sampling is ''feasible'' as that term is defined herein. *See, e.g., Vermont Yankee Nuclear Power Corp.* v *Natural Resources Defense Council,* 435 U.S. 519, 552–54 (1978) (review of agency decision must be made based on information available at the time the decision was made); *ICC* v. *New Jersey,* 322 U.S. 503, 514 (1970).

The Census Bureau is in the process of completing a document which will provide information concerning its assessment of whether using statistically sampling is feasible with respect to the release of P.L. 94–171 data. Although, as the document will indicate, the Census Bureau has determined that the use of statistical sampling is operationally feasible and should improve the accuracy of the census, no final decision will be made with respect to the release of data until after the Bureau has had the opportunity to review whether census operations were conducted in a way that met expectations. This document will be published in the **Federal Register**, along with a proposed regulation that would delegate to the Director of the Census the Secretary of Commerce's authority to make the final, technical decision on what numbers to release and would set forth a process for the Census Bureau's consideration of what numbers to release.

Robert J. Shapiro,

Under Secretary for Economic Affairs.

[FR Doc. 00–15348 Filed 6–14–00; 8:45 am]

BILLING CODE 3510–07–P

[4] Of course, in other contexts where there is no independent requirement that the population court be conducted without the use of sampling (unlike the decennial census, where the statue as construed by the Supreme Court prohibits the use of sampling for apportionment of seat in the House of Representatives), the analysis might also take greater account of the efficiencies that could be gained by substituting sampling for those other methods.