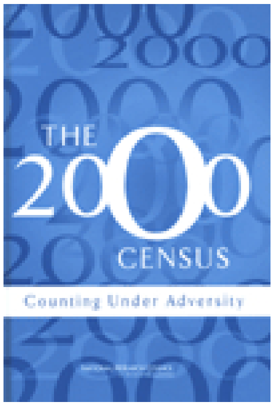

The 2000 Census: Counting Under Adversity

Constance F. Citro, Daniel L. Cork, and Janet L. Norwood, Editors, Panel to Review the 2000 Census, National Research Council

ISBN: 0-309-52998-0, 621 pages, 6 x 9, (2004)

THE 2000 CENSUS

Counting Under Adversity

Panel to Review the 2000 Census

Constance F. Citro, Daniel L. Cork, and Janet L. Norwood, *Editors*

Committee on National Statistics

Division of Behavioral and Social Sciences and Education

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS **500 Fifth Street, NW** **Washington, DC 20001**

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

The project that is the subject of this report was supported by contract no. 50-YABC-8-66010 between the National Academy of Sciences and the U.S. Census Bureau. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number 0-309-09141-1 (book)
International Standard Book Number 0-309-52998-0 (PDF)

Library of Congress Control Number: 2004102206

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, NW, Washington, D.C. 20001; (202) 334-3096; Internet, http://www.nap.edu.

Printed in the United States of America.

Suggested citation: National Research Council (2004). *The 2000 Census: Counting Under Adversity.* Panel to Review the 2000 Census. Constance F. Citro, Daniel L. Cork, and Janet L. Norwood, eds. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Wm. A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. Wm. A. Wulf are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

PANEL TO REVIEW THE 2000 CENSUS

JANET L. NORWOOD *(Chair)*, Chevy Chase, Maryland
ROBERT M. BELL, AT&T Labs–Research, Florham Park, New Jersey
NORMAN M. BRADBURN, National Science Foundation, Arlington, Virginia
LAWRENCE D. BROWN, Department of Statistics, The Wharton School, University of Pennsylvania
WILLIAM F. EDDY, Department of Statistics, Carnegie Mellon University
ROBERT M. HAUSER, Department of Sociology, University of Wisconsin–Madison
RODERICK J.A. LITTLE,* School of Public Health, University of Michigan
INGRAM OLKIN, Department of Statistics and School of Education, Stanford University
D. BRUCE PETRIE, Canadian Institute for Health Information, Ottawa, Ontario

CONSTANCE F. CITRO, *Study Director*
MICHAEL L. COHEN, *Senior Program Officer*
DANIEL L. CORK, *Program Officer*
AGNES GASKIN, *Senior Project Assistant*
MARISA GERSTEIN, *Research Assistant*
SETH HAUSER, *Research Associate*
MICHELE VER PLOEG, *Program Officer*
MEYER ZITTER, *Consultant*

*Served until March 2000.

COMMITTEE ON NATIONAL STATISTICS
2004

Acknowledgments

The Panel to Review the 2000 Census wishes to thank the many people who have contributed to the panel's work and helped make possible the preparation of this final report.

We thank, first, staff of the U.S. Census Bureau who prepared a large number of evaluation reports about the census, the Accuracy and Coverage Evaluation (A.C.E.) Program, and demographic analysis, made informative presentations at panel meetings and workshops, and answered many specific questions about census procedures and evaluations. We thank particularly William Bell, Cynthia Clark, Donald Dalzell, Robert Fay, Philip Gbur, Howard Hogan, Ruth Ann Killion, Louis Kincannon, Joseph Knott, Donna Kostanich, John Long, Susan Love, Mary Mulry, J. Gregory Robinson, and Preston J. Waite. Former deputy director and acting director William Barron, former director Kenneth Prewitt, and former associate director for decennial census John Thompson also made valuable contributions to the panel's meetings and workshops. Rajendra Singh has been very helpful as the Census Bureau's project officer throughout the study.

We further thank the Census Bureau for arranging for panel members and staff to have access to key data files for analysis, beginning in February 2001. Such access was provided not only to the panel but also to congressional oversight groups, under procedures to safeguard confidentiality. The ability to analyze key data sets provided knowledge of census and A.C.E. procedures and evaluations that would not otherwise have been possible for the panel to obtain.

We thank others in the professional community who participated in panel workshops: Barbara Bailar, National Opinion Research Center (retired); Stephen Fienberg, Carnegie Mellon University; David Freedman, University of California, Berkeley; Charles Jones, U.S. Census Monitoring Board, Congressional Members (retired); Graham Kalton, Westat; Jeffrey Passel, Urban Institute; Allen Schirm, Mathematica Policy Research, Inc.; Joseph Sedransk, Case Western Reserve University; Bruce Spencer, Northwestern University; Philip Stark, University of California, Berkeley; Michael Stoto, RAND; Joseph Waksberg, Westat; Martin Wells, Cornell University; Kirk Wolter, National Opinion Research Center (retired); Donald Ylvisaker, University of California, Los Angeles; and Alan Zaslavsky, Harvard Medical School.

We also thank David Harris, University of Michigan, and Joseph Salvo, New York City Department of City Planning, for major contributions to the panel's work. David prepared an insightful paper on the measurement and definition of race and ethnicity in federal statistics and the census. Joe ably chaired a working group on the Local Update of Census Addresses (LUCA) Program commissioned by the panel to evaluate LUCA from the local government perspective.

Throughout its work, the panel benefited from discussions with congressional staff, particularly during the period when oversight authority for the census was vested by the U.S. House of Representatives in a Subcommittee on the Census of the Committee on Government Reform. In particular, we thank David McMillen and Michael Miguel for their insights. We have also benefited from interactions with staff of the U.S. General Accounting Office.

The panel is especially indebted to Constance Citro who, as senior study director, organized the work of the panel and guided its evaluation of the 2000 census. Her wide experience in census issues, her competence in statistical methods, and the clarity of her reasoning have been critical to the successful completion of our interim report and now our final report. We have benefited enormously from her talent and knowledge and feel extremely fortunate to have had her work with us.

The panel was assisted by a very able staff. Daniel Cork played a major role for the panel in conducting analyses of data files from the A.C.E., analyzing 1990 and 2000 census data on mail return rates, drafting text for the report, developing informative graphs of key results, and preparing the report for release. His hard work and contributions, achieved under tight time pressures, were extraordinary. Andrew White, former director of the Committee on National Statistics, served as study director for the panel from November 1998 through March 2000. He was assisted by Michael Cohen, who organized three panel workshops and contributed to the panel's work throughout, particularly to the text on issues of evaluation and imputation methods. Meyer Zitter contributed to the panel's assessments of demographic analysis and the procedures for developing the Master Address File. He also specified and analyzed tables of comparable 1990 and 2000 census item imputation rates for the long-form sample. Michele Ver Ploeg and Marisa Gerstein assisted in data analysis, as did Zhanyun Zhao, University of Pennsylvania. Seth Hauser, now with the U.S. State Department, assisted in the analysis of item imputation rates and drafted text on the measurement of race and ethnicity. Heather Koball, now with the National Center for Children in Poverty at Columbia University, prepared background material for the panel on race and ethnicity and organized and assessed trips for the panel and staff to observe census and A.C.E. operations in January–June 2000. Carrie Muntean, now with the U.S. Foreign Service, prepared background material for the panel on the development of the 1990 and 2000 census address lists and provided invaluable support to the panel's commissioned working group on the LUCA Program. Joshua Dick, Jamie Casey, and Agnes Gaskin provided valuable project assistance to the panel, particularly in making arrangements for the panel's workshops. Christine McShane, senior editor of the reports office of the Division of Behavioral and Social Sciences and Education, made important contributions to the report through her fine technical editing. To all we are grateful.

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the Report Review Committee of the National Research Council (NRC). The purpose of this independent review is to provide candid and critical comments that will assist the institution in making the published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge.

The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

We thank the following individuals for their participation in the review of this report: Vincent P. Barabba, Chairman, Market Insight Corporation and the Internet Home Alliance, Palo Alto, CA; Joseph B. Kadane, Department of Statistics, Carnegie Mellon University; Kenneth Prewitt, Department of International and Public Affairs, Columbia University; Bruce D. Spencer, Department of Statistics, Northwestern University; Michael A. Stoto, Center for Domestic and International Health Security, RAND, Arlington, VA; James Trussell, Office of Population Research, Princeton University; Donald Ylvisaker, Department of Statistics, University of California, Los Angeles; and Alan M. Zaslavsky, Department of Health Care Policy, Harvard Medical School.

Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report before its release. The review of the report was overseen by John C. Bailar III, Professor Emeritus, Department of Health Studies, The University of Chicago. Appointed by the National Research Council, he was responsible for making certain that an independent examination of the report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring panel and the institution.

Janet L. Norwood, *Chair*
Panel to Review the 2000 Census

Contents

List of Figures

List of Tables

List of Boxes

Executive Summary

The decennial census is the federal government's largest and most complex peacetime operation. Mandated in the U.S. Constitution to provide the basis for reapportioning seats in the U.S. House of Representatives, the census has many purposes today: redrawing congressional and state legislative district boundaries; allocating federal and state program funds; planning and evaluating federal, state, and local government programs; providing the basis for updated population estimates; and informing researchers, the private sector, the media, and the public about the characteristics of population groups and geographic localities and how they have changed over time.

The National Research Council's Committee on National Statistics established the Panel to Review the 2000 Census in 1998 at the request of the U.S. Census Bureau. The panel was given a broad charge:

> to review the statistical methods of the 2000 census, particularly the use of the Accuracy and Coverage Evaluation Program and dual-systems estimation, and other census procedures that may affect the completeness and quality of the data.

The panel's findings in this final report cover the planning process for 2000, major innovations in census operations, the completeness of population coverage in 2000, and the quality of both the basic demographic data collected from all census respondents and the detailed socioeconomic data collected from the census long-form sample (about one-sixth of total households). We recommend improvements in the planning process and design for the 2010 census that flow from our evaluations of 2000; however, our recommendations are not intended to be comprehensive in terms of 2010 planning. The Panel on Research on Future Census Methods is assessing the Census Bureau's plans for 2010.

Our overall conclusion is that the 2000 census experienced both major successes and significant problems. Despite problems with the census address list, the completeness of population coverage and the quality of the basic demographic data were at least as good as and possibly better than in previous censuses. However, many of the long-form items on such topics as income and employment, as well as the data for residents of college dormitories, prisons, nursing homes, group homes, and other group quarters, were no better and in some cases worse than in previous censuses. That we found problematic areas should not detract from the success of the 2000 census in providing relevant, timely data for many uses. It should also make clear the importance of a thorough evaluation of the quality of census data and operations for users to understand the census information and for planners to improve future censuses.

THE ROAD TO 2000

The 2000 census planning began in a climate of concern about the perceived failures of the 1990 census—one that saw a substantial decline in public cooperation and, despite higher per household costs than in the 1980 census, resulted in worse coverage of minorities, renters, and children relative to other population groups. The Census Bureau's initial design to remedy these problems for 2000 relied on much greater use of statistical techniques in the census enumeration, but this plan encountered opposition from members of Congress and others. As a result, the Bureau had to contend with externally imposed last-minute changes in design, delayed budget decisions, consequent changes in plans, and insufficient time for operational trials. All of these problems increased not only the costs of the 2000 census but also the risk that it could have been seriously flawed in one or more respects.

In light of this experience, the panel recommends that the Census Bureau, the administration, and Congress agree on the basic design for the 2010 census no later than 2006 in order to permit an appropriate, well-planned dress rehearsal in 2008. In particular, this agreement should specify the role of the new American Community Survey (ACS). Further delay will undercut the ability of the ACS to provide, by 2010, small-area data of the type traditionally collected on the census long-form sample and will jeopardize 2010 planning, which currently assumes a short-form-only census (Recommendations 3.1 and 3.2).

ASSESSMENT OF 2000 CENSUS OPERATIONS

We find that the 2000 census was generally well executed despite many obstacles and even though agreement on the final 2000 design was not reached until spring 1999. The dedication of the Census Bureau staff made possible the success of several operational innovations. Because of the last-minute changes in plans, the ample funding appropriated by Congress in spring 1999 was essential for the successful execution of the census.

Two major successes of the 2000 census were that the decline in mail response rates observed in the 1980 and 1990 censuses was halted and that most operations were well conducted and completed on or ahead of schedule. The use of a redesigned questionnaire and mailing strategy and of an expanded advertising and outreach program contributed to the improved mail response rate, which, in turn, helped reduce the cost and time of follow-up activities. Contracting for selected data operations, using improved technology for capturing the data on questionnaires, and aggressively recruiting enumerators and implementing nonresponse follow-up—all innovations in 2000—contributed to the timely execution of the census.

Two major operational problems of the 2000 census were the error-plagued development of the Master Address File (MAF) and the poorly managed enumeration of residents of group quarters. The use of multiple sources to build a Master Address File—a major innovation in 2000—was appropriate in concept but not well executed. The Bureau's early plans for constructing the MAF assumed that the U.S. Postal Service's Delivery Sequence File would provide most of the needed updates to the 1990 list in mailout/mailback areas, but the Postal Service list was not complete or accurate enough for this purpose. The consequent changes in schedules and operations and the variability in local efforts to update the MAF contributed to census enumeration errors, including a large number of duplicates. The number of duplicates would have been larger yet if the Bureau had not mounted an ad hoc operation in summer 2000 to weed out duplicate MAF addresses.

The census is the only comprehensive source of information on the group quarters population, which constitutes large proportions of some communities. The group quarters enumeration in 2000 was not well conducted, experiencing such deficiencies as poorly controlled tracking of enumerations and errors in assigning group quarters to the correct geographic areas. There was no program for evaluating the completeness of coverage of group quarters residents.

Looking to improve census operations in 2010 from the experience in 2000, we recommend that the Census Bureau give high priority to research in three areas:

(1) *Targeted second questionnaire:* Because a second questionnaire sent to nonresponding households is known to significantly improve response rates—but feasibility issues prevented its use in 2000—the Census Bureau must proceed quickly to work with vendors to develop cost-effective, timely ways to mail a second questionnaire to nonresponding households in the 2010 census, in a manner that minimizes duplicate enumerations (Recommendation 4.1).

(2) *Improved MAF development procedures:* Because a complete, accurate Master Address File is critical not only for the 2010 census but also for the 2008 dress rehearsal, the new American Community Survey, and other Census Bureau surveys, the Bureau must develop more effective procedures for updating and correcting the MAF than were used in 2000. In particular, the Bureau must develop procedures for obtaining accurate information to identify housing units within multiunit structures and redesign its Local Update of Census Addresses partnership program to benefit state and local governments that participate (Recommendation 4.3).

(3) *Better enumeration of group quarters:* Because of the growing importance of group quarters populations, the Census Bureau must thoroughly evaluate and completely redesign the processes related to them for the 2010 census, adapting the design as needed

for different types of group quarters. This effort should include consideration of clearer definitions for group quarters, redesign of questionnaires and data content as appropriate, and improvement of the address listing, enumeration, and coverage evaluation processes for group quarters (Recommendation 4.4).

ASSESSMENT OF POPULATION COVERAGE IN 2000

Much valuable information is available from the 2000 Accuracy and Coverage Evaluation (A.C.E.) Survey about the completeness of the census enumeration for the household population. Changes in estimation methods between the most recent A.C.E. Revision II, completed in March 2003, and the 1990 Post-Enumeration Survey (PES) make it difficult to compare coverage estimates between the two censuses. Nonetheless, there is sufficient evidence to conclude that net undercount rates for population groups were reduced in 2000 from 1990 and, even more important, that differences in net undercount rates between historically less-well-counted groups (minorities, children, renters) and other groups were reduced as well. From this result and from analysis of available information for states and large counties and places, it is reasonable to infer that differences in net undercount rates among geographic areas were also probably smaller in 2000 compared with 1990.

The reduction in differential net undercount was an achievement of the 2000 census. Contributing to this outcome, however, were large numbers of duplicate census enumerations and large numbers of wholly imputed census records (three times as many such records as in 1990, including imputations for members of enumerated households who lacked all basic characteristics as well as imputations of people into housing units that were presumed to be occupied). Both the duplicate enumerations and the imputed records generally accounted for larger proportions of historically less-well-counted groups than of other groups.

Despite the reduced differences in net undercount rates and an estimate that the 2000 census *overcounted* the total population by a small percentage (a first in census history), such groups as black men and renters continued be undercounted in 2000, whereas other groups were overcounted. Reducing such differences will be an important goal for the 2010 census. Also in 2000, gross errors of erroneous enumerations and omissions were almost as high as in 1990. The distribution of coverage errors across geographic areas is hard to discern from the available information; there is no information about coverage of group quarters residents.

The Census Bureau's Three Decisions Against Adjustment

In March 2001, October 2001, and March 2003, the Census Bureau announced that it would not adjust the 2000 census results for incomplete coverage of some population groups (and overcounting of other groups). In the panel's view, all three decisions are justified. The Bureau's March 2001 and October 2001 decisions are justified given: (1) its conclusion in March that evaluation studies were not sufficient to determine the accuracy of the A.C.E. population estimates and (2) its conclusion in October, after further study, that the original A.C.E. population estimates were too high. The Bureau's March 2003 decision not to use the A.C.E. Revision II coverage measurement results to adjust the 2000 census base counts for the Bureau's postcensal population estimates program is justified as well. Although the Revision II estimation work was thorough, innovative, and of high quality, the results are too uncertain to be used with sufficient confidence about their reliability for adjustment of census counts for subnational geographic areas and population groups.

Sources of uncertainty stem from the small samples of the A.C.E. data that were available to correct components of the original A.C.E. estimates of erroneous enumerations and non-A.C.E. residents and the consequent inability to make these corrections for other than very large population groups; the inability to determine which of each pair of duplicates detected in the A.C.E. evaluations was correct and which should not have been counted in the census or included as an A.C.E. resident; the possible errors in subnational estimates from the choice of one of several alternative "correlation bias" adjustments to compensate for higher proportions of missing men relative to women; the inability to make correlation bias adjustments for population groups other than blacks and nonblacks; and the possible errors for some small areas from the use of different population groups for estimating erroneous census enumerations and census omissions.

That estimated net undercount rates and differences in net undercount rates for population groups (and, by inference, subnational areas) were smaller in 2000 than in 1990 contributed to the panel's agreement with the Census Bureau's decision not to use the A.C.E. Revision II results to adjust the 2000 census counts. The smaller the measured coverage errors in the census, the smaller would be the effects of an adjustment on important uses of the census. Because the benefits of an adjustment are less when coverage errors are small, a high level of confidence is needed that an adjustment would not significantly increase the census errors for some areas and population groups. In our judgment, the A.C.E. Revision II estimates, given the constraints of the available data for correcting the original A.C.E. estimates, are too uncertain for use in this context.

Performance of the Accuracy and Coverage Evaluation

The original A.C.E. was well conducted, and, in many respects, an improvement over the 1990 PES, achieving such successes as timely data collection, low missing data rates, and improved quality of matching. However, inaccurate reporting of household residence in the A.C.E.—which also occurred in the census—led the original A.C.E. to substantially underestimate duplicate census enumerations in 2000. Many of these duplications occurred because respondents were confused about or misinterpreted census residence rules, so that, for example, college students were often counted at home as well as at their residence hall, and people with both summer and winter homes were often counted twice.

The Census Bureau commendably dedicated resources to the A.C.E. Revision II effort, which used the original A.C.E. and evaluation data to completely reestimate net undercount (or overcount) rates for several hundred population groups (poststrata). This work exhibited high levels of creativity and effort devoted to a complex problem. From innovative use of matching technology and other evaluations, it provided substantial additional information about the numbers and sources of erroneous census enumerations and, similarly, information with which to correct the residency status of the independent A.C.E. sample. It provided little additional information, however, about the numbers and sources of census omissions.

Demographic Analysis

Demographic analysis uses records of births, deaths, and net immigration to produce updated population estimates for age and sex groups for blacks and for all others. The original 2000 demographic analysis estimates were subsequently revised in October 2001. There are sufficient uncertainties in the revised net immigration estimates (particularly the illegal component) and a revised assumption of completeness of birth registration after 1984, compounded by the difficulties of classifying people by race, so that the revised demographic analysis estimates cannot serve as the definitive standard of evaluation for the 2000 census or the A.C.E. However, the estimates were useful in helping to identify coverage problems in the census and the A.C.E., and the method is important as the basis for postcensal population estimates.

Coverage Evaluation in 2010

A large postenumeration survey is an essential part of a coverage measurement program for the U.S. census. Demographic analysis, which is limited to producing net national-level coverage estimates for blacks and all others by age and sex, cannot substitute for a postenumeration survey. Underscoring the need for a large survey in 2010 is the prospect that the Census Bureau will make extensive use of methods developed from 2000 A.C.E. research on matching census records by name and birthdate to reduce duplicate enumerations in the 2010 census. If that is done and if in 2010 the number of omissions were to be as large as or larger than in 2000, the consequence could be a substantial net undercount of the population and an increase in differential undercoverage among population groups. In those circumstances, it would be essential to have an improved coverage evaluation survey built on the experience of the 2000 A.C.E. to understand the coverage achieved in 2010 and to inform the public about it. Data from an improved survey would be available to develop reasonably reliable estimates of coverage errors, which would be needed not only to explain the accuracy of the census but also to permit adjustment of some or all of the data should that be considered necessary.

We therefore recommend that the Census Bureau and the administration should request, and Congress should provide, funding for the development and implementation of an improved A.C.E. Program for the 2010 census. The A.C.E. survey in 2010 should be large enough to provide estimates of coverage errors that provide the level of precision targeted for the original (March 2001) A.C.E. estimates for population groups and geographic areas. An important area for improvement is the estimation of components of gross census errors (including types of erroneous enumerations and omissions), as well as net error. Understanding gross errors is valuable for identifying areas for improvement in census processes. Other important areas for improvement include the inclusion of group quarters in the A.C.E. (they were excluded in 2000); improved questionnaire content and interviewing procedures about place of residence; a simpler procedure for treating people who moved between Census Day and the A.C.E. interview; the development of poststrata for estimation of net coverage errors by using census results and statistical modeling as appropriate; and the investigation of possible correlation bias adjustments for additional population groups (Recommendation 6.1). We also recommend that the Census Bureau strengthen its program to improve demographic analysis estimates, particularly estimates of net immigration and measures of uncertainty for the demographic results (Recommendation 6.2).

It is clear from the experience with the 2000 A.C.E. that evaluation of coverage errors in the census takes time. It also takes time to evaluate census processes and the quality of the census data for individual items. In the panel's view, adequate evaluation of the census block-level data for congressional redistricting is not possible by the current deadline. Congress should consider moving this deadline to allow more time for evaluation of the completeness of population coverage and quality of the basic demographic items before they are released (Recommendation 6.3). Users should be aware that census counts at the block level—whether based on the enumeration or adjusted for coverage errors—are subject to high levels of variability; these data should be aggregated to larger geographic areas for use.

ASSESSMENT OF COMPLETE-COUNT AND ADDITIONAL LONG-FORM-SAMPLE DATA

The 2000 census data for items asked of everyone (age, sex, race, ethnicity, household relationship, housing tenure) were quite complete at the national level—missing data rates ranged from 2 to 5 percent (including records with one or more missing items and people who were wholly imputed). Rates of inconsistent reporting for these items were also low. However, some population groups and geographic areas exhibited high rates of missing data and inconsistent reporting for the basic items.

Missing data rates for the 2000 census long-form items were moderately to very high (10 to 20 percent or more) for over one-half of the items, which cover such topics as citizenship, veterans status, migration, disabilities, education, work experience, transportation to work, occupation, income, housing characteristics, and others. Missing data rates also varied widely among population groups and geographic areas. By comparison with 1990, missing data rates were higher in 2000 (often substantially so) for most of the long-form items asked in both years. Missing data rates for long-form items for group quarters residents in 2000 were much higher than the rates for household members and considerably higher than the rates for group quarters residents in 1990. Only scattered evidence is available as yet about possible systematic reporting errors for long-form items (e.g., systematic underreports or overreports of income or education).

One reason for higher missing data rates in 2000 than in 1990 is that the 2000 census placed greater reliance on imputation routines to supply values for missing and inconsistent responses, in contrast to the greater reliance on telephone and field follow-up to obtain answers for missing items in 1990. The greater reliance on imputation in 2000 contributed to the timely execution of the census. Whether it produced distributions for characteristics that were less accurate than the distributions that would have been produced from additional follow-up is not known. There is scattered evidence that the imputation procedures sometimes produced untoward results—for example, extremely high unemployment rates for some types of group quarters residents.

We recommend that the Census Bureau should develop procedures to quantify and report the variability of the 2000 long-form-sample estimates due to imputation, in addition to the variability due to sampling and whole-household nonresponse. The Bureau should also study the effects of imputation on the distributions of characteristics and the relationships among them and conduct research on improved imputation methods for use in the American

Community Survey (or the 2010 census if it includes a long-form sample; Recommendation 7.1).

We recommend that the Census Bureau should make users of the 2000 census long-form-sample data products (Summary Files 3 and 4 and the Public Use Microdata Samples) aware of the high missing data rates and measures of inconsistent reporting for many long-form items. The Bureau should also inform users of the need for caution in analyzing and interpreting these data (Recommendation 7.2). Users should note that the Census Bureau has reissued employment status tabulations to provide data for the household population and to exclude group quarters residents because of high missing data rates and imputation errors for them.

We recommend that the Census Bureau's planning for the 2010 census should include research on the trade-offs in costs and accuracy between imputation and additional field work for missing data (Recommendation 4.2). We also recommend that the Census Bureau publish distributions of characteristics and item imputation rates for the 2010 census and the American Community Survey (when it includes group quarters residents) that distinguish household residents from the group quarters population (at least the institutionalized component). Such separation would make it easier for data users to compare census and ACS estimates with household surveys and would facilitate comparative assessments of data quality for these two populations (Recommendation 7.3).

ASSESSMENT OF 2000 CENSUS RESEARCH

The Census Bureau planned two major evaluation efforts for 2000: one set of evaluations addressed population coverage, using results from the census, the A.C.E., and demographic analysis; the other set assessed the full range of census processes and examined measures of completeness and consistency of reporting for content items. Under tight time constraints, the Census Bureau staff charged with evaluating coverage conducted comprehensive and insightful research of high quality. Their results for the A.C.E. (and demographic analysis) were timely, well-documented, and provided useful information to stakeholders and the public. However, the evaluation program for census operations and data quality appears to have experienced difficulties in obtaining usable data for analysis, perhaps because the evaluation staff perspective was not taken into account in designing census databases, such as the MAF. The results from this evaluation program were slow to appear and are often of limited use for understanding the quality of the 2000 census or for planning the 2010 census.

In addition to pursuing improvements for coverage evaluation in 2010 as outlined above, the Census Bureau must materially strengthen the evaluation component for census operations and data quality in 2010 (and in the current testing program). The Bureau should mine data sources created during the 2000 census process, such as the A.C.E. data, extracts from the MAF, a Master Trace Sample of data collection and processing information for a sample of addresses, and a forthcoming match of census records with the March 2000 Current Population Survey, to address important outstanding questions about 2000 data quality and suggest research to improve the 2010 census and the American Community Survey. The Bureau should identify important areas for evaluations in 2010 to meet the needs of users and census planners and set evaluation schedules accordingly; relatedly, it should design and document 2010 data collection and processing systems so that information can be readily extracted to support timely, useful evaluation studies. The Bureau should use graphical and other exploratory data analysis tools to identify patterns (e.g., mail return rates, imputation rates) for geographic areas and population groups that may suggest reasons for variations in data quality and ways to improve quality (such tools could also be useful in real-time management of census operations). The Bureau should accord priority to development of technical staff resources for research and evaluation of both coverage and content in 2010 (Recommendations 9.1 and 9.2).

Finally, the Census Bureau should share preliminary analyses with outside researchers for critical assessment and feedback. To help the Bureau evaluate population coverage and data quality in the 2010 census, the Bureau should seek ways—using the experience with the Panel to Review the 2000 Census as a model—to furnish preliminary data, including microdata, to qualified researchers under arrangements that protect confidentiality (Recommendation 9.3). The panel benefited greatly from the ability to have ready access to detailed evaluation data. In our view, greater sharing with researchers of preliminary research data and results would strengthen the Bureau's evaluation programs and provide additional knowledge to the user community about the strengths and weaknesses of census data.

CHAPTER 1

Introduction and Overview

CONDUCTING A POPULATION CENSUS of the United States is a task of awesome scope and difficulty. It requires massive effort—in a short amount of time—to enumerate and collect data on over 280 million residents at their precise geographic location, striving always to count every resident once and only once. In order to meet this challenge, the decennial census-taking process has evolved into an extremely complex set of operations and components; each new census includes some techniques that have been integral parts of the process for decades and others that are introduced as the newest and possibly best means of improving the count.

The decennial census is formidable in its operations, and the importance of the information generated by the census is similarly impressive. Census data today serve not only the constitutionally mandated purpose of reapportioning seats in the U.S. House of Representatives, but also such purposes as redrawing legislative district boundaries, allocating federal and state program funds, planning and evaluating government programs, informing private- and public-sector decisions, and supporting a wide range of research. Census data not only provide the means of periodically recalibrating the political mechanisms of democracy, but also contribute to an informed citizenry; these data inform people of the civic health of the nation as a whole and allow them to see how their state, city, town, and neighborhood compare with other areas and how they have changed over time.

Due to the scope of census operations and the demands on census data, evaluating a decennial census is a challenging but essential mission, requiring careful scrutiny of every procedure and careful assessment of the effect of each procedure on the quality of the resulting data. The ultimate benchmark against which the results of a census could be compared—namely, an unambiguously true count of the population—is as unknown and elusive to census evaluators as it is to census collectors.

1–A THE PANEL AND ITS CHARGE

In 1998, the U.S. Census Bureau asked the National Research Council's Committee on National Statistics (CNSTAT) to convene a Panel to Review the 2000 Census in order to provide an independent assessment of the 2000 census. A companion CNSTAT Panel on Research on Future Census Methods was convened in 1999 to observe the 2000 census and its evaluation process in order to assess the Bureau's plans for the 2010 census (see National Research Council, 2000a, 2001c, 2003a, 2004).

The charge to our panel is broad:

> to review the statistical methods of the 2000 census, particularly the use of the Accuracy and Coverage Evaluation Program and dual-systems estimation, and other census procedures that may affect the completeness and quality of the data. Features the panel may review include the Master Address File, follow-up for nonresponse, race and ethnicity classifications, mail return rates, quality of long-form data, and other areas.

We conducted a variety of activities to carry out our charge: making observation visits to census offices during 2000; convening three open workshops on issues of coverage evaluation and adjustment; commissioning jointly with the Panel on Research on Future Census Methods a group of local government representatives to evaluate the Local Update of Census Addresses (LUCA) Program; commissioning a paper on race and ethnicity reporting in the census (Harris, 2003); reviewing voluminous evaluation reports and other documents that are publicly available from the Census Bureau; reviewing reports of other groups monitoring the census; and conducting original analysis with aggregate and microdata files provided to the panel by the Bureau under special access arrangements.

Prior to this, our final report, we published three workshop proceedings (National Research Council, 2001e,f,g), the report of the LUCA Working Group (Working Group on LUCA, 2001), three letter reports to the Census Bureau (National Research Council, 1999a, 2000b, 2001d), and an interim assessment of 2000 census operations (National Research Council, 2001a). Appendix A contains a summary of our activities and reports.

1–B OVERVIEW OF THIS REPORT

Our report contains 10 chapters and nine appendixes. The remainder of this first chapter describes sources of error in a census and methods of evaluating the completeness of census coverage and the quality of census data (1–C) and presents 11 summary findings about the 2000 census (1–D).

In our view, it is crucially important that census information be interpreted and evaluated in the context of the goals and uses for the data. Chapter 2 describes major census goals and uses, including the use of census data in congressional reapportionment and redistricting as well as the broader uses of the basic demographic data and the additional data on the long-form sample.

We have divided our analysis of the census as a whole into two major pieces. One is general census operations, which progress from the development of an address list through the completion of census evaluations. The second is coverage evaluation—attempts to assess undercount or overcount in the census and the debate over statistical adjustment of census figures. We present each of these major pieces in two-chapter segments that are parallel in structure. Chapter 3 describes the road to 2000—a summary of the planning cycle of the 2000 census and the complex political and operational environment that surrounded it. Chapter 4 presents our assessment of overall census operations, including the contributions of major procedural innovations to mail response, timeliness, cost, and completeness of response, as well as the enumeration of group quarters residents. Similarly, Chapter 5 introduces general concepts in coverage measurement and statistical adjustment and the evolution of plans for 2000 census coverage evaluation over the 1990s, and Chapter 6 presents our assessment of coverage evaluation in 2000.

In Chapters 7 and 8, we focus on the quality of the census data content. Chapter 7 reviews the quality of the 2000 census basic demographic data and additional long-form-sample data, relative to alternative sources and to past censuses. Chapter 8 addresses the measurement of race and ethnicity (Hispanic origin) in 2000. In Chapter 9, we offer brief comments of a general nature on the organization and management of census operations and the nature of an appropriate research program for 2010. The panel's detailed findings and recommendations are collected and restated for the reader's convenience in Chapter 10.

Appendixes provide information on nine topics: the panel's activities and prior reports (A); basic questionnaire items and additional long-form items in the 2000 census, compared with questionnaire content in the 1990 census and the Census 2000 Supplementary Survey (B); 2000 census operations and major differences from 1990 census operations (C); quality of mail and enumerator returns (D); operations of the 2000 Accuracy and Coverage Evaluation Program (E); the theory and statistical basis of imputation for nonresponse (F); basic (complete-count) data processing (G); long-form-sample data processing (H); and topics covered in experiments and evaluations by the Census Bureau of 2000 census operations and data content (I). A Glossary defines technical terms and acronyms, and a Bibliography provides citations for references in the text and other pertinent references.

1–C EVALUATING A CENSUS

Our focus in this report is the quality of the 2000 census results—not only the completeness of population coverage, which was the primary concern of the U.S. Congress and others leading up to the census, but also the accuracy of the data on population characteristics. Such data include the basic items collected for all persons (about 281 million people) and additional items asked on the long form of about a one-sixth sample of the population (about 45 million people). Where possible, we consider the quality of census results not only at the national level, but also for population groups and subnational areas. We consider how aspects of the design and operation of the census may have affected data quality in 2000 and suggest priority areas for research to improve procedures for the 2010 census. Cost and management issues are also important for a full assessment, and we consider some aspects of them, but our review is not systematic because our charge did not include a management review. More generally, we evaluate some aspects of the 2000 census more fully than others, depending on the importance of the topic, the availability of relevant data, and the expertise of the panel.

In this section, we briefly describe our general approach to the task of evaluating the 2000 census and describe some general principles regarding sources of errors and types of evaluation.

1–C.1 Errors in the Census

At the most basic level, an ideal census evaluation would measure the differences between census-based counts or estimates and their associated true values. An estimated count greater than the true value would be considered a net overcount, and an estimated count less than the truth would be a net undercount. These differences between estimates and (unknown) truth are *errors*, in the statistical sense. Despite the word's colloquial meaning, these errors are not necessarily an indication that a mistake has been made.

Another measure of error would be the sum of the positive deviations and the sum of the negative deviations from the true values for a population group or geographic area. This measure is the gross coverage error, comprising gross overcount and gross undercount, each of which comprises several components. For example, one type of gross error occurs for geographic areas when households and people are assigned to an incorrect location (geocoding error); another type of gross error is duplicate census enumerations. There is not complete agreement on all of the components of gross error—for example, whether to count as errors those people whose status as a correct or erroneous enumeration could not be determined. Furthermore, the definition of gross errors depends on the level of geographic aggregation—in particular, some geocoding errors affect only individual census blocks or tracts but not larger areas. Nonetheless, gross coverage errors are important to examine because they may indicate problems that are not obvious from a measure of net error. For example, net coverage error could be zero for the total population, but large errors of omission could just balance erroneous enumerations. Moreover, to the degree that these two types of coverage error differ among population groups and geographic areas, there could be different net undercounts for population groups and geographic areas even when total net error is zero. Even when gross omissions and erroneous enumerations happen to balance, examination of their components can help identify sources of error that would be useful to address by changing enumeration procedures or other aspects of the census.

Errors in the census can generally be categorized as involving one of two broad types of problems. First, they may result from problems of coverage—that is, some addresses fail to appear once and only once in the Census Bureau's address list, or some individuals fail to be included once and only once in the enumeration. Second, errors may arise due to problems of response, in that responses on a collected questionnaire may be incomplete or inaccurate. Other types of error occur in the census as well. For example, the census long form is distributed to only a sample (17 percent) of the population, and so estimates based on the long form—like all survey results—are subject to sampling error.

Examples of potential errors in coverage are numerous. One natural source of such errors is the list of mailing addresses used by the Census Bureau to deliver census questionnaires, known as the Master Address File (MAF) in the 2000 census. We will discuss the MAF often and in detail later in this report, but for the moment we consider it in the abstract. The MAF was constructed in order to be as thorough as possible, but the dynamic nature of the U.S. population and its living arrangements make it all but impossible for the list to be completely accurate. It is difficult to fully capture additions and deletions to the list that result from construction of new residences, demolition of former residences, and restructuring of existing residences. Identifying all residences in remote rural areas and in multiunit structures is also a major challenge. In the 2000 census, the MAF was built in part with sources that were not used in previous censuses, such as the Delivery Sequence File used by the U.S. Postal Service to coordinate its mail carriers and the direct addition of new addresses from local and tribal governments. Again, the intent was to make the MAF as complete as possible and improve coverage; however, these sources are not guaranteed to be complete, and they may have inadvertently added duplicate addresses to the final address list. All of these complications—representing possible gaps or overages in the address list—may result in either undercounts or overcounts.

Errors in response are similarly numerous and impossible to avoid. Many people simply do not fully cooperate in filling out census forms; this is a particular concern for the census long form, parts of which some respondents may believe violate their privacy. Some households and individuals do not respond to all the questions on the questionnaire; there may also be some degree of intentional misresponse. Some error is also introduced—often unintentionally—by members of the massive corps of temporary census field workers assigned to follow up on nonresponding residents. Another source of census response error is confusion over the questionnaire itself; language difficulties may deter some respondents, while others may not understand who should and should not be included, such as students on temporary visas or away at college. Many individuals have more than one home (examples include "snowbirds" from cold weather states with winter homes in Florida or the Southwest and children in joint custody arrangements), while many others are homeless most or all of the time.

Any evaluation of a decennial census must confront a commonsense but nevertheless critical reality: error is an inevitable part of the census, and perfection—the absence of all error—is an unrealistic and unattainable standard for evaluation. The sources of census error are numerous, and many are difficult to control. In this light, census evaluators must be aware of potential sources of error, estimate their potential effects, and develop strategies to more fully measure and (when possible) minimize errors.

1–C.2 Methods of Evaluation

Acknowledging the inevitability of error is one important point in framing an evaluation of a census; so too is realizing that "evaluation" means many things and spans many different approaches. In this section, we briefly describe several general approaches to evaluation; each of these approaches plays some role in the arguments laid out in the remainder of this report.

Process Evaluations

In such a large and complex operation as the decennial census, it is important to assess the conduct of specific procedures as they contribute to the success of the overall effort. Real-time process evaluations (quality control or quality assurance efforts) are needed to determine that a procedure is being executed as originally specified with minimal variation and on the needed time schedule. If schedule delays or more-than-minimal variation (e.g., different application of rules for accepting proxy enumerations across local census offices) are identified, there will hopefully be sufficient time to correct the problems. For example, sound census practice is to reinterview a sample of each enumerator's work in nonresponse follow-up to be sure that he or she understands the job and is carrying it out in a conscientious manner. An interviewer who fails the reinterviewing standards will be replaced and his or her workload reenumerated as needed.

After-the-fact evaluations of individual procedures are important to determine their effects—positive and negative—on the timing and cost of the census and on data quality. After-the-fact process evaluations cannot assess data quality directly, but they can help answer such questions as whether a procedure may have introduced a systematic bias in an estimate (e.g., a systematic overestimate or underestimate of people who reported being of more than one race in follow-up due to inappropriate cues from enumerators). Such evaluations can also indicate whether a procedure may have introduced variation in data quality (e.g., variation in rates of missing long-form information due to variation in local census office instructions to enumerators on how aggressively to push for the information).

After-the-fact process evaluations can suggest important areas for research to improve operations in future censuses.

Quality Indicators

Having effective real-time quality control and assurance procedures in place for key census operations will contribute to but not ensure high-quality data. First, processes may have been well executed but poorly designed to achieve their goal. Furthermore, even when well-designed operations are carried out as planned, respondent errors from such sources as nonreporting, misreporting, and variability in reporting may result in poor-quality data. As signals of possible problems that require further research, it is common to construct and review a variety of census quality indicators, such as mail return rates, household (unit) nonresponse rates, item nonresponse rates, inconsistency in reporting of specific items, and estimates of duplicates in and omissions from the census count. For census long-form-sample estimates, it is also important to look at variability from sampling and imputation for missing data.

Each of these quality measures must itself be evaluated and refined. For example, the initial measure of duplicates and other erroneous enumerations from the Census Bureau's 2000 Accuracy and Coverage Evaluation (A.C.E.) Program—issued in March 2001—turned out to be substantially lower than a revised estimate released in October 2001, which in turn was lower than the final estimate issued in March 2003. As another example, it is generally assumed that a high nonresponse rate to a questionnaire item impairs data quality. The reason is the twofold assumption that reported values are likely to differ from the values that would have been reported by the nonrespondents and that imputations for missing responses will tend to reflect the values for reporters more than the true values for nonrespondents (as well as add variability). While often true, neither part of this assumption is always or necessarily true (see Groves and Couper, 2002).

Quality indicators (and process evaluations) also need a standard for comparison—for example, what is a "high" versus a "low" nonresponse rate? Such standards are commonly drawn from prior censuses; they may also be drawn from other surveys and administrative records. It is important to evaluate the applicability and quality of the source(s) for the standards. Differences in 1990 and 2000 census procedures complicate the task of comparative evaluation, as do differences in survey procedures from census procedures.

Comparisons of Estimates With Other Sources

Comparison of specific estimates from the census with the same estimates from other sources is another important form of data quality evaluation. For evaluation of the completeness of population coverage, the two traditional comparison sources have been demographic analysis and dual-systems estimation from an independent survey matched with a sample of census enumerations. For evaluation of the accuracy of such estimates as median household income, percentage multirace reporting, or the unemployment rate for specific population groups, comparison sources include household surveys, such as the Current Population Survey and the Census 2000 Supplementary Survey, and administrative records, such as Social Security records. The independent A.C.E. survey is also a comparison source for some estimates, such as people who reported more than one race.

Neither the census nor any comparison source can be viewed as the "truth," and it can often be difficult to establish which sources, if any, may be "better" than the census. A careful assessment of the quality of each source is needed to help determine the relative superiority or inferiority of different sources used for comparisons with census estimates. Also, a careful examination of concepts and procedures for various sources and the census is needed to identify important differences that could affect the validity of comparisons of estimates.

Comparisons With Previous Censuses

An almost inevitable standard of comparison for a census is to compare its execution and outcomes—costs, timing, completeness of population coverage, and other criteria—with past censuses. Typically, the design for a census is shaped by perceptions of the major problems for the preceding census. However, differences in procedures and in the social and political context may affect the appropriateness of comparisons with previous censuses, and such differences need to be interpreted carefully.

Explaining Data Quality Strengths and Weaknesses

An assessment of census data quality is important not only for data users but also for the Census Bureau to help plan future censuses. For this purpose, it is important to conduct analyses to determine one or more underlying explanations for quality successes and problems so that changes can address relevant causal factors.

The ideal method for causal analysis is experimentation in which subjects are randomly assigned to treatment and control groups. Such experimentation is of limited use for understanding census data quality. Experiments were conducted in the 2000 census, but they were few in number, small in scale, and narrow in scope.[1] Census planning tests conducted between census years are often experimental in nature, but they, too, are necessarily limited in scope. Moreover, census tests cannot be conducted under the same conditions as an actual census with regard to publicity, scale of operations, and other features.

So, while experimentation can help identify some reasons for the levels of census data quality achieved nationwide and for areas and population groups, the search for underlying explanations must rely primarily on nonexperimental methods, such as multivariate analysis of quality indicators. We used such methods for analyzing mail return rates as a function of neighborhood characteristics and imputation rates as a function of geographic location and type of census enumeration. While these types of analysis will probably not produce definitive conclusions regarding cause and effect, they are important to pursue for clues to alternative census procedures that merit testing for 2010.

1–D SUMMARY OF FINDINGS: OVERALL ASSESSMENT

To convey our overall assessment of the 2000 census, we provide 11 major findings and a summary conclusion. These findings cover the successes of the census in the face of considerable adversity,

[1] See Section I.3 in Appendix I for a description of the formal experiments conducted to accompany the 2000 census.

problem areas in census operations and data quality, decisions about adjustment of census population counts for coverage errors, and the Census Bureau's two major evaluation programs for 2000 (one on coverage, the other on census processes and content).

1–D.1 Achievements

The decennial census is the federal government's largest and most complex peacetime logistical operation. To be most effective and efficient, design decisions, budget decisions, detailed planning, and operational trials must be completed well before the census is conducted. Yet as we relate in Chapters 3 and 5, the design and planning process for the 2000 census was unusually contentious. Partisan criticism of censuses has occurred in the past, but never before had the Census Bureau faced the need to develop several different operational plans for conducting the census, a budgetary compromise that occurred so late in the planning process, and an atmosphere in which the objectivity of the Bureau itself was under strong attack. In these circumstances, and given the enormous complexity and scale of census operations, the 2000 census was a considerable success. It was carried out in a completely objective manner in which decisions were made on the basis of professional judgment using the best available information.

> **Finding 1.1:** The 2000 census was generally well executed, even though the Census Bureau had to contend with externally imposed last-minute changes in design, delayed budget decisions, consequent changes in plans, and insufficient time for operational trials. The dedication of the Census Bureau staff made possible the success of several operational innovations. Because of the many delays and last-minute changes in census design, the ample funding appropriated in spring 1999 was essential for the successful execution of the census.

The Bureau made innovations in the design of census questionnaires and mailing materials and in the scope and content of advertising and outreach programs that boosted overall mail response rates, even as the population had become more resistant to government requests for information and more concerned about privacy issues. The Bureau also succeeded in hiring the massive workforce required for census field operations, completed follow-up of nonresponding households in a timely manner, and made effective use of new technology for data input (see Chapter 4).

> **Finding 1.2:** The decline in mail response rates observed in the 1980 and 1990 censuses was successfully halted in the 2000 census, and most census operations were well executed and completed on or ahead of schedule.

Finally, the Bureau made progress toward one of its major goals, which was to reduce differential net undercount rates between historically less-well-counted groups (minorities, renters, and children) and others compared with 1990 (see Chapter 6). Such reductions generally imply reductions as well in differential net undercount rates among states and other geographic areas. This assessment is based on comparing the results from the 1990 Post-Enumeration Survey (PES) and the 2000 A.C.E. Program as originally estimated. Subsequent revisions of the A.C.E. estimates are difficult to compare with the PES, but they do not contradict this assessment (see Chapter 6).

> **Finding 1.3:** Although significant differences in methods for estimating net undercount in the 1990 Post-Enumeration Survey and the most recent revision of the 2000 Accuracy and Coverage Evaluation Program make it difficult to compare net undercount estimates, there is sufficient evidence to conclude that the 2000 census was successful in reducing the differences in national net undercount rates between historically less-well-counted groups (minorities, renters, children) and others.

1–D.2 Problems

Although the census was successful in achieving its overall goals for mail response, timely execution of operations, and reduction in differential net undercount, it also experienced problems. They included large numbers of duplicate enumerations—because of undetected duplications in the MAF and people being reported by two households—and high rates of missing data for many content items on the long-form questionnaire (see Chapter 7). Rates of missing data were even higher for residents of group quarters (e.g., college students residing in dormitories, prisoners, nursing home residents, and others).

> **Finding 1.4:** The 2000 census experienced four major problems of enumeration: (1) errors in the Master Address File; (2) large numbers of duplicate and other erroneous enumerations in the census; (3) high rates of missing data for many long-form items; and (4) inaccuracies in enumerating residents of group quarters.

1–D.3 Adjustment Decisions

A decision by the U.S. Supreme Court in 1999 prohibited the use of sampling-based methods for generating census counts used to reapportion the U.S. House of Representatives. But the Census Bureau faced three major decision points as to whether it would use the results of the A.C.E. Program to statistically adjust 2000 census data to reflect estimated coverage errors. In March 2001 the Census Bureau weighed whether to adjust census data to be used for legislative redistricting; in October 2001 the question was whether to adjust census data for such uses as allocating federal funds to states and localities; and in March 2003 the issue was whether to use a revised set of adjusted totals as the base for postcensal population estimates. Each of these decisions was preceded by a wave of intensive evaluation studies dissecting the A.C.E. Program in great detail; at each of these decision points, the Census Bureau ultimately elected not to adjust the census data.

> **Finding 1.5:** In March 2001, October 2001, and March 2003, the Census Bureau announced that it would not adjust the 2000 census results for incomplete coverage of some population groups (and overcounting of other groups). In our judgment, all three of the Bureau's decisions are justified, for different reasons. The March and October 2001 decisions are justified given (1) the Bureau's conclusion in March that evaluation studies were not sufficient to determine the quality of the A.C.E. population estimates and (2) its conclusion in October, after

further study, that the A.C.E. population estimates were too high and an adjustment using A.C.E. results as originally calculated would have overstated the population.

We view the Bureau's March 2003 conclusion as justified because the final A.C.E. estimates that it produced from further analysis, while reflecting high-quality, innovative work, had to make use of incomplete data and incorporate assumptions that cannot be well supported.

Although the A.C.E. Program experienced problems, the program and its supporting research provided important insights into the census process. Moreover, the 2000 A.C.E. experience reaffirmed the utility of in-depth assessment of both coverage gaps and duplications.

Finding 1.6: The Accuracy and Coverage Evaluation Program provided invaluable information about the quality of the 2000 census enumeration. The A.C.E. was well executed, although problems in reporting Census Day residence led to underestimation of duplicate census enumerations in the original (March 2001) A.C.E. estimates.

In the panel's assessment, the 2000 census and A.C.E. experiences suggest three basic findings (or conclusions): first, on the need to examine components of gross error, as well as net coverage error; second, on the need for adequate time to evaluate the completeness of population coverage and the quality of data content; and, third, about the root variability of census totals at the finest grained geographic level.

Finding 1.7: The long-standing focus of key stakeholders on net undercount distracts attention from the components of error in the census , which include duplications, other erroneous enumerations, and omissions. Although the most recently released national net undercount estimate for 2000 is a small net overcount of 0.5 percent of the population (1.3 million additional people counted), there were large numbers of gross errors, almost as many as in

1990. These errors, which would not have been detected without the A.C.E., have greatest impact on estimates for population groups and subnational geographic areas.

Finding 1.8: The experience with the 2000 Accuracy and Coverage Evaluation Program and the evaluation of census processes and data content make clear that useful evaluation requires considerable time. In particular, it appears difficult to complete sufficiently comprehensive assessments of population coverage and the quality of basic characteristics by the currently mandated schedule for releasing block-level census data for use in redistricting (which is 12 months after Census Day).

Finding 1.9: Census counts at the block level—whether adjusted or unadjusted—are subject to high levels of error and hence should be used only when aggregated to larger geographic areas.

1–D.4 Evaluation

Our last two major findings concern the vitally important evaluation programs for the 2000 census, which addressed population coverage and census processes and content.

Finding 1.10: Under tight time constraints, the Census Bureau's coverage evaluation technical staff conducted comprehensive and insightful research of high quality on the completeness of coverage in the 2000 census. Their results for the A.C.E. and demographic analysis were well documented and provided useful information to 2010 census planners, stakeholders, and the public.

Finding 1.11: The Census Bureau's evaluations of census processes and the quality of the data content were slow to appear, are often of limited value to users for understanding differences in data quality among population groups and geographic areas, and are often of limited use for 2010 planning.

1–D.5 Summary Assessment

Because the census is a complex, multifaceted operation, our major assessments (and our more detailed assessments in subsequent chapters) are difficult to reduce to one or two sentences. Overall, we conclude that the 2000 census was at least as good as and possibly better than previous censuses in completeness of population coverage and the quality of the basic demographic data. The census also experienced problems that made some of the data—principally some of the long-form-sample information and the data for group quarters residents—no better and in some cases worse than in previous censuses.

The decennial census is the largest and most difficult of all the nation's statistical programs. The fact that our assessment found areas that need improvement should not detract from the success of the 2000 census in providing relevant, timely data for many uses. Yet it should also make clear the importance of evaluating and documenting the quality of census data and operations not only to facilitate the appropriate use of the data but also to target research and development efforts most effectively for future censuses.

CHAPTER 2

Census Goals and Uses

THE RESULT OF A DECENNIAL CENSUS is a collection of data of various types. One such type is basic population counts, which are used in reapportionment and redistricting, as input to federal funding formulas, and in many other ways. The other prominent type of data product from the census characterizes local areas or population groups from tabulations and analyses of the demographic and socioeconomic variables on the census long form.

These census data are provided to users in a wide variety of data products (see Box 2.1). Such products include aggregate or summary tables for geographic areas ranging from the nation as a whole to individual blocks or groups of blocks. Another data product comprises public use microdata sample (PUMS) files, which are subsamples of individual person records from the census long-form sample, which have been carefully reprocessed to minimize the risk of reidentification of respondents. Both summary and microdata products serve federal, state, and local governments, private-sector organizations, researchers, the media, and, ultimately, the public. Collectively, these data are the only available source of detailed information for small geographic areas and population groups. They are also the only source of information about the entire U.S. population, including not only residents of conventional housing units but also residents of special places and group quarters like military bases, dormitories, nursing homes, and prisons.

Box 2.1 Selected 2000 Census Data Products

A. **TABULATIONS OF BASIC (COMPLETE-COUNT) DATA ITEMS** (age, sex, race, ethnicity, relationship, housing tenure)

- ***Census 2000 Redistricting Data Summary File*** (P.L. 94-171 File) Population counts, total and aged 18 and over, by 63 race categories; by Hispanic origin; and by geography (county, place, minor civil division, census tract, block group, block)
- ***Demographic Profile; Congressional District Demographic Profile*** Selected complete-count characteristics, by state, county, place, minor civil division, census tract; congressional district (106th, 108th Congresses)
- ***Summary File 1*** (SF1, 286 tables in all)
  a. Population counts for 63 race categories and Hispanic, not Hispanic—down to the block level (file for each state)
  b. Population counts for many detailed race and Hispanic categories and American Indian and Alaska Native tribes—down to the census tract level (file for each state)
  c. Selected complete-count characteristics—down to the block level (file for each state)
  d. National-level file—tabulations for states, counties, and places; urban-rural tabulations
- ***Summary File 2*** (SF2, 47 tables in all)
  a. Complete-count characteristics iterated for many detailed race and Hispanic categories and American Indian and Alaska Native tribes—down to the census tract level (file for each state)
  b. National-level file—tabulations for states, counties, and places; urban-rural tabulations

B. **TABULATIONS OF LONG-FORM (SAMPLE) DATA ITEMS**

- ***Demographic Profile; Congressional District Demographic Profile*** Demographic, social, economic, and housing characteristics (three separate tables), by state, county, place, minor civil division, census tract; congressional district (106th, 108th Congresses)
- ***Summary File 3*** (SF3, 813 tables in all)
  a. Population counts for ancestry groups—down to the census tract level (file for each state)
  b. Selected sample population and housing characteristics—down to the census tract and block group levels (file for each state)
  c. National-level file—tabulations for states, counties, places
- ***Summary File 4*** (SF4)
  a. Sample population and housing characteristics iterated for many detailed race and Hispanic categories, American Indian and Alaska Native tribes, and ancestry groups—down to the census tract level (file for each state)
  b. National-level file—detailed tabulations for states, counties, places

Box 2.1 (continued)

C. **PUBLIC USE MICRODATA SAMPLE (PUMS) FILES** (all items for households and persons selected from the census long-form sample)

- **1-Percent Sample Files**, containing about 1 million household and 3 million person records
  Geographic identification: state, large areas of 400,000 or more population
- **5-Percent Sample Files**, containing about 5 million household and 15 million person records
  Geographic identification: state, areas of 100,000 or more population

NOTES: All files available on the Internet and CD-ROM/DVD, profiles also available on paper; all files have been processed to protect confidentiality.

SOURCE: http://www.census.gov/population/www/censusdata/c2kproducts.html [12/1/03].

The sheer scope and variety of census data and products complicates the task of evaluating a census. This difficulty is compounded by two basic truths. First, census results can only be meaningfully assessed in the context in which those data are actually used. For example, a census could provide outstanding population count data but subpar characteristics data: this could happen if serious problems occurred with the census long form. The data from such a census would be perfectly adequate for some uses but would fail to satisfy others. Similarly, whether the data were to be used as counts or shares, or the level of geographic aggregation, might lead to different judgments about the quality of the data. For instance, a hypothetical census that—for some reason—did an excellent job of collecting information from people living in Western states but not elsewhere would provide good counts for the population living in the West but overestimate the West's share of the total U.S. population. Furthermore, changes in census processes could improve the precision of counts while hurting the use of the same data to represent changes in counts over time. For example, the change to allow multiple responses to race and ethnicity questions in the 2000 census may make it possible to capture data on more focused demographic groups, but it may complicate inferences about the relative sizes of minority groups relative to past censuses. A comprehensive evaluation of a census must therefore strive to interpret census results in the context of all of their possible uses.

The second basic truth—and major complication for evaluating a census—is that there is no single, dominant use of census data. While the use of census data to reapportion the U.S. House of Representatives is arguably the primary role of the American census due to the constitutional mandate to do so, it is hardly the exclusive or most demanding use of those data. Hence, there is no single, dominant metric against which census results can be compared in order to unequivocably determine how good they are.

As context for our assessments and recommendations in subsequent chapters, this chapter summarizes the goals of the 2000 census and major uses of the data under four headings: (2–A) use for congressional reapportionment; (2–B) use for congressional and state and local redistricting; (2–C) other uses of the basic demographic data—age, sex, race, and ethnicity—collected from everyone enumerated in the census, including use of updated population estimates in federal funding formulae; and (2–D) uses of the additional data that are obtained on a wide range of topics from questions on the census long form.

2–A CONGRESSIONAL APPORTIONMENT

Censuses play a fundamental role in the U.S. political system by providing counts of the population for each state once a decade for reallocation of seats in the U.S. House of Representatives. This primary role for census-taking is mandated in Article 1 of the U.S. Constitution and elaborated in Title 13 of the U.S. Code (specifically, 13 USC §141b), which stipulates that the Census Bureau must provide state-level population counts to the president by 9 months after Census Day (i.e., December 31 of the census year under the current schedule). Within 1 week of the opening of the next session of Congress, the president must report to the clerk of the House of Representatives the apportionment population counts for each state and the number of representatives to which each state is entitled according to the prescribed formula (2 USC §2a).[1] In turn, the clerk of the

[1] The current reapportionment formula, which uses the method of "equal proportions," was written into law at the time of the 1940 census (Anderson, 1988:189). Apportionment is subject to the constitutional constraint that each state have at least one representative in the House.

House must inform each state governor of the number of representatives to which each state is entitled (2 USC §2a). This schedule was legislated by Congress in 1929 so that reapportionment would occur automatically each decade, thereby precluding what happened after the 1920 census: beset by rural concerns over sharp population growth in urban areas—largely immigration-driven—Congress could not agree on a bill to reapportion seats on the basis of the census results (see Magnuson, 2000b).

Concern over the possible effect of census coverage error on congressional reapportionment—and, more specifically, the effect of competing statistical adjustments to try to correct for said error—has fueled debate over census methodology for years. Although the use of sampling-based census estimates for reapportionment is now prohibited pursuant to a U.S. Supreme Court ruling (see Chapter 3), the potential effects of error and adjustment on apportionment are still viable concerns given the prominence of reapportionment as a use of census data. In studies related to the 1990 census, census data that were adjusted to reflect estimated net undercount shifted one or two House seats between states when input into the "method of equal proportions" formula used for reapportionment compared to results using unadjusted counts. The sensitivity of the apportionment formula to small shifts in population counts was cited by then Commerce Secretary Robert Mosbacher in his decision not to adjust the 1990 census.

The major issue with regard to census data for reapportionment is the definition of who is included in the state counts. Historically, major controversies have involved the treatment of three groups: (2–A.1) noncitizens, (2–A.2) Americans overseas, and (2–A.3) people who are not counted in the census itself but who are estimated to be part of the U.S. population through a coverage estimation program, such as the Accuracy and Coverage Evaluation (A.C.E.) Program implemented for 2000.

2–A.1 Treatment of Noncitizens

Since 1790 the census has had a goal to count all U.S. residents, including people who are not citizens of the United States (except for tourists and other temporary visitors from abroad and people who represent a foreign government). With the rise in the numbers of illegal immigrants in the 1970s and 1980s, lawsuits were filed to exclude them from the apportionment counts (*Federation for American Immigration Reform (FAIR) v. Klutznick* for the 1980 census and *Ridge v. Verity* for 1990; see Passel, 2000). The plaintiffs argued that including illegal immigrants unfairly benefited states with larger numbers of them at the expense of states with fewer of them. The defendants countered that the Constitution did not limit the apportionment counts to citizens or potential voters but included the entire population (originally counting slaves as three-fifths of a person). Defendants also argued that it was impractical to identify and exclude illegal immigrants in the census.

Federal district courts decided both the *FAIR* and *Ridge* cases on the legal grounds of standing. Specifically, the courts ruled that the plaintiffs had not demonstrated injury or harm resulting from the inclusion of illegal immigrants in apportionment counts. The U.S. Supreme Court declined to hear appeals in both cases, upholding the lower court rulings. Although the district court opinions offer some commentary on the constitutional intent to provide representation to everyone and not just to potential voters, this language was not the basis for the decisions. Hence, the issue of including illegal immigrants in the apportionment counts has never been fully resolved.

The 2000 census followed historical practice by striving to include all U.S. residents in the apportionment count. Indeed, the results indicated that the census counted many more illegal immigrants than had been estimated to reside in the country, based on survey and administrative records data (see Section 5–D).

2–A.2 Treatment of Americans Overseas

Rules about counting Americans who live overseas have varied from census to census. Typically, when counted at all, Americans living abroad have been included in an "overseas" population category and not included in the population of specific states for apportionment or other purposes.

Two recent exceptions occurred in 1970 and 1990, when federal military and civilian employees who lived abroad and could be assigned to a "home" state from administrative records were included in the state counts for reapportionment, but not in any other data releases. According to McMillen (2000:33):

> In both cases, the Census Bureau yielded to pressure from Congress to change its procedures to forestall the passage of undesirable legislation. In 1970 the Census Bureau sought to avoid being embroiled in the debate over the Vietnam War. In 1990 the Census Bureau did not want to link the counting of Americans overseas with excluding undocumented aliens from the census.

Since 1995, groups representing Americans abroad have lobbied the Census Bureau to enumerate not only federal government employees who live overseas, but also private citizens who live abroad. The Bureau has resisted such pleas on grounds of the difficulties of finding such people and evaluating the completeness of their coverage. It would also be hard to determine how to allocate them to a home state of residence for inclusion in state apportionment counts. The 2000 census did not attempt to count private citizens who lived abroad but, instead, followed the 1990 practice by enumerating federal military and civilian employees (and their dependents), including them in the state apportionment counts (using the home residence listed for them in administrative records), and tabulating them as "Americans overseas" for other purposes. This population totaled 576,000 people in 2000, or about two-tenths of 1 percent of the total U.S. population.

After narrowly losing the 435th and final seat in the U.S. House of Representatives to North Carolina, the state of Utah filed the first of two challenges against the apportionment totals on the grounds that the Census Bureau had not treated Americans living overseas equally in the count. Specifically, the case held that Mormon missionaries stationed abroad should be treated in the same manner as military and other federal employees stationed oversees (see Box 2.2). The case ultimately failed but—in the wake of the Utah suit and in response to congressional directives for planning for the 2010 census—the Census Bureau in 2004 will test the feasibility of counting private citizens living abroad, using France, Kuwait, and Mexico as test sites. The Census Bureau also plans a second Overseas Census Test in 2006.

Box 2.2 *Utah v. Evans:* Legal Challenges to 2000 Census

Under the reapportionment counts issued by the Census Bureau in December 2000, Utah fell short of the 435th (and final) seat in the U.S. House of Representatives by fewer than 1,000 people; the seat was awarded to North Carolina. As a consequence, Utah filed two legal challenges, one of which was decided by the U.S. Supreme Court and both of which raise issues likely to be revisited as the 2010 census approaches.

Many of the documents filed by the state of Utah in the two cases remain available on the Internet at http://attorneygeneral.utah.gov/highprofileissues.htm [12/1/03].

Utah v. Evans I: Overseas Enumeration

In its first legal filing, Utah challenged the reapportionment counts on the basis that Americans living overseas were treated unequally in the census count. Specifically, the case cited the failure to count missionaries of the Church of Jesus Christ of Latter-day Saints (who temporarily live abroad), even though military personnel and other federal employees stationed overseas are counted. To apply a uniform standard, the case argued that federal employees stationed overseas should be dropped from the apportionment counts.

In April 2001, a three-judge panel of the U.S. District Court for the District of Utah ruled against the state. In a November 26, 2001, order, the U.S. Supreme Court affirmed the ruling without comment and declined to take the case on appeal (Nov. 26, 2001, order list, No. 01-283).

Utah v. Evans II: Whole-Person Imputation

As litigation in the overseas enumeration case proceeded, Utah's legal team developed another challenge to the apportionment count, this time examining imputation of whole persons to the census when even household size is not known (imputation types 3–5, as in Box 4.2 in Chapter 4). Utah argued that whole-person imputation using hot-deck methods (see Appendix G) constituted sampling, which was prohibited for use in apportionment totals by the U.S. Supreme Court in 1999. If those imputations were dropped from apportionment tallies, Utah would win the 435th seat rather than North Carolina, but no other state's apportionment total would be affected.

Utah initially sought to add its case against imputation to the existing litigation on overseas enumeration, but the U.S. District Court for the District of Utah denied amendment to the case. However, it did so not on the merits of the case but rather on the grounds that it would fundamentally affect the character of the lawsuit. Consequently, the imputation challenge was filed in its own right.

Box 2.2 (continued)

In November 2001, a three-judge panel of the U.S. District Court for the District of Utah sided with the Census Bureau and ruled that imputation did not constitute sampling. Utah appealed to the U.S. Supreme Court, which announced in late January 2002 that it would hear arguments on March 27, 2002, on both the court's jurisdiction in the case (i.e., whether Utah had legal standing to bring the case) and on the merits of the case.

On June 20, 2002, Justice Stephen Breyer delivered the opinion of a 5–4 court majority in favor of the Census Bureau. The court found that "imputation differs from sampling in respect to the nature of the enterprise, the methodology used, and the immediate objective sought. [These] differences are of both kind and degree." Consistent with the "actual enumeration" clause in the Constitution, the opinion notes that "in this instance, where all efforts have been made to reach every household, where the methods used consist not of statistical sampling but of inference, where that inference involves a tiny percent of the population, where the alternative is to make a far less accurate assessment of the population, and where consequently manipulation of the method is highly unlikely, [methodological limits under the "actual enumeration" clause] are not exceeded" (*Utah v. Evans*, 536 U.S. 452, 2002).

Justice Sandra Day O'Connor dissented, concluding that imputation is a form of sampling and is, therefore, prohibited under Title 13 of the U.S. Code (see Box 3.1). Accordingly, she declined to comment on the constitutionality of imputation. Justice Clarence Thomas, joined by Justice Anthony Kennedy, also dissented; they agreed with the majority opinion that imputation is not a form of sampling but disagreed on the question of constitutionality, finding imputation and "estimation techniques" inconsistent with their reading of the Constitution's "actual Enumeration" clause and the historical record of the debate surrounding it. Justice Antonin Scalia dissented in full, arguing that Utah lacked standing to bring the case. Even if Utah succeeded in its challenge and the census totals were recalculated, the president could not and would not be compelled to issue a new apportionment slate to Congress. As a result, Scalia concluded, the federal courts are powerless to "redress the injury" claimed by Utah, and the state is ineligible to bring the case. Scalia would dismiss the case on that basis and did not comment on either the legality or constitutionality of imputation in the census process.

2–A.3 Treatment of Uncounted People

The goal of the census is to count each U.S. resident once and only once. However, research on census coverage, which began with analyses of undercounts of draft-age men and young children in the 1940 census, estimated a net undercount of the population in every census from 1950 to 1990. (The net undercount is the difference between the number of missed people who should have been included in the census—omissions—and the number of people who were included in the census in error, such as duplicates and other erroneous enumerations.) Research also estimated higher net undercount rates for some population groups than others, such as higher rates for blacks compared with nonblacks and children compared with adults (see Chapter 5).

Beginning with the 1970 census, the Census Bureau made special efforts to improve coverage of hard-to-count population groups, although research showed that such efforts were only partly effective (see Citro, 2000c). Beginning with the 1980 census, the Bureau worked to develop a dual-systems estimation (DSE) methodology, based on data from a postenumeration survey and a sample of census records, that could be used to statistically adjust census counts for measured net undercount. The Bureau originally planned to use DSE methods to adjust 2000 census state population totals for congressional reapportionment, but a January 1999 decision by the U.S. Supreme Court precluded the use of adjusted totals for this purpose. The Bureau then planned to adjust census counts for other purposes, but that was not done.

The latest Bureau DSE estimates are that the 2000 census *overcounted* the population by about 1.3 million people or about one-half of 1 percent of the population. This small overcount masked not only large numbers of duplicates and other erroneous enumerations, but also large numbers of omissions. In Chapters 5 and 6 we review the long controversy over census adjustment and assess what is known about population coverage in 2000.

On a related note, the second lawsuit filed by the state of Utah challenged the inclusion of certain kinds of census imputations—persons who were not directly enumerated but whose characteristics were imputed instead—in apportionment totals. In June 2002, the U.S. Supreme Court ruled in favor of the Census Bureau's inclusion of imputations (see Box 2.2; we discuss imputation in greater detail in Section 4–D and in Appendixes F, G, and H).

2–B LEGISLATIVE REDISTRICTING

A second major function of the U.S. census, which flows from its constitutional mandate to provide state apportionment counts, is to provide small-area population counts for redrawing congressional and state legislative district boundaries. Census small-area data by age, race, and ethnicity (Hispanic origin) are used to ensure that districts satisfy standards for equal population and the requirements of the Voting Rights Act for equitable representation of population groups. In this section, we review the history of redistricting with regard to population size for three time periods: through 1960, from 1960 to 1990, and since 1990. Although the historical discussion may seem to be a lengthy digression, the various battles over standards for redistricting are crucial to understanding data needs served by the census and have important ramifications for the level of geographic detail needed in census tallies.

It is important to note that, in the following discussion of equality of district size, we speak principally about comparison of district populations *within* states and not equality of population *across* states, among all 435 districts of the U.S. House of Representatives. The latter issue was resolved by the U.S. Supreme Court following the 1990 census; the state of Montana challenged the "equal proportions" method used to allocate seats in the House after it lost its second House seat and Montana's at-large district became the most populous in the nation. The Court ruled that Congress has discretion to allocate House seats as it sees fit so long as good-faith efforts are made to achieve mathematical equality in representation based on state population, even though inequities in district population (necessarily) arise when compared across states (*United States Department of Commerce v. Montana*, 503 U.S. 442, 1992).[2]

[2]Following the 2000 census, Montana remained a single-district state; with a 2000 census apportionment population of 905,316, its single at-large district remains the most populous in the nation.

2–B.1 History of Redistricting Standards

Through 1960

The needs for census data for redistricting have evolved over time. In the 19th century, Congress typically passed a statute at the occasion of each census that required all states, whether or not they gained or lost seats, to redistrict and to establish single-member districts that were contiguous, compact, and as nearly equal in population as practicable (Durbin and Whitaker, 1991:4–8). During this period, the smallest areas for which census counts were published were incorporated towns and villages, which could be quite small in population. Until 1911, with one exception, Congress increased the size of the House of Representatives at each reapportionment, so that no state lost congressional seats. (The exception occurred when the size of the House was decreased from 240 to 223 members following the 1840 census.) In 1911, the size of the House was fixed at 435 seats, at which level it has remained except for a brief two-seat enlargement when Alaska and Hawaii achieved statehood in 1959.

In 1929, Congress required automatic reapportionment of the House of Representatives upon delivery of new census counts, but it set no standards for redistricting. The U.S. Supreme Court held that the omission of such standards was intentional (*Wood v. Broom*, 287 U.S. 1, 1932) and that it was up to the states to develop their own standards until and unless the courts decided to intervene (Durbin and Whitaker, 1991:4–5; see also Anderson, 1988:Ch.6).

From the 1920s through the 1950s, the courts generally declined to intervene in redistricting issues, and congressional and state legislative districts became increasingly unequal in population size. Many states chose not to redistrict after a census unless they gained or lost seats, and those that did often paid little attention to achieving population equality across districts.

1960 to 1990

The landmark "one-person, one-vote" Supreme Court decisions, beginning in the early 1960s, drastically changed the requirements for redistricting. In the first of these cases, *Baker v. Carr* (369 U.S. 186, 1962, which involved Tennessee state legislative districts), the Court held that reapportionment and redistricting matters were subject to judicial review under the equal protection clause of the Fourteenth

Amendment. In *Wesberry v. Sanders* (376 U.S. 1, 1964), the Court held, under Article 1 of the Constitution, that congressional districts must be as nearly equal in population as practicable. In *Karcher v. Daggett* (462 U.S. 725, 1983), by a 5–4 margin, the Court rejected a New Jersey congressional redistricting plan in which the smallest district was only seven-tenths of one percent smaller in population than the largest district. The New Jersey legislature argued that the 0.7 percent difference (about 3,700 people) was the functional equivalent of zero because it was less than the predictable undercount in the census. The majority decision rebutted that argument: "Even assuming that the extent to which the census system systematically undercounts actual population can be precisely determined, it would not be relevant. The census count provides the only reliable—albeit less than perfect—indication of the districts' 'real' relative population levels, and furnishes the only basis for good-faith attempts to achieve population equality" (Parker, 1989:61; see also Durbin and Whitaker, 1991:12; Ehrenhalt, 1983:56–57).

In *Reynolds v. Sims* (377 U.S. 533, 1964), the Supreme Court held that, under the Fourteenth Amendment, both houses of a state legislature must be apportioned on a population basis. Moreover, states should strive for population equality. Generally, however, the courts allowed more deviation among state legislative seats than among congressional districts—deviations below 10 percent in the size of state districts were accepted, and sometimes deviations between 10 and 16 percent, but not deviations greater than 16 percent (Parker, 1989:57–58; see also O'Rourke, 1980:22).

The courts held that the states could use other data sources than the census for redistricting purposes. Over time, however, the states—on their own initiative and prodded by the courts—came to rely almost exclusively on census data to prepare redistricting plans. Thus, when states used other data for redistricting, such as rolls of registered voters, they generally had to obtain census data to demonstrate to the courts that their data would not give a substantially different result from census data (see National Research Council, 1995b:246–247).[3]

---

[3]Massachusetts until 1990 had a requirement in its constitution to conduct a state census every 10 years in years ending in 5; the state census results were used for local redistricting. Beginning in 1992 results from the U.S. census were used.

From the 1970 census the states could obtain population counts for geographic areas as small as city blocks (which were then defined in urbanized areas and in other localities that contracted with the Census Bureau) and for enumeration districts in unblocked areas. However, no special data files or reports were provided specifically to meet redistricting needs, and the boundaries of blocks and enumeration districts often did not match state voting precinct boundaries.

In 1975 Congress required the Census Bureau to provide census population tabulations to state officials for purposes of redistricting within a year after the census date (i.e., under the current schedule, by April 1 of the census year plus one) (Public Law 94-171; codified in 13 USC §141c). States could suggest block boundaries and specify the geographic areas for which they required tabulations, provided that their requirements satisfied Census Bureau criteria and were transmitted to the Bureau on the specified time schedule; if no special areas were identified, the Census Bureau was to provide "basic tabulations of population." In practice, basic tabulations came to mean tabulations for blocks (about 10 million in all), which were identified nationwide beginning with the 1990 census and are the smallest area of geography identified in census data products.

Since 1990

Court cases in the 1990s continued to uphold strict standards of population equality for congressional redistricting, in no instance approving a plan with more than a 0.09 percent deviation in population size from the largest to the smallest district and generally approving plans with 0.01 percent or smaller deviations. Relevant cases are listed in Box 2.3. One commentator (Baker, 1986:275–276) claimed that a majority of the Supreme Court no longer truly supported the ideal of strict mathematical equality for congressional districts but has felt constrained by precedent. He argued that Congress should pass legislation that would permit a reasonable degree of population variance among districts and require other desirable criteria, such as compactness and contiguity. To date, however, the courts have observed precedent on population equality for congressional redistricting, and Congress has not intervened.

Following the 2000 census, courts continued to uphold strict population equality standards for congressional districts. However, they did not reject plans simply for having slightly more variance than alternative plans. Relevant cases are listed in Box 2.4 (note that not all redistricting challenges have played out by the time of this writing).

As in previous decades, courts in the 1990s and 2000s allowed greater population deviation in drawing state legislative districts. Generally, as before, deviations of less than 10 percent did not have to be justified; deviations of between 10 and 16 percent had to have justification to be upheld, and deviations of more than 16 percent were struck down. Relevant cases in the 1990s are listed in Box 2.5. Relevant cases following the 2000 census are listed in Box 2.6.

2–B.2 Voting Rights Act of 1965 and Amendments

The above discussion of data needs for redistricting has focused on total population figures. The civil rights movement of the 1950s and 1960s led to legislation, court decisions, and administrative practices that moved another requirement front and center—namely, the need for data on race and ethnic origin for purposes of legislative redistricting. The Voting Rights Act, originally passed in 1965 (P.L. 89-110) and extended and amended in 1970, 1975, 1982, and 1992, is the primary piece of legislation in this regard (see Laney, 1992, for a history of the act). The act nowhere stipulates the use of census data, but interpretations of the act by the courts and the Justice Department have virtually mandated the use of census data on race and ethnicity for redistricting.

The original intention of the Voting Rights Act was to make it possible for blacks in the South to participate in elections, an opportunity that was often denied them by unreasonable literacy tests and other barriers to registration and voting. The 1965 act prohibited (in §2) under the authority of the Fifteenth Amendment the enactment of any election law to deny or abridge voting rights on account of race or color. It further specified (in §4) that any state or county that had any test or device as a condition for voter registration on November 1, 1964, and in which the number of registered or actual voters fell below 50 percent of the total voting-age population in the 1964 presidential election could not use a literacy test or any other

Box 2.3 Congressional Redistricting Cases on Population Equality, 1990s

- *Hastert v. Board of Elections* (1991) A three-judge panel of the Northern Illinois federal district court approved the Republican congressional redistricting plan partly on the basis that it had a more precisely equal distribution of population (two districts had 1 more person than the ideal of 571,530 people, or 0.0002 percent deviation) compared with the Democratic plan (deviation of 17 people, or 0.003 percent).
- *State ex rel. Stephan v. Graves* (1992) The Kansas federal district court ruled that a congressional district plan with a population deviation of 0.94 percent was unconstitutional because it failed to achieve population equality, citing *Karcher v. Daggett* (1983). The court rejected maintenance of whole counties in each congressional district as justification for the higher population deviation; it approved an alternative plan with an overall deviation of 0.01 percent (69 people).
- *Anne Arundel County Republican Central Committee v. State Administrative Board of Election Laws* (1992) The Maryland federal district court ruled that the General Assembly's congressional redistricting plan was constitutional despite small mathematical population deviations among districts. The overall variance of the plan was 10 people, or 0.00167 percent. The justifications offered by the state (keeping three major regions intact, creating a minority voting district, and recognizing incumbent representation with its attendant seniority in the House of Representatives) were found sufficient to meet the tests under *Karcher*.
- *Puerto Rican Legal Defense and Education Fund v. Gantt* (1992) The New York Eastern District federal court held that a special master's plan met the criteria of a population deviation of less than 0.001 percent (2 people separated the largest and smallest district, or less than 0.0004 percent).
- *Mellow v. Mitchell* (1992) The Pennsylvania Supreme court upheld a plan selected by a court master that had a total variance of 0.0111 percent of the population (57 people), higher than the variance in some other plans. The court ruled that the deviation was "fully justified by the policy of preserving the boundaries of municipalities and precincts;" its decision was upheld in federal district court.
- *Stone v. Hechler* (1992) A three-judge federal district court held that the legislature's plans for West Virginia's three congressional districts were acceptable, even though the population deviation was 556 people, or 0.09 percent, because the plan better preserved the cores of prior districts and made the districts more compact than other plans.

SOURCE: Web site for the Minnesota State Senate: http://www.senate.leg.state.mn.us/departments/scr/redist/redsum [9/1/03].

Box 2.4 Congressional Redistricting Cases on Population Equality, 2000s

- *Graham v. Thornburgh* (2002) A federal 3-judge panel upheld the Illinois legislature's plan, which had a variance of 33 people (0.0049 percent). An alternative plan would have reduced the variance to 29 people, but the court found that the legislature's plan had sufficient justification for a marginally higher variance.
- *Jepsen v. Vigil-Giron* (2002) A New Mexico state district court upheld a plan with a variance of 166 people (0.027 percent) because it made the least change to the previous districts.
- *Vieth v. Commonwealth* (2002) A federal district court rejected the congressional redistricting plan of the Pennsylvania General Assembly because it had a deviation of 19 people (0.0039 percent), split voting precincts, and other problems, and because an alternate plan was available that had a deviation of only 1 person and split no voting precincts. The Assembly thereupon devised a plan with a 1-person deviation.

SOURCE: See Box 2.3.

test or device to screen potential voters. Finally, it provided (in §5) that any covered jurisdiction (i.e., any jurisdiction required to drop voting tests under §4) had to submit "any voting qualification or prerequisite to voting, or standard, practice, or procedure with respect to voting" adopted after November 1, 1964, for "preclearance" to the U.S. Department of Justice or the U.S. District Court for the District of Columbia to determine that there was no abridgement of the right to vote on the basis of race or color.

The 1970 amendments to the Voting Rights Act outlawed literacy tests and other devices in all jurisdictions and extended coverage to jurisdictions that had such tests in November 1, 1968, and in which there was less than 50 percent registration or turnout in the 1968 presidential election. The effect of this provision was to cover subdivisions in northern and western as well as southern states.

The 1975 amendments to the act included a major new provision that extended coverage under the act to protect the voting rights of language minorities on the basis of the equal protection clause of the Fourteenth Amendment. These language minorities were defined to be people of Spanish heritage, American Indians, Asian Americans, and Alaska Natives. The preclearance provisions of the act (i.e., the requirement to submit proposed changes in voting procedures to the Justice Department for approval) were applied to any county,

Box 2.5 State Legislative Redistricting Cases on Population Equality, 1990s

- *Sinkfield v. Bennett* (1997) An Alabama Circuit Court upheld a plan in which one legislative district had a deviation of 10.2 percent because it was a whole county and no other district had more than a 5 percent deviation.
- *Fischer v. State of Board of Elections* (1994) The Kentucky Supreme Court affirmed that maintaining the integrity of counties was as important as population equality for legislative districts, so long as the deviation did not exceed 5 percent.
- *Legislative Redistricting Cases* (1993) The Maryland Supreme Court affirmed that Maryland's constitutional language requiring "substantially equal population" did not impose a stricter standard than the 10 percent rule imposed by the Fourteenth Amendment.
- *Hlava v. Nelson* (1992) A Nebraska district court held that the Nebraska legislature's self-imposed statistical guideline allowing no more than a 2 percent deviation in population size of legislative districts was constitutional.
- *Fund for Accurate and Informed Representation, Inc. v. Weprin* (1992) The Northern District of New York federal court ruled that the legislature's plans for assembly districts met the required standard of less than 10 percent deviation (the deviation was 9.43 percent).
- *Voinovich v. Ferguson* (1992) The Ohio Supreme Court in 1992 held that a senate district was constitutional even though it deviated by 6 percent from the ideal population size because the Ohio constitution allows districts to deviate from the ideal by up to 10 percent when the district constitutes an entire county; the deviation must be less than 5 percent otherwise. The redistricting plan was challenged on other grounds; finally, in *Quilter v. Voinovich* (1994), the Ohio Northern District federal court held that a deviation of 13.81 percent for house districts and 10.54 percent for senate districts fell within constitutional limits because of the desire to respect county boundaries.
- *Ater v. Keisling* (1991) The Oregon Supreme Court held that the secretary of state's decision to adopt a plus-or-minus 1 percent deviation standard was not irrational or contrary to the redistricting provisions of the Oregon constitution.
- *Langsdon v. Millsaps* (1993) A three-judge district court panel held that a 13.9 percent variance for Tennessee house districts was not justifiable to protect county boundaries when an alternate plan had less than a 10 percent variance and split fewer counties.
- *Holloway v. Hechler* (1992) A three-judge federal district court held that a redistricting plan for the West Virginia House of Delegates, with a 9.97 percent deviation, did not violate the Fourteenth Amendment.
- *Gorin v. Karpa* (1991) The Wyoming federal district court struck down a legislative plan that had a deviation of 83 percent for house seats and 58 percent for senate seats in order to protect county boundaries, requiring the legislature to come up with a plan with less than 10 percent deviation, which was done.

SOURCE: See Box 2.3.

Box 2.6 State Legislative Redistricting Cases on Population Equality, 2000s

- *In re 2001 Redistricting Cases* (2002) The Alaska Supreme Court rejected the Redistricting Board's plan that had a deviation for 16 house seats of 9.5 percent because the Alaska constitution had been amended in 1998 to make the state standard "as near as practicable" and thus more exacting than the federal standard.
- *Smith v. Idaho Commission on Redistricting* (2001) The Idaho Supreme Court rejected a legislative redistricting plan with a deviation of 10.69 percent, given that the state had offered no evidence that the disparity resulted from the advancement of a rational state policy. In a later case, the court rejected a revised plan with an 11.79 percent deviation because policies to preserve county and neighborhood boundaries were not consistently followed statewide.
- *Burling v. Chandler* (1992) The New Hampshire Supreme Court adopted a legislative redistricting plan with a deviation of 9.26 percent.
- *Allen v. Pataki* (2002) A New York court upheld a state senate redistricting plan with a deviation of 9.78 percent even though previous deviations had been smaller (1.83 percent in 1972, 5.3 percent in 1982, and 4.29 percent in 1992).
- *Deem v. Manchin* (2001) A state court upheld a plan for West Virginia senate districts with a deviation of 10.92 percent because the plan recognized the legislature's goals to maintain county lines as nearly as possible.
- *Arrington v. Elections Board* (2002) A three-judge federal court rejected all 16 legislative district plans for Wisconsin and adopted one of its own, which had an overall deviation of 1.48 percent.

SOURCE: See Box 2.3.

city, or township for which the Census Bureau determined that more than 5 percent of the voting-age citizens were of a single-language minority, election materials had been printed only in English for the November 1972 elections, and less than 50 percent of all voting-age citizens in the jurisdiction had registered or voted in the 1972 presidential election. This provision covered the states of Alaska, Arizona, and Texas and political subdivisions in eight other states.

The 1982 amendments to the act kept the basic provisions intact but made some changes. The amendments extended the preclearance requirements of the act until 2007 but provided that Congress reexamine them in 1997. Another provision stated that the standard of proof for judging an election law to be discriminatory was no longer discriminatory *intent*, but rather discriminatory *result*. As somewhat of a counterbalance, still another provision stated that minorities had no right to proportional representation, but courts could consider the lack of representation as part of the totality of circumstances in cases brought under the Voting Rights Act.

With regard to data needs for redistricting, §5 of the Voting Rights Act led to the practical necessity for census data on race and ethnicity for redistricting. A key case that supported the use of §5 to review many aspects of state and local electoral systems was *Allan v. Board of Education* (1969), in which the Supreme Court held that such changes as moving from single-member to multimember districts were "practices or procedures" that were subject to review under §5 because they had the potential of "diluting" the black vote. The Justice Department quickly moved to instruct legal officers in covered jurisdictions to clear every change in voting procedure. Whereas only 323 voting changes were received by the Department for preclearance between 1965 and 1969, almost 5,000 were submitted between 1969 and 1975 (Thernstrom, 1979:59). Parker (1989:59–63) notes that challenges to redistricting plans on the grounds that they are racially discriminatory can be brought under §2 of the act as well as the more frequently invoked §5. Hence, although the preclearance provisions of §5 currently apply to fewer than half the states (in some cases, just to selected jurisdictions in the state; see Laney, 1992), all states must worry about the racial composition of legislative districts if they are to avoid challenges under the Voting Rights Act.

From 1965 to 1988, the Justice Department objected most often to municipal annexations that diluted the voting power of blacks, Hispanics, or other protected minority voters; it objected next most often to changes from single-member districts to at-large voting. The third most common objection (made 248 times) was to redistricting plans that lessened the effectiveness of minority votes, for example, such schemes as dividing concentrations of minority voters into adjoining majority-white areas or minimizing the number of minority districts by placing minority voters in as few districts as possible.

Many court challenges to redistricting plans in the 1990s and after the 2000 census invoked the Voting Rights Act. Pursuant to the language in the 1982 amendments denying a right to proportional representation for minorities but allowing representation to be considered, the courts issued a variety of opinions. Yet whether the courts upheld greater or lesser minority representation, the issue of race and ethnic origin of voters played an important role in Voting Rights Act cases. Relevant cases following the 2000 census are listed in Box 2.7.

2–B.3 Implications of Redistricting for Census Data Requirements

To meet court mandates for equal population size congressional districts with a typical standard of less than 0.01 percent deviation, states almost necessarily rely on the block-level data in the P.L. 94-171 file. The average size of a congressional district in 2000 was about 646,000 people, so that 0.01 percent was about 65 people, which compares to the average block size of about 30 people. Many states try to preserve voting precinct and even county boundaries when delineating congressional districts, but block data are needed to meet strict population size standards.

For state legislative districts, block data are also sometimes needed to meet the looser standard of 10 percent deviation because of the smaller size of such districts. As of 2000, the median population size of state house districts was about 40,000 people, with a range from about 423,000 (California) to about 3,000 (New Hampshire) (see Table 2.1). A 10 percent deviation for the median state means that the smallest district must have no fewer than 38,000 people and the largest district no more than 42,000 people, for a difference of 4,000 people—close to the size of an average census tract. For New Hampshire, a 10 percent deviation would require the use of block group or block data to obtain equal population size within a 10 percent range from 2,850 to 3,150 people.

The data files provided by the Census Bureau to the states for redistricting carry not only population counts but also counts of voting-age population by race and ethnicity, which most states use in drawing district boundaries (see National Conference of State Legislatures, 1992:8). The 2000 census P.L. 94-171 data provided to the states by April 1, 2001, included an unprecedented array of race and ethnic origin data for the total population and people age 18 and over for states, counties, minor civil divisions, places, voting districts (when specified by the state), census tracts, block groups, and blocks. In all, the file provided 63 race categories by Hispanic

Box 2.7 Voting Rights Act Redistricting Cases, 2000s

- *Page v. Bartels* (2001) The New Jersey federal district court upheld a legislative redistricting plan that reduced the concentration of blacks to less than a majority of the voting age population in three districts and increased the concentration in a fourth district, arguing that the result was likely to increase by one the number of blacks elected to the legislature because of white and Hispanic support for black candidates. The same court also upheld a senate district in which the black voting age population was raised from 3.9 to 35.3 percent. (The challenger had argued that minority incumbents had been protected but not white incumbents.)
- *Stevenson v. Bartlett* (2001) The North Carolina Supreme Court held that creation of minority state senate and house districts to satisfy the Voting Rights Act and devising a plan that did not cause the opportunities for minorities to regress took precedence over the state constitution's requirement to preserve county boundaries to the extent possible.
- *Bone Shirt v. Hazeltine* (2002) A three-judge federal district court prevented South Dakota from implementing its legislative redistricting plan because Native Americans were packed into a single senate district in which they constituted 86 percent of the voting-age population and because the plan had not been precleared with the Justice Department.
- *Balderas v. State* (2002) The Texas Eastern District federal court found that the legislature could not have created additional Latino majority districts without risking retrogression (dilution) in existing Latino majority districts. It declined to find any necessity to create minority "influence districts."
- *West v. Gilmore* (2002) A Virginia circuit court ruled against the legislature's house and senate redistricting plan for reasons that include that black voters were packed into as few districts as possible in order to minimize their political influence. In some districts, blacks were more than 55 percent of the voting-age population, a percentage that in past elections had enabled minority candidates to win by landslide proportions. The Supreme Court of Virginia reversed the decision, upholding the original plan.

SOURCE: See Box 2.3.

and not Hispanic for each geographic area in order to reflect all of the possible race combinations afforded by the new option of checking more than one race on the 2000 questionnaire.

A difficulty is that the accuracy of total population counts for individual blocks is not high, whether the data are from the enumeration or are adjusted for measured net undercount. Errors in the population counts for blocks can occur because of such factors as misgeocoding—that is, assigning addresses to the wrong block—as well as omissions and duplications or other erroneous enumerations. Errors in adjusted data include sampling error and biases in the adjustment procedure. Inaccuracy in block-level counts may also arise from imputation for nonresponse, proxy response, and other measurement issues. Users must be aware that block data are simply building blocks for larger areas for which relative accuracy can be better ensured.

The potential effect of census error on legislative redistricting is particularly hard to assess, given the intensely political nature of the process. The shrewdness of a mapmaker in piecing together blocks into districts arguably has more effect on any perceived bias in the district than do block-level census errors. However, it is certainly possible that high levels of error in the census could have major effects on districts within states. For instance, errors in the census might affect the urban-rural balance within a state, and any resulting district map could dilute the vote of urban residents at the expense of rural residents—or vice versa. Such outcomes would depend on the average size of the districts, the differential undercoverage rates of major population groups, the proportionate distribution among areas of these population groups, and the number of districts with high rates of census undercoverage.

2–C OTHER USES OF BASIC CENSUS DATA

Census data on age, race, ethnicity, and sex, which are asked of the entire population, have many uses, particularly as they form the basis of small-area population estimates that the Census Bureau develops for years following each census. Currently, the Bureau produces estimates of total population by single years of age, sex, race, and Hispanic origin on a monthly basis for the United States and annually for states and counties as of July 1 of each year. The Bureau also produces estimates of the total population every 2 years for incorporated places and minor civil divisions of counties (in states that have such divisions). Recently, the Bureau began producing biennial estimates of total population and children ages 5 through 17 for school districts. The estimates are produced by updating the census year figures with data from such sources as birth and death records (see Citro, 2000e).

Census-derived population estimates serve a variety of needs of federal, state, and local government agencies and academic and

Table 2.1 Number and Approximate Average Population Size of State Senate and House Districts, by State, 2000

State	Senate Districts		House (Assembly) Districts	
	Number	Average Size	Number	Average Size
Alabama	35	127,000	105	42,000
Alaska	20	31,000	40	16,000
Arizona*	30	171,000	60	86,000
Arkansas	35	76,000	100	27,000
California	40	847,000	80	424,000
Colorado	35	123,000	65	66,000
Connecticut	36	95,000	151	23,000
Delaware	21	37,000	41	19,000
Florida	40	400,000	120	133,000
Georgia	56	146,000	180	45,000
Hawaii	25	48,000	51	24,000
Idaho*	35	37,000	70	18,000
Illinois	59	210,000	118	105,000
Indiana	50	122,000	100	61,000
Iowa	50	59,000	100	29,000
Kansas	40	67,000	125	22,000
Kentucky	38	106,000	100	40,000
Louisiana	39	115,000	105	43,000
Maine	35	36,000	151	8,000
Maryland	47	113,000	141	38,000
Massachusetts	40	159,000	160	40,000
Michigan	38	262,000	110	90,000
Minnesota	67	73,000	134	37,000
Mississippi	52	55,000	122	23,000
Missouri	34	165,000	163	34,000
Montana	50	18,000	100	9,000
Nebraska	49	35,000	—	—

Table 2.1 (continued)

	Senate Districts		House (Assembly) Districts	
State	Number	Average Size	Number	Average Size
Nevada	21	95,000	42	48,000
New Hampshire	24	51,000	400	3,000
New Jersey*	40	210,000	80	105,000
New Mexico	42	43,000	70	26,000
New York	61	311,000	150	127,000
North Carolina	50	161,000	120	67,000
North Dakota*	49	13,000	98	7,000
Ohio	33	344,000	99	115,000
Oklahoma	48	72,000	101	34,000
Oregon	30	114,000	60	57,000
Pennsylvania	50	246,000	203	60,000
Rhode Island	50	21,000	100	10,000
South Carolina	46	87,000	124	32,000
South Dakota*	35	22,000	70	11,000
Tennessee	33	172,000	99	57,000
Texas	31	673,000	150	139,000
Utah	29	77,000	75	30,000
Vermont	30	20,000	150	4,000
Virginia	40	177,000	100	71,000
Washington*	49	120,000	98	60,000
West Virginia	34	53,000	100	18,000
Wisconsin	33	163,000	99	54,000
Wyoming	30	16,000	60	8,000

NOTES: For a 10 percent deviation or less, the smallest district must be 95 to 100 percent of the average district size, and the largest district must be 100 to 105 percent of the average district size. Nebraska has a unicameral legislature.

* Voters elect one senator and two assembly members from a set of legislative districts.

SOURCE: Adapted from National Council of State Legislatures: http://www.ncsl.org/programs/legman/elect/cnstprst.htm [9/1/03]. In states with mixed multi-member districts, district population is for single-member districts.

private-sector users. The estimates are used by the National Center for Health Statistics as denominators for important national and subnational statistics, such as birth and death rates (by age, sex, and race) for the United States, states, and local areas. Currently, federal agencies allocate more than $200 billion of federal dollars to states and other areas by formulas, many of which include population estimates as one of the factors in the formula. For example, the Prevention and Treatment of Substance Abuse Block Grants Program allocates funds to states (estimated $1.6 billion obligated in fiscal 2002) with a formula that includes two equally weighted factors: each state's share of the population ages 18 to 24 (double counting urban residents ages 18 to 24) and each state's share of the population ages 25 to 64 (see U.S. General Accounting Office, 1999b; National Research Council, 2003b; see also Section 2–D.1).

Major federal household surveys, such as the Current Population Survey (source of official employment and poverty statistics) and the Survey of Income and Program Participation, use census-based population estimates as survey controls—that is, the survey estimates are adjusted to agree with national population estimates by age, sex, race, and Hispanic origin. Without reweighting to national estimates, the surveys would underestimate many demographic groups because coverage of the population is typically less complete in household surveys than in the decennial census (see National Research Council, 1995a:App.B). Beginning in 1994 the population controls for weighting survey results included an adjustment for the undercount in the 1990 census, although such an adjustment was not part of the official estimates that were publicly available. Population estimates beginning in 2001 are based on the 2000 census results, not adjusted for estimated net undercount or overcount.

At the local level, census-derived population estimates are used by planning agencies for the development, implementation, and evaluation of programs in a host of areas, such as day care, job training, elderly assistance, and transportation planning. Governments compare administrative records counts of service users (e.g., services for the elderly) to census counts for the relevant populations to document need and to evaluate the reach and effectiveness of programs.

The private sector makes extensive use of census-derived population estimates and year-to-year change in business planning (see

Naymark and Hodges, 2000). For example, growth in the number of teenagers may support a decision to launch a new product aimed at that market; the distribution of the Hispanic population may help guide the placement of Spanish-language advertising, signs, and merchandise; and determining when the population of a growing suburb will reach a threshold for opening a new store can help in site planning and development.

Research uses of basic census data include analyses of residential segregation by race and ethnicity among neighborhoods within and across cities and metropolitan areas and over time by comparison with previous censuses. Another research use of basic census data is to analyze changing age composition and the implications for state and local government finance.

The major issues of concern to our panel with regard to the basic census data (which also include household relationship and housing tenure) concern the quality of the estimates. Specific quality issues that we address (see Chapter 7) include response rates for individual items and the effects of imputation procedures to compensate for nonresponse; consistency of reporting for the same people in the census and other sources (an indicator of reliability); and reporting errors (e.g., reporting age as younger or older than actual age) and the net biases from reporting errors (e.g., the extent to which underreports and overreports of age fail to balance out). Given that people attach different meanings to race and ethnicity, consistency and reporting errors may be particularly problematic for those items.

For census-derived population estimates, an added concern is the quality of the administrative records data that are used to update the census figures to account for births, deaths, and net migration. Specific issues include compatibility of reporting of race and ethnicity among the different data sets and the accuracy of estimates of net illegal immigration. For small-area population estimates, an added issue (which we do not discuss) is the accuracy of the data and methods used to estimate migration flows among areas.

2–D USES OF ADDITIONAL DATA FROM THE LONG FORM

Beginning in 1820, when enumerators were asked to tally the number of noncitizens in each household and the number engaged in agriculture, manufacturing, and commerce, the decennial census has been a vehicle to obtain additional information beyond the basic head count by age, race, and sex. By 1910 the census questionnaire had grown to include 34 questions for household residents, including items on age, race, sex, household relationship, marital status, citizenship, education, health, home ownership, language, occupation, place of birth, service in the Civil War, and—in a last-minute addition by the Congress—language of mother and father (Magnuson, 2000a). The 1910 census also included special questions for American Indians, Alaska Natives, and blind and deaf people.

Beginning in 1940 statistical sampling was used to accommodate additional questions without burdening the entire population: six new questions on socioeconomic status and housing adequacy were asked of a 5 percent sample of the population in 1940 (6.5 million people). The 1960 census, which first used the U.S. Postal Service for questionnaire delivery, introduced the concept of separate "short" and "long" forms. Since then, the short form has included basic items asked of all U.S. residents, and the long form has included the short-form items together with additional items asked of a sample (about one-sixth of the population in 2000, or 45 million people). The added questions on the 2000 census long form (see Appendix B) included 36 population items, covering such topics as marital status, educational attainment, place of birth, citizenship, language spoken at home, English proficiency, ancestry, military service, year moved into residence, various types of disability, responsibility for grandchildren in the home (new item in 2000), current and prior-year employment status, occupation and industry, transportation to work, and income by type. The 2000 census long form also included 26 housing items, covering such topics as market value of owned home, rent, cost of utilities, characteristics of house or apartment (e.g., number of bedrooms, heating fuel), year structure built, ownership finances (mortgage payment, taxes, home insurance), and number of vehicles (including autos, vans, and trucks).[4]

The census long-form-sample data have extensive uses in the public and private sectors. At present the long-form sample is generally the only source of nationwide small-area data (for counties, cities, towns, neighborhoods) on the topics it covers, and the size of

[4]Numbers of questions or items should be taken as approximate: one question may have several parts that refer to distinct items.

the long-form sample also makes it a unique resource for analysis of small population groups, such as elderly minorities. Below we briefly describe uses of long-form data by: (2–D.1) the federal government (see Citro, 2000d; National Research Council, 1995b:Apps. C, G, H, M); (2–D.2) state and local governments (see Gaines et al., 2000; National Research Council, 1995b:Apps. E, G, H); (2–D.3) the private sector (see Naymark and Hodges, 2000; National Research Council, 1995b:App. F; Spar, 2000); and (2–D.4) academia (see National Research Council, 1995b:App. D).

2–D.1 Federal Government Uses of Long-Form-Sample Data

Many federal agency uses of census long-form data are mandated in law, either directly or indirectly in that the census is the only feasible data source to satisfy a mandate. Indeed, no item was included in the 2000 census if it was not deemed to serve an important federal purpose (see Section 3–B.2). Such purposes include implementation of sections of the Voting Rights Act, allocation of federal funds to states and localities, assessment of charges of employment discrimination, and planning, monitoring, and evaluation of federal programs.

- Since §203 was added to the Voting Rights Act in 1975, census long-form data have played a key role in ensuring that localities assist voters who have trouble reading English.
  - §203 required counties, cities, and townships to provide election materials and oral assistance in another language as well as in English in areas for which the Census Bureau determined that 5 percent of the voting-age citizens were of a single-language minority and had an illiteracy rate in English (defined as failure to complete fifth grade) higher than the illiteracy rate in English of the entire nation.
  - The 1982 amendments asked the Census Bureau to investigate the usefulness of 1980 census long-form questions on mother tongue and English-speaking ability for determining coverage under the bilingual assistance provision. On the basis of the Bureau's research, the definition of a covered area became one in which 5 percent of the citizens of voting age spoke a single minority language, said

they did not speak English very well, and had a higher illiteracy rate than the nation as a whole. The result of this change was to reduce the number of covered areas from about 500 following the 1970 census to about 200 following the 1980 census; about 300 covered areas were identified after the 1990 census.

  - The 1992 amendments extended the bilingual voting assistance provision to 2007 and made some additional minor changes to the definition (see Bureau of the Census, 1976; Kominski, 1985, 1992). About 300 areas were identified as covered under the bilingual assistance provision after the 2000 census (*Federal Register*, 67[144, July 26, 2002]:48871–48877).

- Over $200 billion of federal funds are allocated each year to states and localities by means of formulas, many of which directly or indirectly make use of census long-form data (see National Research Council, 2001b, 2003b). Examples include:

  - Medicaid (estimated $145 billion obligated in fiscal year 2002): Reimburses a percentage of each state's expenditures for medical care services for low-income elderly and disabled people and families with dependent children by a formula that uses per capita income estimates from the U.S. Bureau of Economic Analysis (BEA). BEA develops these estimates by using data from a wide range of administrative records, the decennial census long-form sample and other censuses and surveys, and census-based population estimates (as denominators). The Medicaid formula is also used to allocate funds under other programs (e.g., Foster Care-Title IV-E, with estimated obligations of $5.1 billion in fiscal 2002).

  - Title 1 of the Elementary and Secondary Education Act (estimated $9.5 billion obligated in fiscal 2002): Allocates funds to school districts to meet the needs of educationally disadvantaged children on the basis of a formula that includes estimates of poor school-age children. The estimates were previously derived from the most recent census long-form sample; currently, the estimates are obtained from statistical models developed by the Census Bureau, which include census poverty data as one input.

– Special Education Grants to States (estimated $7.5 billion obligated in fiscal 2000): Allocates funds to states for the education of handicapped children in part on the basis of a formula that includes the number of children in the age ranges mandated by the state's program and the number of children in poverty in those age ranges. Census long-form-sample data are the source for the estimates.

– Community Development Block Grants, Entitlement Grants Program ($3 billion authorized in fiscal 2002): Allocates 30 percent of funds to states and 70 percent of funds to localities (metropolitan cities and urban counties) on the basis of the larger amount computed under two formulas. Both formulas use census complete-count and long-form-sample data—total population, poverty population, and overcrowded housing units in the first formula and total population, poverty population, and housing units built before 1940 in the second formula.

– Home Investment Partnership Program (estimated $1.8 billion obligated in fiscal 2002): Allocates funds to states, cities, urban counties, and consortia of local governments on the basis of a formula that uses census long-form-sample data, such as estimates of rental units built before 1950 occupied by poor families.

– Workforce Investment Act Adult Program and Youth Activities (estimated $1 billion obligated in fiscal 2002): These programs allocate funds to service delivery areas (one or more counties or cities of 200,000 or more population) by use of formulas that include census long-form-sample data for specified age groups (people ages 22–72 and people ages 16–21, respectively) in families with income not more than the higher of the official poverty line or 70 percent of the Department of Labor lower living standard income level.

- Many federal agencies use census long-form-sample data for program planning, monitoring, evaluation, enforcement, and development of special statistics and surveys. For example:

  - The U.S. Equal Employment Opportunity Commission regularly uses census long-form-sample labor force data on occupation and industry for ZIP codes and other geographic areas to analyze statistical evidence in class action charges of employment discrimination by age, sex, race, or ethnicity. In addition, federal agencies, as employers, use such data to evaluate their own recruitment and promotion systems.
  - The U.S. Department of Transportation uses census long-form-sample data on disability for traffic analysis zones to monitor compliance with the Federal Transit Act and the Americans with Disabilities Act. The Department has helped develop for every census since 1960 a transportation planning package, which provides local governments with detailed information on commuting flows in their areas (cross-tabulating place of residence by place of work for traffic analysis zones made up of groups of census blocks).
  - The U.S. Department of Housing and Urban Development uses census long-form-sample data on rental housing condition to construct fair market rents for nonmetropolitan counties and all but the largest metropolitan areas (for which it uses the American Housing Survey). Fair market rents are used to determine rent subsidies to eligible families under the Section 8 Housing Assistance Payments Program; for areas for which census data are used, fair market rates are set at a percentile of the distribution of gross rents (including utility costs) for two-bedroom apartments with electricity and complete kitchen and plumbing facilities and into which the occupant moved within the last 5 years.
  - Since 1950, the U.S. Office of Management and Budget has used census long-form-sample data on commuting patterns to help define metropolitan statistical areas (MSAs).

In turn, there are many federal agency applications of MSAs, such as determining eligibility for fund allocations.

- Following every census since 1960, the Science Resources Statistics Division of the National Science Foundation has commissioned a survey by the Census Bureau of a sample of college graduates drawn from the census long-form records, with oversampling of people who reported a science or engineering occupation. People identified as scientists and engineers in the survey are followed up every 2 years to track changes in their training and employment.

2–D.2 State and Local Government Uses of Long-Form-Sample Data

State and local governments use small-area census data extensively for such purposes as fund allocation, program planning and evaluation, facility planning, disaster planning, and economic development and marketing.

- Examples of formula allocation by states include: allocation of over $200 million in Colorado child welfare block grants to counties with a formula that includes census estimates of the county's number of children below 200 percent of the poverty line; allocation of $5 million in Florida Intensive Crisis Counseling funds to counties with a formula that includes census estimates of the number of poor female-headed households with dependent children (see Butcher and Dunton, 1999).

- Examples of state and local government program planning and evaluation include: using census long-form-sample data on English language proficiency, language spoken at home, and years of school completed to identify neighborhoods in which residents may require special assistance in using local government services; developing socioeconomic profiles of census tracts to use in applications for funds from federal housing and other grant programs.

- Examples of facility planning include: using census long-form-sample data on commuting patterns to help redesign bus routes and plan new roads; comparing distributions of census small-area characteristics with corresponding distributions of service users (e.g., people visiting health clinics) to determine where new service facilities are most needed (such studies are often done by geocoding administrative records addresses of service users to census geographic areas).

- Examples of disaster planning include: using census long-form-sample data on vehicle ownership and disability together with basic population data to estimate numbers of people who might need to be evacuated in the case of a disaster (e.g., from an area prone to flooding) and how many of them might need transportation assistance; using census place of work and place of residence data to estimate daytime populations for use in developing disaster planning for employment centers.
- Examples of economic development and marketing include: using census long-form-sample data on educational attainment, occupation, and labor force participation of the adult population to attract prospective businesses by informing them of the availability of an appropriately skilled labor force; using census long-form-sample data on ancestry and language to inform businesses of opportunities for serving particular markets.

2–D.3 Private-Sector Uses of Long-Form-Sample Data

Retail establishments and restaurants, banks and other financial institutions, media and advertising firms, insurance companies, utility companies, health care providers, and many other segments of the business world use census long-form-sample data, as do non-profit organizations. An entire industry developed after the 1970 census to repackage census data, use the data to develop lifestyle cluster systems (neighborhood types that correlate with consumer behavior and provide a convenient way of condensing and applying census data), and supply address coding and computer mapping services to make the data more useful. Typical private-sector uses of census data include site location, targeting of advertising and services, workforce development, and assessment of compliance with government requirements, such as equal employment opportunity.

- Many retail and service businesses use census population, poverty, income, and labor force data for specific locations to study the possibilities for retail outlets.

- Some radio stations relate callers' ZIP codes to census tract profiles to determine areas with residents who are partial to the station's format, using the information to develop a marketing plan to target those areas.

- Some churches compare members' characteristics from a survey to census characteristics for the surrounding community to determine ways in which members are similar to and distinct from the community and use the information to develop an outreach program.

- All banks are required to use median household income and income distributions by census tract to ensure compliance with federal mortgage lending guidelines regarding race, as well as for meeting other regulatory requirements.

- Most employers are required to use local-area data on occupation by sex, age, race, and ethnicity to assess compliance with federal equal opportunity employment and anti-discrimination laws.

2–D.4 Research Uses of Long-Form-Sample Data

Research studies using long-form-sample data have developed basic understanding of key social processes, such as migration flows and the social and economic effects of the aging of the population. They have also contributed important information to policy makers and the public on such topics as demographic changes, trends in educational attainment, concentrations of people in poverty, and local exposure to environmental hazards. Some research applications use tabular data for small areas. Others use public use microdata sample (PUMS) files, which are available for most censuses back to 1850—the first census to obtain data about individuals rather than summary data for whole households (see Box 2.1; Ruggles, 2000).

- The aging of the population has led to heightened research attention on the characteristics of older people and how they are changing over time. Census small-area tabulations support analyses of migration flows and concentrations of the elderly by geographic area in total and for subgroups defined by such characteristics as income, living arrangements, labor force attachment, and extent of disabilities. Microdata sample files support detailed analyses of different groups of the elderly, such as those with different disabilities and the relationship of disability to education, labor force attachment, and income. Such analyses are not often possible with other data sources, which have much smaller sample sizes than the microdata sample files. The availability of these files for previous censuses permits rich historical analyses of the aging of the U.S. population and the different experiences of the elderly over time.

- Census data support research on the process of education and the consequences of education on individuals' life chances. Education researchers use small-area census data on median years of school completed, average income, and employment patterns to understand the challenges facing different school districts. They use microdata sample files to study how age and race groups differ on such factors as educational attainment, income, and housing quality.

- Path-breaking analyses of 1970 and 1980 census small-area tabulations revealed a large increase during the 1970s in the number of urban neighborhoods in which more than 40 percent of the families had incomes below the official poverty line. The number of people living in such areas also increased dramatically, and a large proportion of them were blacks living in a handful of the nation's largest cities. These tabulations stimulated research with census and other data on the correlates and consequences of concentrated urban poverty. For example, detailed longitudinal data for families in the ongoing Panel Survey of Income Dynamics were augmented by census characteristics of the neighborhoods in which the sample families lived, permitting rich contextual analyses of family outcomes.

- Census socioeconomic profiles for small areas, when related to environmental data, permit analyses of the distribution of environmental risks for different population groups and are an essential basis for studies of health effects of such local hazards as toxic waste dumps.

- Census small-area data are an unparalleled resource for analysis of migration flows over time among regions, states, counties, and places and the consequences for changing demographic and economic conditions among different parts of the country. Microdata sample files permit detailed analysis of the characteristics of long-distance migrants versus those who migrate shorter distances or not at all.

CHAPTER 3

The Road to 2000

CONFLICT BETWEEN THE CENSUS BUREAU'S PLANNING process and concerns of key stakeholders, which spanned much of the decade of the 1990s, put the design and successful execution of the 2000 census in jeopardy. In February 1997, the U.S. General Accounting Office added the 2000 census to its list of 25 high-risk activities of the federal government (U.S. General Accounting Office, 1997). The general shape of the 2000 census remained in flux until a 1999 U.S. Supreme Court ruling forced final changes in plan. It is instructive to review the evolution of the 2000 census plan in order to realize the unprecedented difficulties that confronted the Census Bureau staff in 2000.[1]

In this chapter, we review the basic operations and major problems of the 1990 census (3–A). We then outline the major research programs conducted during the 1990s, including work on questionnaire design and address list development (3–B), and complete our history by describing the extended struggle to finalize the design of the 2000 census (3–C). We close the chapter with a fundamental finding and two recommendations that flow directly from the evolution of the 2000 census plan (3–D).

[1]Key sources for this history include: Anderson and Fienberg (1999, 2001); McMillen (1998); National Research Council (1995b, 1997, 1999b); Bureau of the Census (1997); U.S. Census Bureau (1998, 1999b); U.S. General Accounting Office (1992, 2002c).

3–A THE 1990 EXPERIENCE

The natural point to begin the story of one decennial census is to look back at the census that preceded it. Thus, we begin by summarizing the major census operations in 1990 and describing the problems that census experienced. These problems drove the research and planning agenda for 2000.

3–A.1 Major Operations

The 1990 census, broadly similar in structure to the 1970 and 1980 censuses, involved 10 major operations (see also Appendix C).

a. *Set up an organizational structure.* The census structure for 1990 (see Bureau of the Census, 1995b:Ch.1,8) included:

- staff at Census Bureau headquarters in Suitland, Maryland (over a dozen units handled budget, overall management, data processing, and other functions, five of which were directly under the associate director for decennial census, a new position in 1990);
- a back-up computer center in Charlotte, North Carolina;
- the Bureau's Data Preparation Division (later renamed the National Processing Center) in Jeffersonville, Indiana, which handled such operations as coding of place of work and served as a processing office;
- 13 regional census centers and the Puerto Rico Area Office, which managed the local field offices;
- seven processing offices (including the Jeffersonville facility), which handled data capture and also, for large central cities, computerized data editing and telephone follow-up for missing people and content; and
- 449 district offices (plus 9 offices in Puerto Rico), which handled address listing, outreach, nonresponse follow-up, and, in most areas, clerical review and follow-up for missing people and content.

b. *Develop an address list.* Construction of the 1990 Address Control File (ACF) began in 1988 using only a few sources. In urban areas two lists purchased from vendors were field checked and supplemented by a field listing operation. In rural areas a field-generated list was updated by enumerators as they delivered questionnaires. A separate address list was developed for dormitories, nursing homes, and other special places. The Census Bureau's innovative TIGER system was used to assign census geographic area codes to ACF addresses in mailout/mailback areas and to produce maps.[2]

c. *Design short-form and long-form questionnaires and mailing materials.* In 1990, in response to an administration effort to reduce the size and scope of the short and long forms (see Choldin, 1994), several housing items were moved from the short to the long form. The questionnaire format met the requirements of the data capture technology, but the questionnaires were unattractive and hard to read. The only additional mailing was a reminder postcard.

d. *Mail or hand deliver appropriate questionnaires to addresses on the address list.* In mailout/mailback areas—84 percent of total addresses in 1990—Postal Service carriers delivered the questionnaires. In update/leave/mailback areas—11 percent of addresses—Census Bureau enumerators dropped off questionnaires and updated the ACF at the same time. (Update/leave was a new procedure in 1990 designed for areas that largely lacked city-style addresses—street name and house number—for postal delivery.) Census staff enumerated the remaining 5 percent of the household population in person. Separate enumeration procedures were used for such special populations as homeless people, residents of group quarters, and transients.

e. *Carry out advertising and outreach campaigns.* This was intended to boost mail response and follow-up cooperation. The 1990 advertising campaign was conducted on a pro bono basis by a firm selected by the Advertising Council.

f. *Send enumerators to follow up addresses that failed to report.* The nonresponse follow-up (NRFU) operation for 1990 began at the end of April and continued through early August. Subsequent follow-up operations rechecked the classification of virtually all

[2]TIGER stands for Topologically Integrated Geographic Encoding and Referencing; see Glossary and Section 4–E.5.

vacant and nonresidential addresses, relisted and reenumerated households in selected areas believed to be not well covered, and checked the enumeration status of people on parole or probation. Telephone calls and personal visits were also used to obtain responses for missing content items.

These follow-up operations were central to the 1990 design, which emphasized the need to reduce the net undercount of the population and, especially, the difference between net undercount rates for minority and majority groups, by conducting as thorough a field enumeration as possible. In part because the Census Bureau took a position in court cases against statistical adjustment of the 1980 census for net undercount, the 1990 design maximized the use of telephone and field follow-up and minimized the use of computer-based statistical techniques to impute housing units and people for addresses of uncertain status. The Bureau was concerned that such imputation might be viewed as adjustment (see Chapter 6). The 1990 design also incorporated extensive follow-up for missing data content.

g. *Process the basic information* through the steps of data capture, identification and removal of duplicates, editing and imputation, tabulation, and data release. Most questionnaires were routed to local offices for clerical review to identify cases for telephone or in-person follow-up for missing people or items. The completed questionnaires were sent to a processing office and the data captured by an in-house microfilm-to-computer optical scanning operation (write-in responses were keyed by clerks from the paper questionnaires or microfilm images). Computerized routines were used at headquarters for imputation of remaining missing data and preparing data files for dissemination.

h. *Process the long-form-sample information* through the steps of data capture, weighting, coding, editing and imputation, tabulation, and data release. Data processing for the long-form sample was similar to that for the basic items on all records, except that item imputation routines were more complex to handle the greater number of long-form-sample items, and weighting was carried out so that estimates from the long-form sample would represent the entire population.

i. *Conduct coverage evaluation operations.* In 1990, demographic analysis estimates of the population were constructed for age, race, and sex groups at the national level, and a Post-Enumeration Survey (PES) Program was used to develop estimates for additional groups and subnational areas for comparison with the enumerated census counts. The 1990 PES estimated a net undercount of the population of 1.6 percent, which was consistent with the demographic analysis estimate of a 1.9 percent net undercount (revised in October 2001 to 1.7 percent). The Department of Commerce decided not to use the PES results to adjust the 1990 census counts.

j. *Conduct experiments and evaluations.* The 1990 census included a few experiments conducted during the census and evaluations of many aspects of procedures and results.

3–A.2 Major Problems

The 1990 census had important technological achievements, including a computerized address file, the development of the TIGER system for producing maps and coding addresses to census geography, computerized tracking and control of questionnaires, and concurrent processing in which the information was captured and processed on a flow basis (see Bryant, 2000). However, three major problems of the 1990 census captured public, congressional, and stakeholder attention and motivated planning for the 2000 census. They were: (1) unexpectedly low mail response rates; (2) increased costs fueled by the need to hire additional nonresponse follow-up enumerators and stretch out the time for follow-up; and (3) evidence of somewhat worse coverage of the population, particularly of minority groups and children, in comparison with 1980.

Poorer Mail Response

Critical to the efficiency, cost, and quality of a mailout/mailback census is a high level of cooperation from the public in mailing back forms, since that reduces the workload for nonresponse follow-up and provides more complete responses. The Census Bureau uses the mail response rate to estimate the nonresponse follow-up workload. It is the number of mailed-back questionnaires divided by the total number of addresses to which Postal Service carriers and census enumerators delivered questionnaires for mailback.[3] In planning the 1990 follow-up effort, the Census Bureau projected a mail response rate of 70 percent (down from 75 percent in 1980), but the actual rate at the time nonresponse follow-up began was 63 percent (the rate subsequently rose to 65 percent). A difficult-to read form and an inadequate publicity campaign were among the factors believed responsible.

To cope with the unexpectedly low response, the Bureau sought additional appropriations, scrambled to hire sufficient workers for follow-up activities, and took steps to increase productivity (e.g., raising pay rates in selected areas). The nonresponse follow-up operation had been planned to take 6 weeks from when it began in late April; however, the workload took over twice as long to complete, not finishing until the end of July (Bureau of the Census, 1993:6-34,6-36). Subsequent field work for some planned and ad hoc operations also experienced delays. From reviewing housing unit counts by block in the Postcensus Local Review Program, large cities and other jurisdictions claimed that the Bureau had missed significant numbers of housing units. From July through November, the Bureau revisited about 20 million housing units (20 percent of the housing stock) in selected blocks in high-growth areas and for which local governments reported higher counts than the census. In 31 local offices, some or all households were reenumerated because of allegations that enumerators had fabricated data. A "Were You Counted" campaign (similar to the "Be Counted" effort in 2000) extended from June through November. A last-minute special program was implemented to improve the coverage of people on parole or probation, but the operation had low response, was not completed until early December, and, according to subsequent analysis, included a high proportion of erroneous enumerations. Some observers (e.g. Ericksen et al., 1991) concluded that the 1990 census was poorly executed overall.

Higher Costs

In inflation-adjusted fiscal 2000 dollars, the cost of the census rose from $24 per household in 1980 to $32 per household in 1990. (The

[3]See Glossary and Box 4.1 for the distinction from the mail return rate.

census had cost only $13 per household in 1970; see U.S. General Accounting Office, 2001b:6, and Table 3.1 below.) Part of the additional cost in 1990 was due to the unanticipated additional workload for nonresponse follow-up and other field operations.

Worse Coverage

The additional money and effort for the 1990 census did not buy improved coverage: the net undercount, measured by demographic analysis, increased from 1.2 percent of the population (2.8 million people) in 1980 to 1.9 percent of the population (4.7 million people) in 1990. Net undercount rates for blacks increased from 4.5 percent in 1980 to 5.7 percent in 1990; net undercount rates for nonblacks increased from 0.8 percent in 1980 to 1.3 percent in 1990 (National Research Council, 1995b:Table 2-1).[4] Because net undercount rates increased more for blacks than others, the difference in coverage rates for blacks and nonblacks widened from 3.7 percentage points in 1980 to 4.4 percentage points in 1990, reversing the historical trend toward declining net undercount and the narrowing of differences among important groups. The October 2001 revised demographic analysis estimates for 1990 do not change this story—the revisions lowered the total net undercount rate from 1.9 percent to 1.7 percent, the black net undercount rate from 5.7 to 5.5 percent, and the nonblack net undercount rate from 1.3 to 1.1 percent, leaving the difference between the black and nonblack net undercount rates unchanged at 4.4 percentage points (Robinson, 2001b).

3–B RESEARCH AND DEVELOPMENT TOWARD 2000

The perceived failures of the 1990 census—low public response, high costs, and worse coverage—drove planning for the 2000 census, which began in late 1991 accompanied by oversight from an unprecedented array of outside groups (see Section 3–C.1). The Bureau devoted significant attention and resources to research in four major areas: (1) how to make the census address list more complete

[4]Net undercount estimates for 1990 from demographic analysis differ somewhat from those from the PES; see Chapter 5. The demographic analysis estimates are used for 1990 for comparison with 1980 for which no accepted PES results are available (see National Research Council, 1985:Ch.4).

by taking advantage of prior work by the Bureau and other organizations; (2) how to make the questionnaire more attractive and the mailing materials more compelling in order to bolster response; (3) how to streamline nonresponse follow-up to reduce costs and time and possibly reduce coverage errors; and (4) how to implement a coverage evaluation survey of a size and on a schedule to permit the results to be used to adjust the census counts for reapportionment, redistricting, and other uses of the data. Another line of research—on the possible uses of administrative records (e.g., tax records, Social Security records) to assist in the census enumeration—proceeded to the point of assembling administrative records databases for several sites and testing their completeness in a 1995 census test. However, administrative records research was dropped except for an experiment as part of the 2000 census (see National Research Council, 2000a).

3–B.1 Address List Research: Development of the Master Address File

Almost before the 1990 census data collection was completed, the director of the Census Bureau endorsed the concept of maintaining a Master Address File (MAF) on a continuous basis from one census to the next. The expectation was that a continuous updating process would, in the end, be cheaper and would produce a more accurate and complete address list than the process used since 1970 of building a new list from scratch each decade a few years prior to the census. The importance of an accurate MAF was underscored by research estimating that two-thirds of the people missed in the 1990 census were in housing units or structures that were missed, presumably because of problems in the ACF; the other one-third of missed people were members of enumerated households (Childers, 1993).

In urban areas two sources were expected to provide the bulk of the updates to the 1990 census address list: (1) the U.S. Postal Service Delivery Sequence File (DSF), which is the master list of addresses for mail delivery maintained by the Postal Service, organized by carrier route; and (2) input from local governments from reviewing the address list for their area. Field work in selected areas would be used to spot check the completeness of the list. In rural areas lacking city-style addresses a field listing operation was planned, supplemented by local government review. This MAF development concept quickly won wide acceptance (see, e.g., Commerce Secretary's 2000 Census Advisory Committee, 1995; National Research Council, 1994, 1995b).

Because of confidentiality constraints on sharing address lists, legislation was needed to implement the concept. The 1994 Census Address List Improvement Act (P.L. 103-430) authorized the Postal Service to share the DSF with the Census Bureau. It also set up a process whereby the Census Bureau was required to allow interested local governments to review the addresses for their area, swearing in local officials as census agents for this purpose. (Local review efforts in 1990 and 1980 limited participating localities to review of preliminary housing unit counts by block for their areas, not individual addresses.)

At 6-month intervals beginning in 1995 (later than originally hoped), the Bureau received the latest DSF, which included newly constructed units, to match to and update the MAF. However, internal Census Bureau evaluation suggested that the DSF files were not updated at the same rate across all areas of the country, missed many addresses for new construction, and often did not identify individual addresses in multiunit structures. Moreover, planning to involve local governments in address review fell behind schedule. The Bureau developed a Program for Address List Supplementation (PALS), which invited local governments to send address files to the Bureau for matching to and updating the MAF, but participation was low and files arrived in such variable formats (including paper) that they were not usable.

Hence, in September 1997 the Bureau announced a revised and more costly plan for producing a 2000 MAF of acceptable quality. Instead of conducting field checks of the MAF in selected areas, the new plan called for blanket field canvassing operations in 1998 and 1999 similar to those carried out for the 1990 census. Local governments in urban areas would be given the opportunity to review MAF addresses for their jurisdiction in the 1998 Local Update of Census Addresses (LUCA 98) Program; governments in rural areas would be given housing unit counts by block to review in the 1999 Local Update of Census Addresses (LUCA 99) Program.

3–B.2 Questionnaire and Mailing Package Research

Questionnaire research had high priority in the early 1990s because of the perception that the format and length of the 1990 questionnaires, particularly the long form, had adversely affected response. Mail response rates fell slightly between 1970 and 1980 (78 to 75 percent) and dropped sharply between 1980 and 1990 (75 to 65 percent).

To control the number of questions, the Census Bureau decided that items could not be asked of everyone unless block-level counts for the items were mandated in legislation or federal regulations and should not be asked of those in the long-form sample unless legal or regulatory authority could be cited for those data for census tracts or larger geographic areas.[5] Consequently, most housing items that were complete-count items in 1990 were relegated to the 2000 long-form sample or dropped; marital status was also moved to the long form. Ultimately, the 2000 short form had only five population items and one housing item, compared with six population items and seven housing items in 1990 (see Appendix B).

The long form had both additions and deletions. Some 1990 long-form items were dropped entirely for 2000, but Congress mandated a new set of long-form questions on responsibility for care of grandchildren in the home. On balance, the 1990 and 2000 long forms included about the same number of questions, but the 2000 long form (and short form) provided space for characteristics of only six instead of seven household members as in 1990. Additional household members' names could be listed but not their characteristics. Mail returns that indicated more than six persons in the household were to be contacted by telephone to obtain basic characteristics for the additional members. (For enumerations obtained in the field, the questionnaire included space to record characteristics of five household members; enumerators were to fill out continuation forms for additional household members.)

The issue of how much room the questionnaires would provide

[5]See http://www.census.gov/dmd/www/content.htm [12/1/03] for the legal or regulatory authority for each of the 2000 census questions, as compiled by the Census Bureau in March 1998. Input from stakeholder groups, such as the Association of Public Data Users, was instrumental in retaining some items on the 2000 long form that were not specifically mandated or required for federal programs.

for individual household member characteristics was debated over the decade. Leaving space for detailed data from seven household members had merit but, when combined with efforts to enhance the readability and the attractiveness of the questionnaire, could make the document look long and burdensome and discourage response. Early testing used questionnaires that provided space for characteristics of only five household members, with the intent to use telephone follow-up to obtain characteristics for additional household members when necessary. However, providing space for only five members was believed likely to lead to undercount of members of large households, so a compromise was reached on providing space for six members. The desire to limit the length of the 2000 census forms also made it infeasible to implement some findings from a 1994 Living Situation Survey, which suggested cues that might improve coverage of such hard-to-count groups as young minority men who floated among several households (see Martin, 1999).

With regard to the format of the questionnaire and mailing package, the Census Bureau conducted several experiments early in the 1990s to identify formats that would maximize response. A decision to use optical mark and character recognition technology for data capture made it easier to create more visually appealing questionnaires in 2000 than those used in previous censuses (see Appendix C.5).

In a 1992 test, mail response to a user-friendly short form was 3.4 percentage points higher than response to the type of short form used in 1990; the difference in response rates for areas that were hard to enumerate in 1990 was even greater, 7.6 percentage points. A short form with fewer questions than the 1990 form further boosted mail response (Dillman et al., 1993). The results from that experiment and a second experiment suggested that the use of more mailings could substantially increase response. It appeared that sending an advance letter (not used in 1990) could increase response by 6 percentage points; that sending not only an advance letter, but also a reminder postcard (used in 1990) could increase response by a total of 13 percentage points; and that sending a second questionnaire to nonrespondents (not used in 1990) could increase response by another 10 percentage points (Dillman et al., 1994). Yet another test in 1993 demonstrated that stressing the mandatory nature of filling out the questionnaire on the envelope was effective in encouraging response, while emphasizing the benefits of the data or their confidentiality was not particularly effective (National Research Council, 1995b:120–121). Finally, a 1993 test that focused on long-form response rates found a positive effect on response rates from appealing forms and multiple mailings. Disturbingly, however, it found an 11 percentage point difference between short-form and long-form response rates; the difference was 15 percentage points in hard-to-enumerate areas (Treat, 1993). The Census Bureau hoped that the increased publicity in a census environment would narrow the differences between short-form and long-form response rates, but that did not happen in 2000 (see Section 4–B.1).

The Bureau decided to adopt most of the strategies that the questionnaire and mailing package research supported for 2000—specifically, creating more attractive questionnaires, mailing an advance letter and reminder postcard (in mailout/mailback areas), and emphasizing on the envelope that responses to the census are required by law. However, the Bureau had not conducted operational research on how to deliver replacement questionnaires only to nonrespondent households (a targeted mailing) on the scale and time schedule that would be required for 2000. At the time of the census dress rehearsals in 1998, vendors informed the Bureau that a targeted mailing was not feasible within the short period of time allowable for the operation. Instead, replacement questionnaires would have to be mailed to all households on the MAF (a blanket mailing).

Blanket mailing of replacement questionnaires to all households was tested for the first time in the 1998 dress rehearsals. The second mailout increased mail response rates for the two sites using mailout/mailback methods by 8 percent (increasing response rates from about 47 percent to about 55 percent), but there was a considerable amount of duplication as well as concern that a blanket second mailing would appear wasteful. Consequently, the Census Bureau chose not to send a second questionnaire at all in 2000.

3–B.3 Nonresponse Follow-up Research

Nonresponse follow-up and related activities amounted to 54 percent of total census costs in 1990 (National Research Council, 1995b:51), or $1.8 billion of $3.3 billion total costs (in fiscal 2000 dollars). Consequently, there was great interest after 1990 to investigate ways to reduce these costs and increase the timeliness of the follow-up effort and thereby possibly reduce coverage errors as well. The Census Bureau made a commitment in the mid-1990s to use sampling and imputation for the nonresponse follow-up operation. That is, instead of enumerating every household that failed to mail back a questionnaire, the Bureau would enumerate only a sample of such households. Some form of imputation would be used to estimate the characteristics of nonsampled households from the characteristics of sampled households in the same neighborhoods (see Wright, 2000). In addition, the Bureau would also visit a sample of addresses that the Postal Service reported were vacant. To improve follow-up operations, the Bureau also conducted research on ways to increase enumerator hires and productivity by varying wage rates according to local labor market conditions.

Debate ensued about the particular sample design to use. The lower the sampling rate, the more money and time would be saved in enumeration, but the more troubling the appearance to Congress, other stakeholders, and the public that so many nonresponding households would not be visited. Congress and other stakeholders expressed interest in a design in which every nonresponding household would receive at least one visit, with sampling used only for households that could not be reached during that first visit. Such a design, however, saved little on costs and added complexity to the field operations. The Bureau determined instead to use "direct sampling," that is, to sample at the start of nonresponse follow-up, using higher sampling rates for census tracts with lower mail response rates and vice versa so that a total of 90 percent of households in each census tract would either have mailed back a questionnaire or been enumerated in person. (Units identified as vacant by the Postal Service would be sampled at a rate of 30 percent.) Consistent with a recommendation of a National Research Council (1997:40) panel, the Bureau decided that no tract would be sampled at a rate lower than 1 in 3, even though that might result in achieving more than a 90 percent overall response rate. The reason was not to unduly increase the sampling variability of the estimates for high mail response tracts.

The same National Research Council panel concluded that the results from sampling for nonresponse follow-up would be of equal or better quality than the results from a complete follow-up enumeration effort for such important purposes as drawing congressional district boundaries (National Research Council, 1999b:26–30). Three reasons were: (1) follow-up with sampling could be completed in a more timely manner, reducing coverage errors due to people moving between Census Day and follow-up; (2) follow-up with sampling could possibly reduce the use of proxy respondents (neighbors and landlords), thereby obtaining more complete and accurate data; and (3) the error introduced from sampling variability could be reliably determined for any level of geography and would be smaller for larger areas (e.g., congressional districts built up from individual blocks). The panel report further explained why it would be impossible to use sampling for nonresponse follow-up as an opportunity for political manipulation to gain a desired outcome, either in the design or the field implementation.

The sample design the Census Bureau adopted for nonresponse follow-up—known as sampling for nonresponse follow-up or SNRFU—would not save as much money or time as a design with a lower sampling rate overall. It appeared reasonable, however, for the first use of sampling as part of the census enumeration. Ultimately, the design could not be used for 2000 because of the January 1999 Supreme Court decision that precluded the use of sampling for reapportionment (see Section 3–C.4).

3–B.4 Coverage Evaluation and Adjustment Research

The final major area of research for the 2000 census was to determine the most effective way to measure the completeness of population coverage in the census and be able to use the coverage measurement results to adjust the enumeration for net undercount. By the mid-1990s the Census Bureau accepted the view expressed by National Research Council panels and others that it was not possible to reduce net undercount—particularly the higher undercount rates for minorities and children compared with white non-Hispanic people and adults—by traditional enumeration methods, even with substantial resources spent on special coverage improvement programs (National Research Council, 1994, 1995b; Commerce Secretary's 2000 Census Advisory Committee, 1995). Given a mandate to improve coverage and contain costs, it would be necessary to use statistical estimation to account for the hard-to-enumerate population.

It was expected that demographic analysis would continue, as in the past, to provide a benchmark for the 2000 census results for age, race, and sex groups at the national level; however, demographic analysis could not provide reliable estimates for subnational geographic areas or for race and ethnicity groups other than blacks and nonblacks. Little research attention was devoted to demographic analysis in the 1990s.

Research concentrated on refining the methodology for dual-systems estimation (DSE) using the results of an independent population survey (P-sample) matched with a sample of census enumerations (E-sample) to estimate the population. There was also research to compare DSE with a method called Census Plus, in which enumerators would revisit a sample of households to obtain a roster of household members and immediately reconcile that roster with the roster from the census. When weighted, the union of the two rosters (after making any deletions or additions) would provide an estimate of the population.

A major test in 1995, which incorporated sampling for nonresponse follow-up, Census Plus, and refined DSE procedures, determined that Census Plus was itself subject to undercounting. Hence, subsequent research focused on further improving the DSE methodology, particularly the ability to conduct the independent survey and other DSE operations on a schedule to permit using the results to adjust the state population counts for reapportionment by December 31, 2000. This use of DSE for census adjustment was termed Integrated Coverage Measurement (ICM). The Census Bureau proceeded with plans for ICM but—as discussed in Section 3–C.4—a 1999 Supreme Court decision forced the move from ICM to a more limited-scale Accuracy and Coverage Evaluation (A.C.E.) Program.

3–B.5 Efficacy of Research

This brief review indicates the substantial amount of research that the Census Bureau devoted to planning innovations in major areas of census operations for 2000. The questionnaire and mailing package research was particularly effective in that the 2000 design incorporated most of the results of this research (e.g., multiple mailings and readable questionnaires), and these innovations apparently helped stem the historical decline in mail response rates (see Section 4–B.1). The coverage evaluation research supported the basic DSE survey-based design and led to innovations in the A.C.E. that promised to achieve gains in accuracy and timeliness compared with the 1990 PES. Many of these gains were in fact achieved, although a failure to detect large numbers of duplicate enumerations compromised the A.C.E. population estimates (see Section 6–C.2).

Research in the other two priority areas had less of an impact. The research on sampling for nonresponse follow-up had to be discarded because the Supreme Court decided that a sample-based census was unlawful. The MAF research did not proceed far enough in the early 1990s, with the result that additions to the program and changes in the schedule—late implementation of the full block canvass, rushed plans for local review, and questionnaire labeling before verification of most of the addresses submitted by local governments—had to be made late in the decade.

3–C DETERMINING THE 2000 DESIGN

We now turn to the political and advisory processes that shaped the 2000 census design over the 1990s. In the first half of the decade, the balance of stakeholder opinion provided impetus to the Census Bureau to plan an unprecedented use of statistical methods to reduce costs and improve the completeness of population coverage in 2000. In the second half of the decade, a battle ensued between those who welcomed and those who feared statistical adjustment of the census counts. The result was to delay a final decision on the 2000 design and the provision of needed funding until very late in the decade.

3–C.1 1991 to 1996

The concern over the perceived failings of the 1990 census led the Census Bureau, Congress, and stakeholder groups to take an unprecedented interest in beginning to plan for 2000 before release of data products had even been completed for 1990. In late 1991 the secretary of commerce established a 2000 Census Advisory Committee, consisting of over 30 representatives from a wide range of associations representing business, labor, minority groups, data users, scientific professions, state and local governments, and others (see Commerce Secretary's 2000 Census Advisory Committee, 1995, 1999). This committee met several times a year over the decade (and continues to meet on planning for the 2010 census). Also meeting regularly during the 1990s on 2000 census issues were the Bureau's long-established Advisory Committee of Professional Associations and advisory committees for minority groups (Citro, 2000a). Oversight committees in Congress were active in holding hearings and taking such steps as passing the 1994 Address List Improvement Act to facilitate development of the MAF for 2000.

Two Committee on National Statistics panels were established in 1992 to address 2000 census planning. One panel was convened at the behest of Congress to undertake a thorough review of data requirements and alternative designs for 2000 (National Research Council, 1995b); the other panel was convened at the Census Bureau's request to consider detailed methodology (National Research Council, 1994). Subsequently, a third CNSTAT panel was organized in 1996 to comment periodically on the Bureau's maturing plans for 2000; this panel issued its final report in 1999 (National Research Council, 1999b). All three panels supported the Bureau's research on DSE and sampling for nonresponse follow-up as ways to contain costs and improve accuracy in the 2000 enumeration.

The Bureau initially presented a range of alternative designs to these groups and others, such as the U.S. General Accounting Office and the Department of Commerce Inspector General's Office. It then widely shared the results of research and testing as it homed in on a design.

3–C.2 Original Design for 2000

As early as 1993 the Bureau adopted the overall concept of a "one-number census" for 2000, in which a combination of counting, sampling, use of administrative records, and estimation would be used to produce the best possible population data for reapportionment, redistricting, and other uses of census data (Miskura, 1993). In contrast, the 1990 census used a dual strategy, in which two sets of population totals were produced—from the enumeration and adjusted for coverage errors using the PES—and an ex post facto decision was made about which set to use (see Chapter 5).

In early 1996, following a major test in 1995, the Bureau settled on a specific design for 2000, which would have used sampling for nonresponse follow-up and Integrated Coverage Measurement (see Sections 3–B.3 and 3–B.4). Other innovative features of the 2000 design compared with 1990 included development of the Master Address File from multiple sources in order to reduce census omissions because of overlooked addresses (see Section 3–B.1); use of redesigned mailing materials and questionnaires to promote mail response (see Section 3–B.2); expansion of advertising and outreach and use of paid advertising; aggressive recruitment of enumerators; adoption of new data capture technology; and hiring of contractors for data capture and other operations instead of conducting all operations with census staff. Finally, the design deemphasized field work, not only as a way to improve completeness of coverage but also as a way to obtain data content, relying heavily on computer editing and imputation in place of additional follow-up for missing household data. Indeed, the 2000 design focused primarily on coverage and much less on data content in comparison with the 1990 design.

3–C.3 1996 to 1998: Two Dress Rehearsals

Views in the Congress and the executive branch on the merits of the Bureau's 2000 census plan—particularly, the proposed use of SNRFU and ICM—divided along partisan lines: the Clinton administration and House Democrats generally supported the use of sampling while House Republicans generally opposed its use (see McMillen, 1998). Complicating matters throughout 1996 and into 1997 was that the Bureau had not yet issued a detailed plan of operations for the census (see U.S. General Accounting Office, 2002c). In Title VIII of an emergency supplemental appropriations act, Congress ordered the Census Bureau to provide a detailed plan within 30 days of enactment; the act was signed into law as P.L. 105-18 on June 12, 1997. The Census Bureau plan was delivered July 12 (Bureau of the Census, 1997), but it was discovered to have an error in calculation of the likely savings from sampling for nonresponse follow-up, which hurt credibility with Congress.

Conflict between the two sides resulted in a delay of funds with which to conduct the 1998 dress rehearsal. In fall 1997 compromise legislation stipulated that unadjusted data would be released from the 2000 census in addition to any adjusted data. It also provided for expedited judicial review of the legality of the use of sampling for the census and established a 2000 Census Monitoring Board to consist of four members appointed by House and Senate Republican leaders and four members appointed by President Clinton in consultation with House and Senate Democratic leaders. The Monitoring Board was proposed as a way to address the concerns of some in Congress that the Census Bureau might manipulate the census data for political gain. The congressional and presidential appointees each had their own budgets and staffs. They were charged to report periodically to Congress through September 2001 and then go out of existence. In none of its reports did either the congressional or presidential members of the Monitoring Board cite evidence or make arguments suggesting partisan intent in the choice of design features or implementation of the 2000 census (see U.S. Census Monitoring Board Congressional Members, 1999, 2001; U.S. Census Monitoring Board Presidential Members, 2001a,b).

With the mechanism of the Monitoring Board in place, there was agreement that the dress rehearsal and other necessary census planning activities would be fully funded and that the dress rehearsal in one of the planned sites would be conducted without the use of either SNRFU or ICM. The House of Representatives also established a Subcommittee on the Census in the Committee on Government Reform and Oversight, which began census oversight activities in early 1998 and ceased operations at the end of 2001.[6]

The Census Bureau consequently adjusted its planning to allow for two possible census designs—one based on SNRFU and ICM, the other implementing follow-up on a 100 percent basis and using a postenumeration survey to evaluate coverage. The Bureau implemented each design in a dress rehearsal in spring 1998: the rehearsal in Sacramento, California, used SNRFU and ICM; the one in Columbia, South Carolina, and surrounding counties, used 100

[6]At the end of 2001, authority for census issues was vested in a Subcommittee on Civil Service, Census, and Agency Reorganization. Subcommittee structures were revised when the chairmanship of the full House Committee on Government Reform changed hands in the 108th Congress. Census issues are now in the jurisdiction of the Subcommittee on Technology, Information Policy, Intergovernmental Relations, and the Census.

Box 3.1 *Department of Commerce v. U.S. House of Representatives:* Sampling in the 2000 Census

In 1997, amid debate over the role of sampling methods in the 2000 census, Congress passed an amendment to the Census Act, Title 13 of the U.S. Code, prohibiting "sampling or any other statistical procedure, including any statistical adjustment," in deriving counts for congressional apportionment (H.R. Conf.Rep. 105-119). However, the appropriations bill containing this provision was vetoed by President Clinton. Subsequently, Congress revised the proposed language, dropping the prohibition on sampling but requiring the Census Bureau to prepare to test a census design "without using statistical methods" (111 Stat. 217). The same legislation also required expedited judicial review of cases filed by parties who argued that they would be harmed by implementing sampling methods in the census; among the parties explicitly authorized for this fast-track review were members of Congress or either house of Congress.

Two subsequent legal challenges were heard by federal district courts: *Glavin v. Clinton*, filed in Virginia by four counties and residents of 13 states, and *U.S. House of Representatives v. Department of Commerce*, filed in the District of Columbia. In both cases, the district courts concluded that the Census Bureau's proposed use of sampling for purposes of generating apportionment counts violated the Census Act; having concluded such, neither court addressed the question of whether the sampling methods were unconstitutional. The two cases were consolidated on appeal and heard by the U.S. Supreme Court on November 30, 1998.

On January 25, 1999, Justice Sandra Day O'Connor's opinion for a 5–4 Court majority concurred with the lower court findings and prohibited the use of sampling-based methods in tabulating population counts for reapportioning the U.S. House of Representatives (*Department of Commerce v. United States House of Representatives*, 525 U.S. 316). Chief Justice William Rehnquist and Justices Antonin Scalia, Anthony Kennedy, and Clarence Thomas joined the key portions of the O'Connor opinion, with Justice Stephen Breyer joining the section regarding the standing of the litigants to bring their suits; one subsection of the O'Connor opinion on the lack of discussion of relevant sampling issues during congressional debate was not adopted by a majority of the Court.

Census Act language passed in 1957 provided that "[e]xcept for the determination of population for apportionment purposes, the [secretary of commerce] may, where he deems it appropriate, authorize the use of the statistical method known as 'sampling' in carrying out the provisions of this title" (this initial authorization allowed for the first long-form sample involving questions not asked of all respondents; see Section 2–D). The majority Court opinion concluded that subsequent revisions and interpretations of the language must be interpreted in the "context . . . provided by over 200 years during which federal statutes have prohibited the use of statistical sampling where apportionment is concerned. In light of this background, there is only one plausible reading of the [Census Act language, as revised in 1976]: It prohibits the use of sampling in calculating the population for purposes of apportionment."

Box 3.1 (continued)

The Court continued that, because the methodology was prohibited by existing statute, "we find it unnecessary to reach the constitutional question presented"—that is, whether sampling is unconstitutional rather than simply unlawful for purposes of deriving apportionment counts. The applicability of sampling-based methods for purposes other than apportionment was not at issue.

Principal dissents were filed by Justices John Paul Stevens and Stephen Breyer. Stevens argued that Census Act language enabling the secretary of commerce to "take a decennial census of population . . . in such form and content as he may determine, including the use of sampling procedures and special surveys" constitutes an "unlimited authorization" that is not impaired by the "limited mandate" contained elsewhere in the act. Breyer's dissent also concluded that the Census Act does not prohibit sampling for purposes of apportionment counts, but on different grounds. Specifically, Breyer argued that the Census Act "permits a distinction between sampling used as a substitute and sampling used as a supplement;" interpreting the Census Bureau's proposed ICM plan as a complement to rather than substitution for a "traditional headcount," Breyer found no prohibition in the statute. Justice Stevens extended his argument to the issue of constitutionality of sampling methods, arguing that "the words 'actual Enumeration' require post-1787 apportionments to be based on actual population counts, rather than mere speculation or bare estimate, but they do not purport to limit the authority of Congress to direct the 'Manner' in which such counts should be made."

percent nonresponse follow-up with a postenumeration survey. A third dress rehearsal in Menominee County, Wisconsin (which includes the Menominee Indian Reservation), was a hybrid using 100 percent nonresponse follow-up and ICM.

3–C.4 1999: Supreme Court Decision

In January 1999 the U.S. Supreme Court ruled that existing law (in Title 13 of the U.S. Code) prohibited the use of sampling in producing population counts for congressional reapportionment. Further detail on the Court's decision is given in Box 3.1. The ruling did not preclude the use of sampling to adjust census data for other purposes.

Following the Supreme Court decision, the Census Bureau again revised its plans. The final design for the 2000 enumeration was announced by Director Kenneth Prewitt at a press conference in February 1999, little more than a year before Census Day on April 1, 2000 (see Section 3–C.5). After intense debate, Congress approved the design and provided the full amount of funding—over $4.5 billion—requested by the Census Bureau for fiscal year 2000. Total appropriations for fiscal years 1991–2003 provided a budget of over $7 billion for the 2000 census.

3–C.5 The 2000 Census Design

The final design for 2000 dropped the use of SNRFU and ICM and adopted 100 percent follow-up of nonresponding households. It also expanded the number of field offices to manage the enumeration (from the originally proposed number of 476 to 520 offices, excluding Puerto Rico) and expanded the advertising program. For these reasons, the fiscal year 2000 budget request for census operations requested an increase of 61 percent over the amount originally requested (U.S. General Accounting Office, 1999a:4).

The final design, however, retained other features originally planned for 2000, including the emphasis on computer imputation to supply missing data for households and persons in place of repeated follow-up attempts and the use of multiple sources to develop the MAF. The MAF development procedures were modified in the late 1990s to respond to problems in the original plans, and a special operation to reduce duplicate enumerations resulting from duplicate MAF addresses was added to the census in summer 2000.

The final 2000 design included the Accuracy and Coverage Evaluation Program, which fielded an independent household survey smaller than that planned for ICM (300,000 compared with 700,000 households), but larger than the 1990 PES (165,000 households). The plan was to use the A.C.E. results, if warranted, to adjust the census counts for redistricting and all other purposes except reapportionment of the Congress. The first adjusted results under this scheme would be released by April 1, 2001. As described in Chapter 5, this plan was ultimately not carried out.

3–D FINDINGS AND RECOMMENDATIONS

The lack of agreement among key stakeholders about an appropriate design adversely affected the planning and decision-making process for the 2000 census and added to its costs. The level of rancor and mistrust between the Census Bureau and Congress, and between parties within Congress on census issues, reached distressing heights. An overstretched Census Bureau staff spent much energy late in the decade on developing plans for alternative designs—with and without extensive use of statistical sampling and estimation—and consequently could not make firm plans for implementation. The Bureau staff also faced unprecedented demands for information and testimony on an almost daily basis in response to an extremely large number of congressional, administrative, and nongovernmental oversight groups. The 1998 dress rehearsal was a large-scale comparative test, not a rehearsal of an agreed-on plan to work out the operational kinks. Uncertainty about funding impeded hiring of needed staff and resolution of specific design elements until less than a year before the census was to begin (see Waite et al., 2001).

The last-minute decision on the design and provision of funding put the census in substantial jeopardy. Specific procedures for several operations were finalized very late. Likewise, many data management and processing systems were specified and developed at the last minute and implemented almost as soon as they were completed, without benefit of advance testing. The cost of the census escalated greatly compared with 1990—from $32 per household enumerated to $56 per household enumerated (in fiscal 2000 dollars). The cost increase was due not only to the decision to conduct a basically traditional census with complete field follow-up for household coverage, but also to the lateness of the design decision, which meant that careful planning for cost efficiency was not possible. To get the job done, the Census Bureau poured on resources; there was no opportunity for careful consideration of how to constrain costs at the same time. (See Table 3.1 for a summary of the total costs of the past four decennial censuses.)

To the Bureau's credit, the census was conducted on time and within the specified budget ($400 million of the $7 billion was not spent). Most operations went smoothly (see Chapter 4). However, some of the problems in the census—such as difficulties in enumerating group quarters (see Section 4–F)—very likely can be traced, at least in part, to the haste with which the Bureau had to develop specific operational plans to implement the final agreed-on design.

Table 3.1 Decennial Census Costs, Total and Per Housing Unit, 1970–2000 (in constant fiscal year 2000 dollars)

Census	Period	Costs (billions of dollars)	Cost per Housing Unit (dollars)
1970	1964–1973	$0.9	$13
1980	1974–1983	2.2	24
1990	1984–1995	3.3	32
2000	1991–2003	6.6	56

SOURCE: U.S. General Accounting Office (2001b:Table 1).

In structuring this report, we defer most of our findings and recommendations to specific chapters devoted to assessment, but this chapter's narrative of the convoluted development of the 2000 census plan is the natural point to highlight a most basic finding:

> **Finding 3.1:** The lack of agreement until 1999 on the basic census design among the Census Bureau, the administration, and Congress hampered planning for the 2000 census, increased the costs of the census, and increased the risk that the census could have been seriously flawed in one or more respects.

Looking ahead to the next census cycle, it is in the interest of the Census Bureau, Congress, and the public that the design and planning process for the 2010 census proceed as smoothly as possible. Of course, there should be continued input and critical review of the Bureau's planning and research from a wide range of stakeholders. Yet there should be a commitment by all parties to reach a decision on the basic design by 2006 so that the dress rehearsal in 2008 can be as effectively planned as possible.

The need to reach early closure on the design of the 2010 census is heightened by the fundamental departure of the Census Bureau's emerging 2010 census plan from the 2000 census model. In particular, the Census Bureau has proposed replacing the census long-form-sample with the American Community Survey (ACS). The ACS is anticipated to include 250,000 housing units per month—sufficient sample size to provide detailed population profiles for counties, cities, and other local areas. Yearly estimates would be produced for areas of at least 65,000 population; estimates for areas with smaller populations would be based on aggregating the data over 3- or 5-year periods.

Our companion Panel on Research on Future Census Methods reviewed the current plans for the ACS. The panel concluded that significant work must still be done by the Census Bureau in evaluating the relative quality of ACS and census long-form-sample estimates and in informing data users and stakeholders of the features and the problems of working with moving average-based estimates. However, the panel concluded (National Research Council, 2003a:7):

> We do not see any looming flaw so large in magnitude that full ACS implementation should be set aside. We therefore encourage full congressional funding of the ACS. It is important, though, that Congress recognize that funding of the ACS should be viewed as a long-term commitment. Cuts in funding in subsequent years (and with them reductions in sample size) will impair the overall quality of the survey, with first and most pronounced impact on the ability to produce estimates for small geographic areas and population groups.

In concordance with the Panel on Research on Future Census Methods, and consistent with the history of the 2000 census plan, our panel agrees that the replacement of the census long-form sample is a sufficiently sweeping design choice that agreement must be secured early to facilitate a successful 2010 census.

> *Recommendation 3.1:* **The Census Bureau, the administration, and Congress should agree on the basic census design for 2010 no later than 2006 in order to permit an appropriate, well-planned dress rehearsal in 2008.**

> *Recommendation 3.2:* **The Census Bureau, the administration, and Congress should agree on the overall scheme for the 2010 census and the new American Community Survey (ACS) by 2006 and preferably earlier. Further delay will undercut the ability of the ACS to provide, by 2010, small-area data of the type traditionally collected on the census long form and will jeopardize 2010 planning, which currently assumes a short-form-only census.**

Chapter 4

Assessment of 2000 Census Operations

We begin our evaluation of the 2000 census by describing two of its three most important achievements (4–A): first, halting the historical decline in mail response rates and, second, implementing data collection, processing, and delivery in a smooth and timely manner. (We describe and assess the third important achievement—the reduction in differential coverage errors among important population groups—in Chapters 5 and 6.) We then assess the contribution of seven major innovations in census operations, not only to the realization of the above-cited achievements, but also to data quality and cost containment. By quality, we mean that an estimate from the census accurately measures the intended concept, not only for the nation as a whole, but also for geographic areas and population groups. With regard to costs, although cost reduction drove the research and planning for the original 2000 design (see Chapter 3), it was not an explicit goal for the final 2000 design.

The seven innovations that we review include two that facilitated public response—namely, redesigned questionnaires and mailing materials and paid advertising and expanded outreach (4–B, see also Appendix C.2, C.4); three that facilitated timeliness—namely, contracting for data operations, improved data capture technology, and aggressive recruitment of enumerators and implementation of nonresponse follow-up (4–C, see also Appendix C.3, C.5); one that facilitated timeliness but may have had mixed effects on data quality—namely, greater reliance on computers to treat missing data (4–D, see also Appendix C.5); and one that was sound in concept but not well-implemented—namely, the use of multiple sources to develop the Master Address File, or MAF (4–E, see also Appendix C.1). We do not assess an eighth major innovation in 2000, which was the expanded use of the Internet for release of data products to users; see http://factfinder.census.gov [12/12/03].

The final section in this chapter (4–F) assesses the problem-plagued operations for enumerating residents of group quarters. Sections 4–B through 4–F end with a summary of findings and recommendations for research and development for 2010.

4–A TWO MAJOR OPERATIONAL ACHIEVEMENTS

4–A.1 Maintaining Public Cooperation

The 2000 census, like censuses since 1970, was conducted primarily by delivering questionnaires to households and asking them to mail back a completed form. Procedures differed somewhat, depending on such factors as type of addresses in an area and accessibility; in all, there were nine types of enumeration areas in 2000 (see Box C.2 in Appendix C). The two largest types of enumeration areas in 2000—mailout/mailback and update/leave/mailback or update/leave—covered 99 percent of the household population; together, they constituted the mailback universe. The goal for this universe for the questionnaire delivery and mail return phase of the census was to deliver a questionnaire to every housing unit on the MAF and motivate people to fill it out and mail it back.

The Census Bureau expected that mail response would continue to decline, as it had from 1970 to 1990, due to broad social and economic changes that have made the population more difficult to enumerate. These changes include rising numbers of new immigrants, both those who are legally in the country and those who are not, who may be less willing to fill out a census form or who may not be able to complete a form because of language difficulties; increasing amounts of junk mail, which may increase the likelihood that households will discard their census form without opening it;[1] and larger numbers of households with multiple residences, making it unclear which form they should mail back.

The Bureau budgeted for a decline in the mail response rate (mail returns as a percentage of all mailback addresses) to 61 percent in 2000, compared with a 70 percent budgeted rate in 1990. As the 2000 census got under way, the Census Bureau director initiated the '90 Plus Five campaign, challenging state, local, and tribal government leaders to reach a mail response rate at least 5 percent higher than their 1990 mark. The Internet was used to post each jurisdiction's goal and actual response rates day-by-day through April 11, 2000.

Maintaining the budgeted mail response rate was key to the Bureau's ability to complete nonresponse follow-up on time and within budget, and, if the rate improved over the budgeted figure, that would be beneficial. Estimates produced in conjunction with the 1990 census were that each 1 percentage point decline in the mail response rate increased the nonresponse follow-up workload and costs of that census by 0.67 percent (National Research Council, 1995b:48). In addition, evidence from the 1990 census, confirmed by analysis of 2000 data, indicated that returns obtained in the field, on balance, were less complete in coverage and content than mail returns (see Appendix D and Chapter 7).

Success in the 2000 Mail Response

A significant achievement of the 2000 census was that it halted the historical decline in the mail response rate (see Box 4.1). The rate achieved at the time when the workload for nonresponse follow-up was identified (April 18, 2000) was 64 percent—3 percentage points *higher* than the 61 percent rate for which the Bureau had budgeted follow-up operations. By the end of 2000 the mail response rate was 67 percent (Stackhouse and Brady, 2003a:2, Table 4). A significant proportion of the mail returns received after April 18, 2000, had arrived by April 30, so that the Census Bureau could perhaps have saved even more follow-up costs if it had set a later date to establish the nonresponse follow-up workload. However, the Bureau was not

[1] A post-1990 census study of household behavior found that 4 percent of households that reported receiving a questionnaire in the mail discarded it without opening it (Kulka et al., 1991).

Box 4.1 Mail Response and Return Rates

The *mail response rate* is defined as the number of households returning a questionnaire by mail divided by the total number of questionnaires sent out in mailback areas. Achieving a high mail response rate is important for the cost and efficiency of the census because every returned questionnaire is one less household for an enumerator to follow up in the field.

The *mail return rate* is defined as the number of households returning a questionnaire by mail divided by the number of occupied housing units that were sent questionnaires in the mailback areas. This rate is an indicator of public cooperation. Achieving a high mail return rate (at least to the level of 1990) is important because of evidence that mail returns are more complete than enumerator-obtained returns.

In 2000, because of the alternative modes by which households could fill out their forms, the numerator of both "mail" responses and "mail" returns included responses submitted on the Internet, over the telephone, and on "Be Counted" forms. The denominator of the mail response rate included all addresses on the April 1, 2000, version of the MAF, covering both mailout/mailback and update/leave areas. The denominator of the mail return rate excluded addresses on the MAF that field follow-up determined were vacant, nonresidential, or nonexistent and included addresses for occupied units that were added to the MAF after April 1.

Final Rates, 1970-2000 Censuses (as of end of census year)

	1970	1980	1990	2000
Mail response rate:	78%	75%	65%	67%
Mail return rate:	87%	81%	75%	78%

Differences in Final Mail Return Rates: Short and Long Forms

Return rates of long forms are typically below the return rates of short forms. This difference widened substantially in 2000.

	1970	1980	1990	2000
Short-form rate:	88%	82%	76%	80%
Long-form rate:	86%	80%	71%	71%
Difference	2%	2%	5%	9%

NOTES: 1980 and 1990 rates shown here differ slightly from those in Bureau of the Census (1995b:1-24). Mail response and return rates are not strictly comparable across censuses because of differences in procedures used to compile the address list and in the percentage of the population included in the mailback universe (about 60 percent in 1970 and 95 percent or more in 1980–2000).

SOURCE: National Research Council (1995b:Table 3.1, App. A) for 1970 and 1980 rates; Stackhouse and Brady (2003a,b:v) for 1990 and 2000 rates.

willing to risk possible delays in completing nonresponse follow-up by delaying the start of the effort (U.S. General Accounting Office, 2002a).

In 1990, the mail response rate at the time when the workload for nonresponse follow-up was identified (late April) was 63 percent—1 percentage point lower than the corresponding 2000 rate and 7 percentage points lower than the 70 percent rate for which the Bureau had budgeted follow-up operations. By the end of 1990 the mail response rate was 65 percent, 2 percentage points lower than the final 2000 rate (Stackhouse and Brady, 2003a:2).

Most of the improvement in mail response in 2000 was because of improved response to the short form. At the time when the nonresponse follow-up workload was specified (April 18, 2000), the long-form mail response rate was 12 percentage points below the rate for short forms (54 percent and 66 percent, respectively). When late mail returns are included, the gap between short- and long-form mail response rates was reduced from 12 to 10 percentage points (Stackhouse and Brady, 2003a:17,18).

The 2000 census was also successful in stemming the historical decline in the mail return rate (mail returns as a percentage of occupied mailback addresses), which is a more refined measure of public cooperation than the mail response rate (see Box 4.1). At the end of 2000, the final mail return rate was 78 percent, compared with a final mail return rate of 75 percent in 1990 (Stackhouse and Brady, 2003b:v). Again, most of the improvement in mail return rates in 2000 was because of improved response to the short form: the final mail return rate for short forms was 80 percent in 2000 compared with 76 percent in 1990. For long forms, the final mail return rate was 71 percent in both years, so that the gap between short-form and long-form mail return rates in 2000 was 9 percentage points, compared with only 5 percentage points in 1990.

Mail Return Patterns

A tabular analysis of all 2000 mail returns (Stackhouse and Brady, 2003b:Tables 8, 10, 12, 14, 16) found considerable variation in mail return rates for population groups:

- Total mail return rates increased as a function of age of householder: the final mail return rate for householders ages 18–24 was 57 percent, climbing to a rate of 89 percent for householders age 65 and older.

- White householders had the highest total final mail return rate of any race group—82 percent. Householders of other race groups had final mail return rates of 75 percent or lower; the lowest rates were for black householders (64 percent), householders of some other race (63 percent), and householders of two or more races (63 percent).

- Hispanic householders had a lower total final mail return rate (69 percent) than non-Hispanic householders (79 percent); Hispanic householders had a larger difference between short-form and long-form return rates (71 percent short-form rate, 57 percent long-form rate) than did non-Hispanic householders (81 percent short-form rate, 72 percent long-form rate).

- Total final mail return rates were similar for most household size categories, except that two-person households had a higher rate (82 percent) than other categories and households with seven or more members had a lower rate (72 percent) than other categories, largely because of differences in long-form return rates for these two household types (76 percent for two-person households and 58 percent for households with seven or more members).

- Owners had a higher total final mail return rate (85 percent) than renters (66 percent).

Exploratory regression analysis conducted by the panel of preliminary total mail return rates for 2000 and 1990 for census tracts found a strong positive relationship in both years of a tract's return rate to its 1990 percentage population over age 65 and a strong negative effect in both years of a tract's return rate to its 1990 hard-to-count score, estimated percentage net undercount, percentage people in multiunit structures, and percentage people who were not high school graduates.[2] Geographic region effects, which differed somewhat between 1990 and 2000, were also evident. While analysis found no variables that explained changes in mail return rates

[2]See National Research Council (2001a:App.B). The data sets used were U.S. Census Bureau, Return Rate Summary File (U.S.), provided to the panel February 26, 2001, and 1990 Planning Database (U.S. Census Bureau, 1999a).

for census tracts from 1990 to 2000, graphical displays identified clusters of tracts with unusually large increases and decreases in return rates between the two censuses, suggesting that local factors may have had an effect in those areas. Thus, large clusters of tracts that experienced 20 percent or greater declines in mail return rates from 1990 to 2000 were found in central Indiana; Brooklyn, New York; and throughout Kentucky, Tennessee, and the Carolinas. At the other extreme, tracts that experienced increases of 20 percent or more in mail return rates were concentrated in the Pacific census division (particularly around Los Angeles and the extended San Francisco Bay Area) and also in New England. In contrast, mail return rates for tracts in the Plains and Mountain states were very similar between 1990 and 2000.

Further investigation of the characteristics of clusters of tracts that experienced unusually large increases or decreases in mail return rates and of local operations and outreach activities in those areas would be useful to identify possible problems and successes to consider for 2010 census planning. More generally, research is needed on patterns of mail response and the reasons for nonresponse.

4–A.2 Controlled Timing and Execution

The Census Bureau met its overriding goal to produce data of acceptable quality from the 2000 census for congressional reapportionment and redistricting by the statutory deadlines. This outcome may appear unremarkable because the Bureau has never failed to produce basic population counts on time. However, as we saw in Chapter 3, success was by no means a foregone conclusion given the serious problems that hampered planning and preparations for the census, which included lack of agreement on the 2000 design until early 1999.

Not only were the basic 2000 data provided on time, but most of the individual operations to produce the data, such as nonresponse follow-up and completion of edited data files, were completed on or ahead of schedule.[3] Of the 520 census offices (excluding Puerto Rico), 68 completed nonresponse follow-up within 6 or fewer weeks, 290 finished within 7 weeks, 462 finished within 8 weeks, and all

[3]The 2000 Accuracy and Coverage Evaluation operations were also carried out on or ahead of schedule; see Chapter 6.

520 finished within 9 weeks—a week ahead of the planned 10-week schedule for completion (U.S. General Accounting Office, 2002a:6).[4] Few instances occurred in which operations had to be modified in major ways after Census Day, April 1, 2000. One exception was the ad hoc unduplication operation that was mounted in summer 2000 to reduce duplicate enumerations from duplicate addresses in the MAF (see Section 4–E).

In contrast, the 1990 census experienced unexpected problems and delays in executing such key operations as nonresponse follow-up, which took 14 weeks to finish, and the Census Bureau had to return to Congress to obtain additional funding to complete all needed operations. The basic 1990 data were released on time, however, and delivery schedules for more detailed products were similar for 1990 and 2000 (see Bureau of the Census, 1995a:Ch.10; http://www.census.gov/population/www/censusdata/c2kproducts.html [1/10/04]).

4–B MAJOR CONTRIBUTORS TO PUBLIC RESPONSE

The Census Bureau had three strategies to encourage response to the 2000 census. Two of the three strategies very likely contributed to the success in halting the historical decline in mail response and return rates—a redesigned questionnaire and mailing package (4–B.1) and extensive advertising and outreach (4–B.2). The third strategy, which we do not discuss further, was to allow multiple modes for response. This strategy had little effect partly because the Census Bureau did not actively promote alternative response modes. Specifically, the Bureau did not widely advertise the option of responding via the Internet (an option available only to respondents who had been mailed the census short form) because of a concern that it could not handle a large volume of responses. The Bureau also did not widely advertise the option of picking up a "Be Counted" form at a public place because of a concern about a possible increase in the number of responses that would require address verification and unduplication. Finally, telephone assistance was designed pri-

[4] Subsequently, every housing unit in the workload of one district office, in Hialeah, Florida, was reenumerated because of problems that came to light in that office, and selected housing units were reenumerated in seven other offices for which problems were identified (see Appendix C.3.b).

marily to answer callers' questions and only secondarily to record responses. As it turned out, of 76 million questionnaires that were returned by households, 99 percent arrived by mail and only 1 percent by other modes (see Appendix C.2.b). In particular, only about 600,000 Be Counted and 200,000 telephone questionnaires were obtained. At the end of a process of geocoding, field verification, and unduplication, the Be Counted and telephone operations added about 100,000 housing units that would otherwise not have been included in the census count (see Vitrano et al., 2003a:29).

4–B.1 Redesigned Questionnaire and Mailing Package

Strategy

Based on its extensive research in the early 1990s (see Section 3–B.2), the Census Bureau redesigned the census short-form and long-form questionnaires and mailing materials for 2000 as part of its effort to encourage the public to fill out questionnaires and return them in the mail (or by an alternative response mode). The questionnaires were made easier to read and follow; the respondent burden of the short form was reduced by providing room for characteristics for six household members instead of seven as in 1990 and by moving questions to the long form; an advance letter was sent out in mailout/mailback areas; the envelope with the questionnaire carried a clear message that response was required by law; and a reminder postcard (also used in 1990) was sent out in mailout/mailback and update/leave areas.

Evaluation

The success of the 2000 census in maintaining mail response and return rates at 1990 levels and above was of major importance for its success in terms of timely, cost-effective completion of operations. It seems highly likely that the changes to the questionnaire and mailing package and the use of an advance letter—despite or perhaps even because of the publicity due to an addressing error in the letter (see Appendix C.2.a)—contributed to maintaining the response and return rates, although how large a role each of these elements played in this achievement is not known. It is also unknown whether the redesigned questionnaire and mailing package had greater effect on some population groups and geographic areas than others, or whether the effect was to boost return rates about equally across population groups and geographic areas and thereby maintain differences in response rates from past censuses.

One disappointment of the mailing strategy and other initiatives to encourage response in 2000 was that they had much less effect on households that received the long form than on those that received the short form. As noted above, at the time when the nonresponse follow-up workload was specified (April 18, 2000), the long-form mail response rate was 12 percentage points below the rate for short forms. This difference was double the difference that the Bureau expected (U.S. General Accounting Office, 2000c:5) and far larger than differences between long-form and short-form response rates seen in previous censuses. One of the Bureau's questionnaire experiments in the early 1990s, using an appealing form and multiple mailings, presaged this outcome: it found an 11 percentage point difference between short-form and long-form response rates (Treat, 1993), perhaps because the user-friendly long form tested was longer than the 1990 long form (28 versus 20 pages). The 2000 long-form had about the same number of questions as the 1990 long form but twice as many pages.

When late mail returns are included, the gap between short- and long-form mail response and return rates in 2000 was somewhat reduced. In other words, some households that received the long form lagged in completing and mailing it back in. However, the wide gap at the time of nonresponse follow-up was important because it meant that a higher proportion of the workload comprised long forms, which experience in 1990 and 2000 demonstrated were more difficult to complete in the field than short forms.

A drawback of the 2000 census mailing strategy was that the plan to mail a second questionnaire to nonresponding households had to be discarded, even though dress rehearsal results suggested that doing so might have increased mail response by 8 percentage points above what was achieved (Dimitri, 1999), or even by 10–11 percentage points as estimated from earlier tests (see Section 3–B.2). The Census Bureau did not work with vendors early in the decade to develop a feasible system for producing replacement questionnaires. Consequently, its plans for a targeted second mailing were stymied when the vendors selected for the dress rehearsal were not prepared to turn around the address list on the schedule required. Experience in the dress rehearsal suggested that mailing a second questionnaire to every address (instead of only to nonresponding households) would generate adverse publicity and increase the number of duplicate returns that would need to be weeded out from the census count, so no second mailing was made in 2000.

4–B.2 Paid Advertising and Expanded Outreach

Expansion of Efforts in 2000

A second important element of the Census Bureau's strategy in 2000 to reverse the historical decline in mail response rates and to encourage nonrespondents to cooperate with follow-up enumerators was to advertise much more extensively and expand local outreach and partnership programs well beyond what was done in the 1990 census. An integral part of the advertising strategy was to pay for advertisements instead of securing them on a pro bono basis. The advertising ran from November 1, 1999, to June 5, 2000, and included separate phases to alert people to the importance of the upcoming census, encourage them to fill out the forms when delivered, and motivate people who had not returned a form to cooperate with the follow-up enumerators. Advertisements were placed on TV (including a spot during the 2000 Super Bowl), radio, newspapers, and other media, using multiple languages. Using information from market research, the advertisements stressed the benefits to people and their communities from the census, such as better targeting of government funds to needy areas for schools, day care, and other services.

The Census Bureau also hired 560 full-time-equivalent partnership and outreach specialists in local census offices (three times the number hired in 1990—see U.S. General Accounting Office, 2001a:11), who worked with community and public interest groups to develop special initiatives to encourage participation in the census. The Bureau signed partnership agreements with over 100,000 organizations, including federal agencies, state and local governments, business firms, and nonprofit groups. (Estimates vary as to the number of partnerships—see Westat [2002c], which reports on a survey of a sample of about 16,000 partners on their activities, receipt of materials from the Census Bureau, and assessment of the materials' effectiveness.)[5] Finally, the Bureau developed a special program to put materials on the census in local schools to inform schoolchildren about the benefits of the census and motivate them to encourage their adult relatives to participate.

The advertising campaign appeared very visible and appealing, and, similarly, partnerships with local communities for outreach were more extensive and seemed more vigorous than in 1990. Correspondingly, costs of advertising and outreach quadrupled between 1990 and 2000, increasing from 88 cents per household in 1990 to $3.19 per household in 2000 (in constant fiscal 2000 dollars; U.S. General Accounting Office, 2002a:11).

Evaluation

We view it as likely that both the advertising and the local outreach efforts contributed to maintaining—even improving—the mail response and return rates in 2000 compared with 1990. However, linking the advertising campaign, much less specific advertisements or community-based communications, to individual behavior—and measuring the magnitude of the effects—is typically very difficult in market research, and the census is no exception. Similarly, the program to put materials about the census in local schools probably contributed to awareness of the census and may have contributed to participation, but evaluation cannot determine that (Macro International, 2002).

Three studies, one commissioned from the National Opinion Research Center (NORC), a second commissioned from the University of Michigan, and a third carried out by Census Bureau staff using data collected by InterSurvey, Inc. (now known as Knowledge Networks), shed light on the effects of advertising and outreach in 2000.

NORC Study The NORC study of the 2000 census marketing and partnership efforts (Wolter et al., 2002) obtained data from a total of

[5] See also U.S. General Accounting Office (2001a:17), which notes that a database developed by the Census Bureau to track the activities of partnership programs had problems keeping up-to-date records, which reduced its usefulness for management and subsequent evaluation.

10,000 households interviewed by telephone in three separate cross-sectional waves in fall 1999, winter 2000, and spring 2000. Each wave's sample comprised a population or core sample (about one-half of the total) and samples of American Indians, Asians, and Native Hawaiians. Under the most favorable calculation, the response rate to wave 1 was only 48 percent because it used a random-digit dialing telephone design with no field follow-up. Response rates for waves 2 and 3, which used samples from the MAF, were somewhat better, at 65 and 68 percent, respectively (Wolter et al., 2002:Table B-1, alternate response rate #4).

The study found that overall awareness of communications about the census increased significantly over time and was greater after the marketing program than before for the total sample and for six race/ethnicity groups analyzed separately. Examining types of communications, most populations recalled TV advertisements in greater proportions than they did magazine or other types of advertisements, and most were more aware of census job announcements, signs or posters, and informal conversations than they were of other types of community-based communications. Among Asian groups, English speakers were more likely to be aware of census communications than were non-English speakers.

With the exception of the American Indian population, the NORC survey found significant associations between awareness of the census and the development of more positive beliefs about the census prior to Census Day (e.g., "it lets government know what my community needs"). There was little evidence that census beliefs shifted after Census Day. Four race/ethnicity groups—non-Hispanic whites, blacks, Asians, and Native Hawaiians (but not Hispanics or American Indians)—became more likely to say they would return the census form after the marketing campaign began. Higher individual awareness of communications about the census also correlated with a greater intention of returning the questionnaire for each of the six groups except American Indians, for which small sample size may have been a problem for analysis.

However, a cross-sectional logistic regression analysis of households interviewed in waves 2 and 3 of the NORC survey and matched with their census returns found limited evidence of an association between awareness of census communications and the likelihood that households would actually complete and mail back a census form instead of having to be followed up by an enumerator. (The study was not designed to support causal inference.) In a model developed for respondents in the wave 2 core sample, there were no significant main or interaction effects on mailback propensities of indexes of awareness of mass media and community-based communications. In the model developed for respondents in the wave 3 core sample, who were interviewed toward the very end of the publicity and outreach campaign, there were a few significant interaction effects. Specifically, an index of awareness of mass media communications increased mailback propensities for race/ethnicity groups in the "other" category in the regression model (non-Hispanic American Indians, Asians, Native Hawaiians, and all others) compared with non-Hispanic whites. An index of awareness of community-based communications increased mailback propensities for people ages 55 and older compared with younger people; it also increased mailback propensities for blacks compared with non-Hispanic whites (Wolter et al., 2002:Table 97).

Comparison of the NORC survey results for 2000 with an Outreach Evaluation Survey conducted for 1990 is difficult. It appears that the 2000 marketing and partnership effort was more effective in raising awareness of the census than the 1990 effort; also, the 2000 program appeared more effective than the 1990 program in creating favorable attitudes that the census cannot be used against you. Respondents in both years exhibited generally similar positive levels toward the importance of participating in the census.

Michigan Study The Michigan study obtained data from Surveys of Privacy Attitudes conducted of about 1,700 households in July through October 1999 and of about 2,000 households in April through July 2000. (The Gallup Organization collected the data by telephone using random-digit dialing.) Response rates to the two surveys were 62 and 61 percent, respectively (Singer et al., 2001:14).[6]

The results showed greater awareness of census data uses on the part of respondents in 2000 than in 1999 and, over the same time period, a significant decline in the belief that census data were likely

[6]The two Michigan surveys were similar to Surveys of Privacy Attitudes conducted in 1995 and 1996 (see Singer et al., 2001, for results from all four surveys).

to be misused for law enforcement—changes the analysts attributed to the effects of the advertising and outreach programs.

Responses to a question in the 2000 survey on exposure to publicity about the census broke down as follows: 30 percent of the sample reported no exposure to publicity; 27 percent reported exposure to positive publicity only (e.g., the importance of being counted); 20 percent reported exposure to negative publicity only (e.g., concerns about privacy or confidentiality and also hearing controversy in the media about answering the long form, as discussed in the next section); and 22 percent reported exposure to both positive and negative publicity. Compared with the group who reported no exposure to publicity (referred to as the "control group" in subsequent analysis), the other three groups were significantly more likely to be knowledgeable of the census. Those exposed to positive publicity only were also significantly more likely to report a variety of positive attitudes about the census than the control group. Those exposed to both positive and negative publicity generally had more positive attitudes toward the census than the control group, but they were also significantly more likely than the control group to consider the census an invasion of privacy. Those exposed to negative publicity only were similar to the control group except for a greater awareness of the census (Singer et al., 2001:71).

As in the NORC study, respondents to the 2000 Survey of Privacy Attitudes were matched to their census forms, although a smaller-than-expected number of cases were successfully matched. (The analysts attempted to compensate for this and other biases in the matched sample.) Similar to the NORC study, a logistic regression model of matched cases in the 2000 Michigan survey showed no direct relationship of exposure to publicity and the propensity to complete and mail back the census form—indeed, the only group that was significantly more likely than the control group to mail back a form was the group exposed to negative publicity.

However, positive publicity may have had indirect effects on response in that exposure to such publicity generally led to more favorable attitudes toward the census, which in turn increased mail-back propensities. In particular, those who believed that census data are not misused for any of three enforcement purposes (identifying illegal aliens, keeping track of troublemakers, and using census answers against respondents) mailed back their census forms at a rate of 86 percent. In contrast, those who believed that census data are misused for all three of these purposes mailed back their census forms at a rate of only 74 percent. The NORC study found a smaller difference that was not significant: respondents to the wave 2 NORC survey who had mildly or highly favorable attitudes toward the census mailed back their forms at a rate of about 80 percent compared with a rate of 73 percent for those with unfavorable attitudes toward the census; the comparable percentages for respondents to the wave 3 NORC survey were 75 and 68 percent, respectively (Wolter et al., 2002:Table 79). The NORC study did not use an index of census attitudes in logistic regression models.

InterSurvey Analysis Martin (2001) analyzed five cross-sectional surveys conducted by InterSurvey, Inc., over WebTV between March 3 and April 13, 2000, with sponsorship from private foundations. The data are subject to substantial nonresponse and sample selection biases, but they help fill in a picture of the effects of hearing negative publicity about the census long form. Several prominent politicians (including then presidential candidate George W. Bush and then Senate majority leader Trent Lott) and radio talk show hosts commented on the intrusiveness of the long form and were widely quoted in the press in late March, about the time when census forms were being mailed to the public.[7] Martin found that mistrust in government, receipt of a census long form, and hearing about the long-form controversy were strongly associated with respondents' level of privacy concerns about the census. In turn, Martin found that respondents who received a long form or who had privacy concerns were less likely to return their census form or to fill it out completely if they did return it. The Michigan study (discussed above) reported similar results from the 2000 Survey of Privacy Attitudes regarding the negative effects of receipt of a long form and of a high score on an index of general privacy concerns on mailback propensities (Singer et al., 2001:80).

[7]See, e.g., "Census 2000 too nosey? Republicans criticize long-form questions," posted on CNN.com, March 31, 2000; http://www.cnn.com/2000/US/03/31/census.01/. See also "Nosy Census?," transcript of March 30, 2000 *NewsHour with Jim Lehrer*, posted at http://www.pbs.org/newshour/bb/fedagencies/jan-june00/census_3-30.html.

4–B.3 Mail Response: Summary of Findings and Recommendations

The evidence on the effectiveness of the Census Bureau's two major strategies to improve mail response to the census in 2000—redesigned questionnaire and mailing materials, and paid advertising and extensive outreach—is not definitive. In particular, there appear to have been few direct effects of advertising and outreach on mailback propensities, although there is evidence of indirect effects from favorable publicity inducing more favorable attitudes toward the census that, in turn, stimulated response. There is also evidence of positive effects of mass media or community-based communications for a few population groups, such as older people and blacks. Based on the higher response and return rates achieved in 2000 compared with 1990, however, it seems reasonable to conclude that the two strategies were successful, particularly the redesign of the questionnaire and mailing materials and particularly for the short form. Response to the long form may have been adversely affected by negative publicity, as well as by such factors as the length of the questionnaire.[8]

> **Finding 4.1:** The use of a redesigned questionnaire and mailing strategy and, to a more limited extent, of expanded advertising and outreach—major innovations in the 2000 census—contributed to the success achieved by the Census Bureau in stemming the decline in mail response rates observed in the two previous censuses. This success helped reduce the costs and time of follow-up activities.

Looking to ways to further enhance mail response to the 2010 census, we note the strong base of evidence for the positive effects of mailing a second questionnaire to nonresponding households. Preliminary results from a 2003 test survey corroborate the positive effects found in earlier research (Angueira, 2003). Because handling a replacement questionnaire on the scale of a census presents logistical challenges, it is essential for the Census Bureau to begin work

[8]See Edwards and Wilson (2003), which summarizes findings from the NORC survey, the InterSurvey analysis, and process evaluations of the advertising and outreach programs, and reaches similar conclusions.

immediately on ways to surmount those challenges so that nonresponding households receive a second questionnaire on a timely basis and in a manner that minimizes the likelihood of duplicate enumerations.

> *Recommendation 4.1:* **The Census Bureau must proceed quickly to work with vendors to determine cost-effective, timely ways to mail a second questionnaire to nonresponding households in the 2010 census, in order to improve mail response rates, in a manner that minimizes duplicate enumerations.**

The Census Bureau also included in its 2003 test survey an emphasis on alternative response modes, including the Internet and telephone using interactive voice response. If widely accepted, such methods could lead to significant savings in costs and time for follow-up and data capture in 2010, even if there is no change in the overall response rate. Whether these methods would increase overall returns is not as clear, as they may mainly attract respondents who would fill out and mail back their questionnaires in any case.

Although mail response and return rate patterns have been extensively analyzed, there is much that is not known about factors that explain differences in return rates among population groups and geographic areas and over time. The Census Bureau now has available substantial data on local area characteristics from the 1990 and 2000 censuses that could be used to analyze geographic clusters of high and low return rates with a goal of suggesting testable innovations that might improve return rates in 2010, particularly for low-response areas. The Bureau also has extensive operational and characteristics data on a sample of 2000 census enumerations—a compilation known as the Master Trace Sample—that could be used to analyze demographic and socioeconomic return rate patterns with the goal of suggesting testable strategies for improving return rates for low-response groups. Such research should be pursued (see also Chapter 9).

4–C MAJOR CONTRIBUTORS TO TIMELINESS

4–C.1 Contracting for Operations

Strategy

A major innovation for the 2000 census was the use of outside contractors for key operations. One outside vendor was contracted to develop the systems for checking-in questionnaires and putting the data into computerized form (data capture). Another contractor was hired to manage data capture at three processing centers; the Census Bureau's long-established National Processing Center at Jeffersonville, Indiana, was the fourth data capture center. Also, outside vendors were used to provide telephone questionnaire assistance and to carry out telephone follow-up for questionnaires that were identified as possibly incomplete in coverage (e.g., households that reported more members than the number for which they answered individual questions—see Appendix C.5). The Census Bureau rarely used contractors for major operations in censuses prior to 2000, preferring a strategy of performing all operations in house (with some technical assistance from contractors), so the decision to contract for such large important operations as data capture marked a significant departure from past practice.

Evaluation

Information with which to assess the performance of specific contractors on a range of dimensions is not currently available, except insofar as performance can be inferred from such outcomes as timely and accurate data capture. With regard to data capture (see Section 4–C.2), the system development contractor was commended by the U.S. General Accounting Office (2000c:16–18) for following good practices for software development in the context of the work to reconfigure the system to separate the completion of data capture for long-form-sample items from the basic items.

4–C.2 Improved Data Capture Technology

Strategy

The Census Bureau early on made a decision to use new data capture technology to replace its in-house FOSDIC system (Film Optical

Sensing Device for Input to Computers). First developed for the 1960 census, FOSDIC was improved at each successive census through 1990 (see Salvo, 2000; Bureau of the Census, 1995b:Ch.8). FOSDIC basically involved microfilming census questionnaires and using optical mark recognition (OMR) software to read light spots representing respondents' filled-in answer circles and write the output onto computer files. Write-in answers (e.g., name, race, occupation, industry, place of work) were captured by having clerks key the data from the paper forms into specified fields shown on computer screens (for keying write-in race entries, the clerks worked from microfilm readers).

For 2000 the Bureau decided to contract for the adaptation of commercial optimal mark recognition and optical character recognition (OMR/OCR) technology in which questionnaires would be scanned directly into the computer and read by the OMR/OCR software. Clerks would supplement the process by keying data from computer images when the OCR technology could not make sense of the responses.

The new technology performed well enough in a 1995 test to be used in the 1998 dress rehearsal and well enough in the dress rehearsal to be recommended for use in 2000 (see Conklin, 2003:1). There remained, however, a considerable element of risk in the decision to use OMR/OCR in the 2000 census because of limited testing on the necessary scale prior to late 1999. Indeed, during testing that was conducted in early 2000, some problems were identified in the accuracy of the optical mark/character recognition, and changes were made to improve the accuracy rate and reduce the number of items that had to be keyed or rekeyed from images by clerks (U.S. General Accounting Office, 2000b). In addition, the large number of different census forms and delays in finalizing questionnaires and making prototypes available to the OMR/OCR contractor made development of the technology and the associated flow management software more difficult and costly (Brinson and Fowler, 2003).

Finally, the data capture system was redesigned at the last minute to postpone capture of long-form content that required keying by clerks. Both short forms and long forms were run through the OMR/OCR scanning technology and keying was performed, when necessary, for the basic items. However, keying of additional long-form responses that the automated technology could not process was performed in a second, separate operation. This change was made on the basis of operational tests of keying from images, which demonstrated that keying could not occur fast enough to process the additional long-form items and keep to the schedule for producing data for reapportionment and redistricting (U.S. General Accounting Office, 2000d).

Evaluation for Timeliness

During actual production, the new technology performed on schedule; there were no delays that affected field or other operations. Top census officials have attested that the decision to complete the long-form data capture process at a later stage was essential to timely completion of the basic census counts. Without the ample funding that was ultimately appropriated for the census (see Section 3–C.4), it would have been difficult to make this decision because of the increased costs needed to handle the long-form questionnaires twice.[9]

Evaluation for Accuracy

An important issue is the accuracy of the new technology in capturing responses. Citing a study by the Rochester Institute of Technology Research Corporation (2002), an overall assessment of the 2000 data capture system concluded that the system—which processed over 152 million questionnaires—exceeded specified performance goals for both mail and enumerator forms (Titan Corporation, 2003:11). Conklin (2003) provides detailed data on error rates for the new technology. In this study, the images and data capture information from the 2000 census processing were obtained for 1.5 million forms (divided equally among short and long forms, including English mailed-out/mailed-back forms, Spanish mailed-out/ mailed-back forms, English update/leave/mailed-back forms, and English enumerator-obtained forms, plus four form types for Puerto Rico). Clerks then reprocessed the information by keying all of the responses from the computer images of the questionnaires. The 2000 data capture system was evaluated against the results of the independent rekeying operation to determine how well the system captured the intended response as determined by the clerks. For

[9]Kenneth Prewitt, former Census Bureau director, personal communication.

check-box responses, the content captured had to agree exactly with the clerical determination of the intended content to count as correct; for write-in responses, the agreement did not have to be letter for letter—a matching algorithm determined when the correspondence was within specified limits.

Overall findings of the evaluation (Conklin, 2003:6) include:

- Optical mark recognition had an error rate of 1.2 to 1.5 percent (97 percent confidence interval) for all check-box responses that the technology considered readable;

- Optical character recognition had an error rate of 1.0 to 1.1 percent (97 percent confidence interval) for all write-in responses that the technology considered readable (79 percent of such responses); and

- Key from image (KFI) had an error rate of 4.8 to 5.3 percent (97 percent confidence interval) for the responses that the OMR or OCR technology rejected as unclear.

No data are available for comparing the accuracy of the 2000 data capture technology with the 1990 technology, but a study for the dress rehearsal (cited in Conklin, 2003) gave a performance standard for errors using the 1990 system of 2 percent. The OMR and OCR error rates in 2000 were well below that rate (1.3 percent and 1.1 percent, respectively).

The Conklin (2003) analysis for 2000 examined OMR, OCR, and KFI error rates by type of form, data capture center, and regional census center. Excluding forms for Puerto Rico, key findings include:

- Averaged across type of technology (OMR, OCR, KFI), error rates for short-form content grouped into five categories (name, demographic, race, ethnicity, housing) were 3 percent or lower for all categories for all types of forms in the evaluation, with the exception of name, for which error rates were 4 percent or more on three of the eight form types: enumerator-obtained short forms (4.2 percent median error rate), Spanish mailed-out/mailed-back long forms (4.6 percent median error rate), and Spanish mailed-out/mailed-back short forms (7.1 percent median error rate).

- Averaged across type of technology, mailed-back forms had slightly higher error rates than enumerator-obtained forms for some important items, such as race and ethnicity, perhaps because enumerators were instructed on how to write answers to facilitate accurate data capture.

- Averaged across type of technology, error rates for long-form content grouped into five categories (occupation, income, education, military, disability) were 3 percent or lower for all categories for all types of long forms in the evaluation, with the exception of the military items on enumerator-obtained long forms, which had a median error rate of 3.4 percent.

- There were no significant differences in error rates across the four data capture centers. This finding is important because it indicates that data capture site was not a source of geographic variation in the quality of the census data.

- Forms originating from regional census center areas with above-average concentrations of immigrants had high error rates for name fields.

- Analysis of individual items (defined as a specific response for a specific person line number on a specific form—2,996 in all) identified 150 person-items that had high error rates ranging from 8 percent to 91 percent. These 150 person-items appeared on at least 500 records in the sample; person-items with high error rates based on smaller sample sizes were excluded, as were person-items with high error rates on the Puerto Rico forms (Conklin, 2003:Table 8). Conklin (2003) recommends that these items be the subject of further investigation to improve data capture technology or questionnaire design or both for 2010.

The overall assessment of the 2000 data capture system in Titan Corporation (2003) has many recommendations for improvements in the system for 2010. They include changing the questionnaires to facilitate automated data capture, better and more timely specification of contractual requirements for data capture systems and management, and integrating the development of the 2010 data capture system with questionnaire design and printing (e.g., choosing a background color that makes it easier to distinguish responses than the color used in 2000).

4–C.3 Aggressive Recruitment of Enumerators and Implementation of Follow-Up

Critical to the success of the 2000 census was the ability to field a timely, complete follow-up of nonresponding households to enumerate them, to determine that the address was a vacant unit (and obtain some information about the type of housing), or to determine that the address should not have been included in the MAF (because the structure was commercial, had been demolished, or another reason). Nonresponse follow-up was a major problem in the 1990 census because the mail response rate not only dropped below the rate in 1980, it also dropped 7 percentage points below the budgeted rate. The Bureau had to seek additional funding, scramble to hire enough enumerators, and take much more time to complete the effort than planned.

Recruitment Strategy

In 2000, fears of a tight labor market that could make it difficult to hire short-term staff led the Bureau to plan aggressive recruitment of field staff from the outset. Generous funding made it possible for the Bureau to implement its plans, which included directing local offices to recruit twice as many enumerators as they expected to need; this funding also allowed the Bureau to offer part-time work schedules and above-minimum wages (which differed according to prevailing area wages).

Pressures to Expedite Follow-Up

The Census Bureau not only encouraged local offices to hire more enumerators than they expected to need, it also encouraged them to expedite nonresponse follow-up efforts. The U.S. General Accounting Office (2002a:23) documented that the Bureau "developed ambitious interim stretch goals. These goals called on local census offices to finish 80 percent of their nonresponse follow-up workload within the first 4 weeks of the operation and be completely finished by the eighth week," even though the master schedule allowed 10 weeks for completion.

Evaluation for Timeliness

The Bureau's recruitment strategy and imposition of stretch goals for nonresponse follow-up were very successful in terms of timely completion of the workload. Every local office attracted at least three applicants for each enumerator position to be filled, and about 80 percent of the offices achieved their recruiting goal, which was to hire twice as many enumerators as they were likely to need. Pay for enumerators that exceeded locally prevailing wages together with effective regional office management were strong determinants of local office recruiting performance (Hough and Borsa, 2003:8,9, citing Westat, 2002b). A midstream assessment of nonresponse follow-up concluded that it was going well in most offices (U.S. General Accounting Office, 2000c). A large percentage of offices finished within the stretch-goal target of 8 weeks, and all offices finished within 9 weeks—a week ahead of schedule. The timely completion of nonresponse follow-up was a major achievement of the 2000 census, which an evaluation commissioned by the Census Bureau attributed primarily to the success in recruiting sufficient numbers of qualified applicants and retaining enumerators as long as they were needed (Hough and Borsa, 2003:12, citing Westat, 2002a).

Evaluation for Accuracy of the Population Count

One might expect that the success in completing nonresponse follow-up ahead of schedule, and, similarly, in fielding a more focused coverage improvement follow-up effort than in 1990 (see Appendix C.3), would contribute to fewer duplicates and other types of erroneous enumerations in 2000 than in 1990. In 1990, questionnaires for occupied housing units with later check-in dates (the date of entering the Census Bureau's processing system) were more likely to include erroneous enumerations than were returns checked in earlier. Specifically, the percentage of erroneous enumerations increased from 2.8 percent for questionnaires checked in through April 1990 (largely mail returns), to 6.6 percent, 13.8 percent, 18.8 percent, and 28.4 percent, respectively, for those checked in during May, June, July, and August or later (largely, enumerator-obtained returns).[10]

Although the correlation between timing of receipt and accurate coverage of household members on a questionnaire may be spurious, there are several plausible reasons to support such a relationship. For example, people who moved between Census Day and follow-up could well be double-counted—at both their Census Day residence and their new residence (e.g., people in transit from a southern winter residence to a northern summer residence or college students in transit between home and dormitory around spring or summer vacation). More generally, the later a household was enumerated, the less accurately the respondent might have described the household membership as of Census Day.

Given the delays in nonresponse follow-up in 1990, it appears that as much as 28 percent of the workload was completed after June 6, when erroneous enumeration rates were 14 percent or higher. Almost 10 percent of the nonresponse follow-up workload was completed in July or August. (These workload percentages do not include coverage improvement follow-up.) We have only limited information on the relationship of erroneous enumerations to the timing of enumeration in the 2000 census. Estimates from the 2000 E-sample indicate that nonresponse follow-up enumerations overall had a higher percentage of erroneous enumerations (15.5 percent) than did mail returns overall (5.3 percent).[11] Also, in the A.C.E. Revision II estimation, late enumerator returns (defined as those received after June 1) had higher estimated erroneous enumeration rates than did early enumerator returns (see U.S. Census Bureau, 2003c:Table 6). Furthermore, 90 percent of the nonresponse follow-up workload was completed by June 6 and 98 percent by June 20 (Moul, 2003:App.G). Some returns for occupied units (2.4 million) were obtained through coverage improvement follow-up (which began at the end of June), but these were less than 10 percent of the total

[10]These estimates were derived from the 1990 Post-Enumeration Survey—see Ericksen et al. (1991:Table 2).

[11]Tabulations by panel staff from U.S. Census Bureau, E-Sample Person Dual-System Estimation Output File, provided to the panel February 16, 2001; weighted using the median value of TESFINWT within households. TESFINWT is the final weight, adjusted for the targeted extended search, for the E-sample cases that were included in the dual-systems estimation (see Appendix E.5).

number of returns obtained for occupied units in nonresponse and coverage improvement follow-up (see Moul, 2002, 2003). Hence, although we cannot be sure, it is possible that the speedier completion of nonresponse follow-up in 2000 helped reduce erroneous enumerations.

Intensive analysis of the Accuracy and Coverage Evaluation (A.C.E.) data demonstrated a large number of duplicate enumerations in 2000 (see Chapter 6), which could have been partly due to errors in follow-up processes. Research on the effects of both nonresponse follow-up and coverage improvement follow-up on correct versus erroneous enumerations would be useful.

Evaluation for Effects on Differential Coverage of Population Groups

One evaluation of both the nonresponse follow-up and coverage improvement follow-up efforts in 2000 found that enumerator returns included higher percentages of traditionally hard-to-count groups (children, renters, minorities) compared with mail returns (Moul, 2003:Tables 18–22). Therefore, as was also true in 1990, these operations were important for helping to narrow differences in net undercount rates among groups. Thus, for household members enumerated in 2000:

- 45 percent of nonresponse and coverage improvement enumerations were of people living in rented units, compared with 25 percent renters on mail returns;
- 51 percent of nonresponse and coverage improvement enumerations were of men, compared with 48 percent men on mail returns;
- 59 and 57 percent of nonresponse and coverage improvement enumerations, respectively, were of people under age 35, compared with 45 percent people under age 35 on mail returns;
- 18 percent of nonresponse and coverage improvement enumerations were of Hispanics, compared with 12 percent Hispanics on mail returns;
- 18 and 17 percent of nonresponse and coverage improvement enumerations, respectively, were of blacks, compared with 10 percent blacks on mail returns.

Evaluation for Completeness of Data Content

The U.S. General Accounting Office (2002a) conducted a study to determine whether the pressure from headquarters on local census offices to speed up the completion of nonresponse follow-up may have resulted in higher percentages of incomplete enumerations. The GAO examined data on when local offices finished their workload to see if early closers had higher percentages of "partial" or "closeout" cases (as recorded by enumerators on the questionnaires). Partial interviews were defined as having less than the minimum amount of information for a complete interview but at least housing unit status and, for occupied units, the number of residents (see Moul, 2003:3 for the requirements for a complete interview). Closeout interviews were those obtained once a crew leader's assignment area within a local census office reached a 95 percent completion rate at which time enumerators made a final visit to each missing address to try to obtain a complete interview, or, at a minimum, the unit status and number of residents.

The GAO analysis found no relationship between the week of completion and the percentage of partial or closeout interviews. Nor did the analysis find a relationship between the week of completion and the percentage of residual workload,[12] or between week-to-week spikes or surges in production and the percentage of closeout or partial interviews.

These findings suggest that the pace of the follow-up work did not reduce the quality of the data. However, it would be useful to supplement this analysis with a study of the possible relationship between the speed of completion and the percentage of proxy interviews conducted with a neighbor or landlord and between the speed of completion and missing data rates, particularly for long-form items.

4–C.4 Timeliness: Summary of Findings

The Census Bureau employed a variety of innovative strategies in 2000 to facilitate the timely execution of the census, including not

[12]The residual workload consisted of addresses in the original nonresponse follow-up workload for which questionnaires were not processed through data capture; local offices had to conduct residual nonresponse follow-up on these cases (122,000 of an original workload of 42 million) at a later date.

only strategies to improve mail response (see Section 4–B), but also strategies to facilitate data capture and ensure an adequate work force for nonresponse follow-up, as discussed in Section 4–C. A concern with an aggressive strategy for completion of nonresponse follow-up is that it could have led to higher rates of missing and erroneous data. The evidence to date suggests that the use of new data capture procedures and technology and aggressive goals for enumerator recruitment and work completion were important innovations that had positive effects on timeliness while not impairing data quality.

> **Finding 4.2:** Contracting for selected data operations, using improved technology for capturing the data on the questionnaires, and aggressively recruiting enumerators and implementing nonresponse follow-up were significant innovations in the 2000 census that contributed to the timely execution of the census.

4–D TIMELINESS VERSUS COMPLETENESS: GREATER RELIANCE ON COMPUTERS TO TREAT MISSING DATA

4–D.1 Strategy

The 2000 census used computers whenever possible to replace tasks that were previously performed in part by clerks or enumerators. Notably, questionnaires went directly to one of the four processing centers for data capture instead of being reviewed first by clerks in local census offices, as occurred for much of the workload in 1990. Editing and imputation of individual records to supply values for missing responses to specific questions or reconcile inconsistent answers were handled entirely by computer; there was no clerical review or effort to telephone or revisit households to obtain more content information as occurred in 1990 and previous censuses. Mail returns for households with more than six members and some other returns that appeared not to have not filled out the basic information for one or more members were followed up by telephone to collect the information for the missing members. However, in contrast to 1990, there was no attempt to collect missing items for already enumerated household members, and there was no field follow-up when telephone follow-up was unsuccessful (see Appendix C.5).

After completion of all follow-up procedures, computer routines were used as in previous censuses, not only to impute responses for individual missing items but also to supply census records for household members that were missing all basic characteristics and for whole households when not even household size was known for the address. These imputation routines used records from neighboring households or people who matched as closely as possible whatever information was available for the household or individual requiring imputation (see Appendixes F, G, and H). The advantages expected from greater computerization of data processing included savings in cost and time to complete the data records. Also, it was expected that computer systems for editing and imputation would be better controlled and less error-prone than clerical operations.

4–D.2 Evaluation

Evaluation of Computer Data Processing Systems for Timeliness

The 2000 census computer systems for data processing appear to have worked well. Although delays in developing specific systems occurred because of the delays in determining the final census design, completion of software at the last minute appears to have had little adverse effect on the timing of other operations. Software development problems did delay the implementation of the coverage edit and telephone follow-up operation by a month (the operation began in late May instead of late April—see Appendix C.5.b). Moreover, the requirements for data processing operations were not often fully specified in advance, and much of the software was developed without adequate testing and quality assurance procedures, which put many data processing steps at risk (see Alberti, 2003:32–33; U.S. General Accounting Office, 2000a; U.S. Department of Commerce, Office of Inspector General, 1997). Fortunately, there were no actual breakdowns in the performance of critical software systems, although errors occurred in specific routines, such as the processing of information on forms obtained from enumerators about occupancy status and write-in entries for race, which are discussed below.

Evaluation of Imputation Procedures for Accuracy of the Population Count

The 2000 census, like previous censuses, included records for people who were wholly imputed because they lacked even basic information. In some cases, every household member lacked basic information, and the household required "whole-household imputation" or "substitution" (the Census Bureau's term), which was performed by duplicating the record of a nearby household.

Box 4.2, under subheadings "B" and "C," defines the five types of situations in which whole-person imputation (as distinct from individual item imputation, "A") may be required in the census (see also Appendix G). Table 4.1 shows the number and percentage of the population of each of these types of imputations for the 1980, 1990, and 2000 censuses. In total, wholly imputed people represented a small percentage of the population in each year: 1.6 percent (3.5 million people), 0.9 percent (2 million people), and 2.2 percent (5.8 million people), respectively, in 1980, 1990, and 2000. The larger numbers of wholly imputed people of all types in 2000 compared with 1990 help explain part of the reduction in net undercount rates between children and adults, renters and owners, and minorities and non-Hispanic whites. If these people had not been imputed into the census, then there would have been a net undercount of about 1.7 percent overall (0.5 percent estimated net overcount minus 2.2 percent whole-person imputations), and differences in net undercount rates among population groups would have been wider (see Section 6–C.1).

The largest percentage of wholly imputed people in 2000 were those in mail return households that did not provide characteristics for all of their members (type 1). These people were genuine household enumerations for which the computer was used to impute their basic characteristics, item by item, on the basis of knowing a great deal about the household. The much larger number of type 1 imputations in 2000 compared with 1990 and 1980 was a direct outcome of the decision in 2000 to reduce the space for reporting household member characteristics on the questionnaire and to follow up households with more members than room to report them by telephone only, with no field follow-up as occurred in 1990. Type 1 imputations were not required for large households enumerated

Box 4.2 Imputation Types for Basic (Complete-Count) Characteristics

Following data capture from census questionnaires and the various levels of nonresponse follow-up, the Census Bureau uses editing and imputation procedures to fill in apparent gaps. The level of imputation required depends on the number of household members who are *data defined*—that is, whose census records have at least two basic data items reported (counting name as an item). In descending order of known information—and ascending order of required imputation—the various types of imputation performed in the census are:

(A) *Item imputation.* All members of a household are data defined, but some basic items are not reported or are reported inconsistently; these missing values are supplied through "hot-deck" imputation (termed *allocation* by the Census Bureau), or through procedures called assignments or edits. Broadly speaking, edit and assignment procedures make use of other information provided by the person; imputation procedures make use of information from other household members or a similar individual in a similar, nearby household.

(B) *Whole-person imputation.* At least one member of a household is data defined as in (A), but not all members are so defined.

1. *Individual person(s) imputed in an enumerated household.* For the members of the household who are not data defined, all basic information is imputed or assigned, item-by-item, on the basis of information about the other household members. An example in 2000 would be a household of seven members that had data reported for six members, and the telephone follow-up failed to obtain information for the seventh person on the household roster (the mail questionnaire allowed room to report characteristics for only six members instead of seven as in 1990). This type 1 imputation is called *whole-person allocation* by the Census Bureau.

(C) *Whole-household imputation* There is no data-defined person at the address. Imputation is performed using information from a similar, nearby household or address. Collectively, types 2–5 below are termed *whole-household substitution* by the Census Bureau.

2. *Persons imputed in a household for which the number of residents is known* (perhaps from a neighbor or landlord), but no characteristics are available for them.
3. *Persons imputed in a housing unit known to be occupied* for which there is no information on household size.
4. *Persons imputed in a housing unit for which occupancy status and household size have to be imputed first* (from among housing units for which occupancy or vacancy status is not known).
5. *Persons imputed in a housing unit for which housing unit status and occupancy status have to be imputed first* (from among addresses for which not even status as a housing unit is known).

Types 3–5 were the focus of legal action by the state of Utah. In June 2002, the U.S. Supreme Court declined to characterize such imputation as "sampling," and hence permitted its use to contribute to state population counts for congressional reapportionment (see Box 2.2).

Table 4.1 People Requiring Imputation of All Basic Characteristics by Type of Imputation, 2000, 1990, and 1980 Censuses

	2000[a]	1990[b]	1980[b]
As Percentage of Household Population:			
Whole-Person Imputations in Enumerated Households (type 1)	0.90	0.20	0.10
Whole-Person Imputations in Wholly Imputed Households			
Characteristics (type 2)	0.83	0.64	1.17
Count, Occupancy Status, Housing Status	0.43	0.02	0.34
Count (type 3)	0.18	—	—
Occupancy Status (type 4)	0.10	—	—
Housing Status (type 5)	0.15	—	—[c]
Subtotal, types 2–5	1.26	0.66	1.51
Total, types 1–5	2.16	0.86	1.61
Number of Persons (millions):			
Whole-Person Imputations in Enumerated Households (type 1)	2.330	0.373	0.152
Whole-Person Imputations in Wholly Imputed Households			
Characteristics (type 2)	2.270	1.547	2.580
Count, Occupancy Status, Housing Status	1.171	0.054	0.761
Count (type 3)	0.496	—	—
Occupancy Status (type 4)	0.260	—	—
Housing Status (type 5)	0.415	—	—[c]
Subtotal, types 2–5	3.441	1.601	3.341
Total, types 1–5	5.771	1.974	3.493

NOTES: See Box 4.2 for definitions of imputation types. —, not available separately.

[a] Tabulations by panel staff of U.S. Census Bureau, File of Census Imputations by Poststratum, provided to the panel July 30, 2001 (Schindler, 2001).

[b] Calculated from Love and Dalzell (2001). Whole-person imputations in enumerated households include a small number of whole-person imputations for group quarters residents.

[c] Housing status imputation (type 5) was not used in 1980.

in person because of the use of continuation forms (see Section 3–B.2). However, in list/enumerate areas, because of a data processing error, the continuation forms for large households were lost, and the imputation process may not have imputed all of the people on the lost forms back into the census (Rosenthal, 2003b:6).

The numbers of people requiring imputation in households for which only occupancy status and household size were reported (type 2) were roughly similar in all three censuses. The larger number of wholly imputed people in households of types 3–5 (for which, at a minimum, household size had to be imputed) in 2000 compared with 1990 and even 1980 is difficult to explain, although processing errors may have contributed to this outcome (see end of this section). Information is not available with which to evaluate the reasonableness of the numbers of type 4 and 5 whole-household imputations as percentages of the numbers of addresses for which imputations could have been possible; that is, the Census Bureau has not provided the denominators with which to calculate imputation rates for these two categories.

Whole-person and whole-household imputations varied by geographic area. In particular, the small number of people imputed in 2000 when it was not even clear whether the address was a housing unit (type 5) were concentrated in rural list/enumerate areas (e.g., the Adirondacks region of New York State, rural New Mexico).[13] In these areas, enumerators developed an address list and enumerated the units at the same time. Although there was a follow-up operation to recheck the status of units that the enumerators classified as vacant (Hough and Borsa, 2003:21), the status of some units was apparently not resolved.

A question is how many of the imputed people in households for which a household size had to be imputed (types 3–5) represented correct enumerations—some of them may have been erroneous in that the household did not contain as many people as were imputed into it. Alternatively, not enough people may have been imputed into some households. A related question is whether the basic characteristics assigned to imputed people in all of the imputation types accurately reflected the true distributions.

[13]From analysis by panel staff of U.S. Census Bureau, Census Tract Imputation File, provided to the panel April 4, 2002.

There are limited pieces of evidence that suggest some problems in the imputation of whole persons. An administrative records experiment conducted in five counties in which national records were matched with 2000 census addresses found that, when the linked census record was not imputed, 51 percent of the matches agreed on household size, compared with only 32 percent agreement when the linked census record was a whole-household imputation. Moreover, while the household size discrepancies involving links of administrative records to nonimputed census households were symmetric, those involving links to imputed census households were asymmetric. Thus, 41 percent of imputed census households were larger in size than the linked administrative households, while 27 percent of imputed census households were smaller in size than the linked administrative households (Bauder and Judson, 2003:24). What fraction of the discrepancies were due to errors in the imputations or in the administrative records is not known.

In addition, Alberti (2003) documents two data processing problems that resulted in larger numbers of households requiring occupancy status imputation (type 4) and housing status imputation (type 5). First, because of an error in processing enumerator forms, about 145,000 housing units were classified as occupancy status unknown, when they were all most probably vacant at the time of the census and none of them should have been imputed as occupied (Alberti, 2003:34–35). These units represented 74 percent of all of the units eligible for occupancy imputation (type 4 imputation; Alberti, 2003:26). Second, because over 2.3 million group quarters responses that indicated a usual residence elsewhere were erroneously sent to the process for verifying these and other newly identified addresses (see Section 4–F.2), the verification process encountered delays. As a consequence, the data for 207,000 housing units were not included in the census and their status as a housing unit (or as a record to be deleted from the census) had to be imputed. These units represented 70 percent of all of the units eligible for housing status imputation (type 5 imputation; Alberti, 2003:33–34).

Evaluation of Imputation Procedures for Accuracy of Basic Data

Response to the basic data items for enumerated (data-defined) people on the short and long forms—age, sex, race, ethnicity, household relationship, and housing tenure—was good in 2000. Rates of missing data for individual basic items were low (1–4 percent), as was also generally the case in 1990 (although some population groups and geographic areas had higher basic item imputation rates—see Sections 7–B and 8–C.2). Moreover, it was often possible to infer a missing value from other information on the household's own questionnaire instead of having to use information from neighboring households. The decision to capture names for all enumerations helped in this regard, as names could be used in assigning sex and, in some instances, ethnicity. The Bureau made additional improvements in its editing and imputation routines for missing and inconsistent basic items (see Appendix G), although an error in the processing of write-in entries for race led to an overestimate of people reporting more than one race (see Chapter 8). Assuming that the Bureau otherwise maintained good quality control of the basic data editing and imputation specifications and implementation in 2000, the use of computer routines to provide values for specific missing items should have had little adverse effect on the quality of the basic items. The resulting data products are more complete and therefore useful for a broader range of purposes.

Evaluation of Imputation Procedures for Accuracy of Additional Long-Form Data

In contrast to the responses for basic items, and of considerable concern, is that missing data rates for the additional long-form-sample population and housing items were high in many cases and generally higher than the comparable rates in 1990 (see Section 7–C and Appendix H). The Census Bureau relied for imputation of these items on procedures that it has used for many censuses with little evaluation of their appropriateness or effectiveness. We recommend research on imputation procedures in Chapter 7 (see also Appendix F). Below we recommend tests of the effects on response accuracy of the 2000 strategy of relying entirely on computerized imputation for missing values compared with the intensive telephone and field follow-up efforts used in 1990 to reduce the extent of missing data. A third alternative of following up in the field a sample of households with missing data should also be investigated, as it could balance considerations of cost and timing and the need for accuracy better than the two extremes.

4–D.3 Reliance on Imputation: Summary of Findings and Recommendations

A major design strategy for the 2000 census was to limit as much as possible costly, time-consuming follow-up (particularly in-person follow-up) for missing data and instead to rely more heavily than in 1990 on computer-based imputation. Two considerations initially led to the adoption of this strategy. The first was the desire to contain census costs. The second was the desire to allow enough time in the schedule for completion of all coverage evaluation activities so that the state population totals for reapportionment could incorporate the results of a large independent survey under the originally proposed Integrated Coverage Measurement design (ICM—see Section 3–C.2). Even after the ICM design was discarded, the Census Bureau retained its plan to limit follow-up in favor of more extensive use of imputation.

In contrast, the 1990 census design emphasized the role of repeated telephone and field follow-up to reduce not only the extent of missing content but also the number of addresses for which no information was obtained about occupancy status or the number and characteristics of residents. In this way, the Census Bureau hoped to defuse possible arguments on whether the use of whole-household imputation constituted statistical adjustment (see Sections 3–A.1 and 6–C.1).

A concern about greater reliance on imputation in place of follow-up efforts is that the imputed data would contain more error than the information that could have been obtained by recontacting the household. To date there is little evidence on this point. Given the higher rates of imputation in 2000 than in 1990, particularly for individual people and for many long-form items, the Census Bureau should conduct experiments to test the relative costs and accuracy of more imputation versus more follow-up before deciding whether to continue the 2000 strategy in 2010 (see also Appendix F on imputation methods). Assessing relative accuracy is not easy because of the problem of establishing a reference data source that can serve as a "gold standard." A possible method for tackling this problem is to use multiple data sources, including household surveys, administrative records, and content reinterviews (see Chapter 9). Such research should investigate the cost-effectiveness of field follow-up of a sample of households and people with missing data. The data from such a follow-up effort should permit the development of imputation models that much more accurately reproduce the characteristics of nonrespondents, compared with models that are limited (as in 2000) to using respondents' characteristics for imputation to nonrespondents.

Finding 4.3: The greater reliance on imputation routines to supply values for missing and inconsistent responses in 2000, in contrast to the greater reliance on telephone and field follow-up of nonrespondents in 1990, contributed to the timely completion of the 2000 census and to containing the costs of follow-up. It is not known whether the distributions of characteristics and the relationships among characteristics that resulted from imputation (particularly of long-form content) were less accurate than the distributions and relationships that would have resulted from additional follow-up.

Recommendation 4.2: **Because the 2000 census experienced high rates of whole-household nonresponse and missing responses for individual long-form items, the Census Bureau's planning for the 2010 census and the American Community Survey should include research on the trade-offs in costs and accuracy between imputation and additional field work for missing data. Such research should examine the usefulness of following up a sample of households with missing data to obtain information with which to improve the accuracy of imputation routines.**

4–E MASTER ADDRESS FILE DEVELOPMENT: FLAWED EXECUTION

4–E.1 Strategy and Implementation

With the bulk of the population enumerated by mailout/mailback or update/leave/mailback techniques, the quality of the 2000 address list was essential to the completeness and accuracy of population coverage. The Census Bureau early in the 1990s made a decision to partner with other organizations and use multiple sources to develop the MAF. Contributing to the 2000 census MAF (termed the Decennial MAF or DMAF by the Census Bureau) were the 1990 address list augmented by updates from the U.S. Postal Service (in mailout/mailback areas), a full block canvass by Census Bureau staff, input from localities that participated in the Local Update of Census Addresses (LUCA) Program, and census field operations (see Vitrano et al., 2003a:3–5).

The goal of using multiple sources to build as complete a list as possible was a worthy one. However, because many of the procedures were new, implementation was not always smooth:

- The decision to conduct a complete (instead of targeted) block canvass in urban areas in 1998 was made late in the decade and required substantial additional funding to implement. The change became necessary when the Census Bureau learned that the U.S. Postal Service Delivery Sequence File (DSF) was not accurate enough or sufficiently up to date to serve as the primary means of updating the 1990 census address list (see Section 3–B.1)

- The decision to provide localities in city-style-address areas an opportunity to add addresses for units newly constructed in January–March 2000 was made even later, in March 1999.

- Original plans for a sequential series of steps in the LUCA Program, involving back-and-forth checking with localities, had to be combined under pressures of time, and many LUCA components experienced delays by the Census Bureau; see Table 4.2. Except for the stage of appealing to the U.S. Office of Management and Budget, localities were not given additional time for their review.

- Questionnaire labeling had to occur before the Bureau had the opportunity to check most of the addresses supplied by LUCA participants, increasing the opportunity for erroneous enumerations, such as duplications.

- Integration of the address list for special places (group quarters) with the housing unit address list was delayed, and, con-

Table 4.2 Original and Actual Timelines for the Local Update of Census Addresses (LUCA) Program

Planned Dates	Actual Dates	Activity
LUCA98[a]		
November 1997	February 1998	Census Bureau sent invitation letters to eligible localities
April–August 1998	May 1998–March 1999	Bureau sent initial materials (address list, maps) to participants
May–December 1998	May 1998–June 1999	LUCA participants conducted initial review of materials from Bureau
Not part of original plan	January–May 1999	Bureau conducted full (instead of targeted) block canvassing
May–October 1999	July–December 1999	Bureau verified participants' addresses in the field (reconciliation) (original plan was to send results to localities to obtain feedback before sending final determination materials)
March–November 1999	October 1999–February 2000	Bureau sent detailed feedback/final determination materials to participants
Original deadline, January 14	November 1999–April 3, 2000	LUCA participants filed appeals (addresses were visited in coverage improvement follow-up)
April–August 1998	December 1999–April 2000	Special Places LUCA (Bureau did not complete Special Places list until November 1999)
Added operation	January–March 2000	Participating localities submitted new construction addresses (addresses were visited in coverage improvement follow-up)
LUCA99[b]		
July 1998–February 1999	July 1998–February 1999	Census Bureau field staff listed addresses
September–October 1998	September–October 1998	Bureau sent invitation letters to eligible localities
January–April 1999	January–August 1999	Bureau sent initial materials (block counts, maps) to participants
January–April 1999	January–October 1999	LUCA participants conducted review of initial materials from Bureau
March–May 1999	May–October 1999	Bureau verified participants' block counts in the field (reconciliation)
March–June 1999	September 1999–February 2000	Bureau sent detailed feedback/final determination materials to participants
Original deadline, January 14	October 1999–April 3, 2000	LUCA participants filed appeals for specific addresses
January–April 1999	December 1999–April 2000	Special Places LUCA

[a] This program was conducted in areas with mostly city-style addresses.
[b] This program was conducted in areas with mostly rural route and post office box addresses.

SOURCE: Adapted from Working Group on LUCA (2001:Figure 1).

sequently, local review of the special places address list was delayed and often did not occur.

- Significant numbers of errors in assigning individual group quarters to geographic areas occurred, perhaps because they were coded to the address of the special place (e.g., a university) and not to the location of the specific group quarters.

- In some mailout/mailback areas Postal Service workers returned several million questionnaires as undeliverable because the address label used a city-style street name and house number when the Postal Service delivery address was to a post office box. The Census Bureau had to arrange for enumerators to redeliver the questionnaires to these addresses, which should have been included in the update/leave operation. They were instead included in the mailout/mailback operation because of inappropriate designation of the "blue line" separating mailout/mailback from update/leave areas (see Appendix C.1.a). Inaccuracies in designating the blue line also caused problems for LUCA participants (Working Group on LUCA, 2001:Ch. 4).

The Bureau recognized early on that the MAF was at risk of including duplicate and other erroneous addresses. The risk of omitting valid addresses was also present, but MAF procedures were expected to reduce the level of omissions from previous censuses. An increased risk of including duplicate addresses in the 2000 MAF resulted not only from the planned use of multiple sources but also from the operational problems just reviewed. To minimize duplication, the Bureau used a combination of field checking and internal consistency checks of the MAF file (see Appendix C.1.d). One computer check was performed prior to nonresponse follow-up; another check—not included in the original plans—was performed in summer 2000.

The special summer unduplication operation resulted from evaluations of the MAF conducted by Population Division staff between January and June 2000, in which MAF housing unit counts were compared to estimates prepared from such sources as building permits. The results led the Census Bureau to conclude that there were probably still a sizable number of duplicate housing unit addresses on the MAF despite prior computer checks. Field verification carried out in June 2000 in a small number of localities substantiated this conclusion.

Consequently, the Bureau mounted an ad hoc operation to identify duplicate MAF addresses and associated census returns. Housing unit and person records flagged as likely duplicates were deleted from the census file and further examined. After examination, it was determined that a portion of the deleted records were likely to be separate housing units not already included in the census, and they were restored to the census file. At the conclusion of the operation, 1.4 million housing units and 3.6 million people were permanently deleted from the census; 1 million housing units and 2.4 million people were reinstated (see Table 4.3, which charts additions to and deletions from the MAF due to census operations in 2000).

4–E.2 Overall Assessment of MAF Development

We assess the MAF development process as sound in concept but flawed in execution. On the plus side, all elements of the MAF development process were completed despite the various implementation problems noted. Moreover, delays in component MAF operations did not appear to affect other census operations, such as mailout, field follow-up, and data processing.[14]

With regard to accuracy, however, errors detected in the MAF required an ad hoc unduplication operation at the height of census processing—a laudable effort, but one that subsequent evaluation determined to contain some flaws (Mule, 2002a). Of the 1.4 million housing unit records deleted from the census, 0.3 million were *not* in fact duplicates; in contrast, of the 1 million potential duplicate housing unit records that were reinstated, 0.7 million were in fact duplicates that should not have been reinstated.[15] On net, the unduplication operation failed to delete an additional 0.4 million duplicate housing unit records. However, some of the reinstated duplicates were retained to compensate for omissions in small mul-

[14]However, the special unduplication effort complicated estimation for the Accuracy and Coverage Evaluation Program, as reinstated people could not be included in the A.C.E. matching process (see Section 6–A.4).

[15]This estimate is derived from estimates by Mule (2002a) of people who were erroneously reinstated from the special unduplication operation.

Table 4.3 Additions to and Deletions from the 2000 MAF from Major Census Operations in 2000

Addresses Added from Field Operations	
During questionnaire delivery (update/leave, update/enumerate, list/enumerate areas)	2.3 million
During nonresponse and coverage improvement follow-up	1.7 million
Total addresses added from listed operations	*4.0 million adds*
Addresses Deleted as Nonexistent or Duplicative from Computer Checks Prior to Nonresponse Follow-Up	
Nonexistent deletions (mainly LUCA addresses that had not been verified)	2.5 million
Duplicate deletions (from LUCA and other sources)	1.1 million
Addresses Deleted as Nonexistent or Duplicative from Nonresponse and Coverage Improvement Follow-Up	5.4 million
Addresses Deleted as Duplicative from Ad Hoc Computerized Unduplication Operation in Summer 2000	
Addresses deleted originally	2.4 million
Addresses restored upon further examination	1.0 million
	1.4 million net deletions
Total deletions from listed operations	*10.3 million deletions*
MAF at Completion of Census	115.9 million housing unit addresses

NOTE: The housing unit coverage study (Barrett et al., 2001, 2003) estimated additional duplicates and other erroneous enumerations, as well as omissions, in the MAF; see text.

SOURCE: Farber (2001b:Tables 1, 2); Miskura (2000a); Nash (2000); Mule (2002a); for details of operations, see Appendix C.1.

tiunit structures that lacked individual apartment addresses. For example, an apartment designated as unit "A" on the MAF by the Census Bureau may have picked up and completed the questionnaire for unit "B," while the "B" household failed to respond. In follow-up the Census Bureau would return to apartment "A" and obtain a second interview. Reinstating the second questionnaire for "A" would effectively result in a whole-household imputation for "B" (see Hogan, 2000b).

Duplicates and other erroneous housing unit addresses in the MAF that were not involved in the special unduplication operation were probably present in the MAF, and the MAF probably omitted some valid addresses as well. A housing unit coverage study, which developed dual-systems estimates for housing units from the A.C.E. data (Barrett et al., 2001:Table 2), estimated that 1.5 percent of occupied housing units in the A.C.E. E-sample of census enumerations (excluding reinstated records) were erroneously enumerated (about 1.6 million units, weighted up to national totals) and that 2.6 percent of occupied housing units in the independent A.C.E. P-sample were omitted from the MAF (about 2.7 million units, weighted). The dual-systems estimate of occupied units, when compared with the census count (including reinstated records), produced an estimated net undercount of 0.3 percent, or about 0.4 million net occupied units omitted from the census.[16]

By type of structure, the housing unit coverage study estimated that small multiunit structures (those with two to nine housing units) had greater percentages of erroneous enumerations and omissions than either single-family units or large multiunit structures (Barrett et al., 2001:Table 6; see also Jones, 2003a; Ruhnke, 2003). This finding accords with evidence from the special unduplication study and the experience of LUCA participants.

[16]The estimated net undercount of occupied housing units has not been fully reconciled with the estimated net overcount of people. See Robinson and Wolfgang (2002:i–ii), who find broad agreement among coverage patterns for regions and housing tenure between Revision II A.C.E. and the housing unit coverage study, but differences for some race/ethnicity groups. Revision II A.C.E. also estimates greater reductions in differential coverage errors between 1990 and 2000 than does the housing unit coverage study (which was based on the original A.C.E. and never reestimated).

All of these results are subject to sampling and nonsampling error and should be interpreted cautiously. Nonetheless, assuming that the estimates are roughly accurate, it appears that the 2000 MAF may have had a negligible percent net undercount of occupied units. A small net coverage error, however, masked large numbers of gross errors of omission and erroneous inclusion—as many as 2.3 million erroneously enumerated occupied units (an estimated 0.7 million units that should have been deleted in the special unduplication operation, but were not, and 1.6 million erroneously enumerated units estimated from the A.C.E.), and as many as 2.7 million omitted units. The omitted units included not only 1.4 million units that were never on the MAF but also 0.4 million units that were incorrectly dropped from the MAF before Census Day, 0.6 million units that were incorrectly deleted from the MAF on the basis of such operations as nonresponse follow-up, and an estimated 0.3 million units that were incorrectly deleted from the MAF in the special unduplication operation (Vitrano et al., 2003b:Table 33). Without the special unduplication operation, even with its problems, the gross errors in the MAF would have been larger yet.[17]

Among the erroneously enumerated occupied housing units identified in the Housing Unit Coverage Study were census units in A.C.E. block clusters that were incorrectly geocoded to a block in a ring of surrounding blocks; such geocoding errors were estimated as 0.4 percent of total housing units (Barrett et al., 2001:Table 10). A study that looked for geocoding errors for census housing units in individual A.C.E. blocks by searching the entire tract and one ring of surrounding census tracts estimated that 4.8 percent of total housing units were misgeocoded (Ruhnke, 2003:iv). Most or all of these geocoding errors, however, would not contribute to gross errors for larger geographic areas (e.g., towns, cities, counties, states).

We do not know whether the errors in the MAF contributed more or less to population coverage errors than did omissions and erroneous inclusions of people within correctly identified housing units. We also know little about the variability in the accuracy of the

[17]Even higher percentages of errors of erroneous enumeration and omission occurred for vacant housing units (Barrett et al., 2001:Table 2). See Vitrano et al. (2003b:65–73) for an analysis of gross errors in the MAF.

MAF across population groups and geographic areas that could be due to the LUCA Program and other factors.

4–E.3 Variable Participation in LUCA

All counties, places, and minor civil divisions on the Census Bureau's list of functioning governmental units—39,051 units in all—were eligible to participate in the Local Update of Census Addresses (LUCA) Program; either or both LUCA 98 (conducted in city-style-address areas) or LUCA 99 (conducted in areas with large numbers of rural route and post office box addresses). However, preliminary data show that only 25 percent of eligible units fully participated and that rates of full participation varied across several dimensions (see Table 4.4; Working Group on LUCA, 2001:Ch.2). By full participation, we mean that they informed the Census Bureau of needed changes to the address list for their area; see Box 4.3.

Factors that relate to participation include:

- *Geographic region.* Jurisdictions in some areas of the country participated at higher rates than those in other areas (56 and 37 percent, respectively, of jurisdictions in the Pacific and Mountain states participated, compared with only 18 and 19 percent, respectively, of jurisdictions in the New England and West North Central states).

- *Population size.* Jurisdictions with larger populations participated at higher rates than those with smaller populations (75 percent of those with 1 million or more people participated, compared with only 14 percent of those with 1,000 or fewer people).

- *Type of government.* Places and counties participated at higher rates (37 and 27 percent, respectively) than minor civil divisions (15 percent).

- *Type of program.* Areas eligible for LUCA 98 or both LUCA 98 and LUCA 99 participated at higher rates (42 and 37 percent, respectively) than areas eligible only for LUCA 99 (14 percent).

A multiple regression analysis found that, among counties and places that signed up to participate in LUCA, the estimated 1990

Table 4.4 Participation of Local Governments in the 2000 Local Update of Census Addresses (LUCA) Program

	LUCA 98 Only		LUCA 99 Only		Both LUCA 98 and LUCA 99					
						Percent Participated in				
Category	Number Eligible	Percent Partici-pated	Number Eligible	Percent Partici-pated	Number Eligible	LUCA 98 Only	LUCA 99 Only	LUCA 98 and 99	Total Number Eligible	Percent Participated, One or Both
Total	9,044	41.6	21,760	14.2	8,247	17.7	7.6	12.0	39,051	25.4
Geographic Division										
New England	518	41.5	1,047	6.7	66	12.1	3.0	1.5	1,631	18.1
Middle Atlantic	2,034	43.3	2,133	16.2	662	19.2	7.0	12.7	4,829	30.7
East North Central	3,944	35.0	3,187	10.6	3,483	15.4	5.3	5.1	10,614	24.6
West North Central	548	40.3	9,437	13.3	1,414	17.4	10.7	19.1	11,399	18.8
South Atlantic	697	52.7	1,681	24.1	585	18.1	11.5	21.4	2,963	36.1
East South Central	411	35.5	1,046	11.0	427	14.8	7.3	11.7	1,884	21.5
West South Central	241	37.8	1,914	14.1	886	13.5	9.8	14.0	3,041	22.7
Mountain	113	68.1	966	23.7	328	35.7	6.4	22.3	1,407	36.7
Pacific	538	72.1	349	20.1	396	33.6	9.6	21.0	1,283	55.5
Population Size (1998 est.)										
1,000 or fewer	1,743	26.0	15,100	12.1	1,436	14.2	10.5	6.3	18,279	14.8
1,001–10,000	4,550	40.9	6,080	18.9	4,044	17.8	6.6	11.7	14,674	30.4
10,001–50,000	2,157	50.2	563	21.1	1,827	17.6	8.1	12.6	4,547	41.8
50,001–100,000	364	64.7	17	23.5	444	19.6	7.4	16.9	825	52.7
100,001–1,000,000	217	58.5	0	—	469	25.2	6.4	23.2	686	56.0
1,000,001 or more	13	69.2	0	—	27	25.9	7.4	44.4	40	75.0
Government Type										
County	122	46.7	982	17.1	1,956	10.3	9.3	10.6	3,060	26.7
Minor civil division	3,624	29.8	9,887	7.7	3,082	13.7	4.2	5.2	16,593	15.4
Place	5,298	49.6	10,891	19.8	3,209	25.9	9.9	19.4	19,398	33.8

NOTES: See Box 4.3 for definition of participation in LUCA 98 and 99. Not all regions have minor civil divisions. The analysis excludes tribal governments, the county-level municipios of Puerto Rico, and three places for which 1998 population estimates were unavailable.

SOURCE: Tabulations by panel staff from preliminary U.S. Census Bureau data (LUCA 98 and LUCA 99 spreadsheets, June 2000), modified by assigning county codes to minor civil division and place records and augmenting the file with 1998 population estimates and variables from the Census Bureau's 1990 Data for Census 2000 Planning (1990 Planning Database, on CD-ROM) (see Working Group on LUCA, 2001:Tables 2-2, 2-3).

Box 4.3 Defining Participation in the Local Update of Census Addresses (LUCA) Program

It is not straightforward to determine participation in LUCA. On the preliminary data records available to the Working Group on LUCA (2001:Ch.2), the Census Bureau coded eligible jurisdictions into one of four categories:

- The government did not participate at all (0);
- The government signed a confidentiality agreement, which entitled it to receive materials from the Census Bureau for review (1);
- The government signed a confidentiality agreement and "returned materials" to the Census Bureau (2);
- The government signed an agreement, returned materials, and, in the submission for a LUCA 98 jurisdiction, provided at least one address action record (addition, correction, deletion, or notation that the address was outside the jurisdiction or nonresidential), or, in the submission for a LUCA 99 jurisdiction, challenged the address count for at least one block (3).

In an analysis of the LUCA 98 Program with final Census Bureau data, Owens (2003:v, 9) used a broad definition of participation, under which 53 percent of eligible governments participated by signing the required confidentiality agreement. These are governments in categories 1–3; they include approximately 92 percent of 1990 housing units in eligible areas. However, about one-third of the number "participating" did not return any address changes to the Census Bureau, so that only 36 percent participated fully in LUCA 98 as the Working Group on LUCA defined participation—that is, category 3 only. Some governments in categories 1 and 2 that did not provide address changes may have been satisfied with the MAF for their areas, but, more likely, they did not have time or resources to conduct a full review (see Owens, 2003:14).

In Table 4.4, we define participation in LUCA 98 and LUCA 99 in the same manner as the Working Group on LUCA—that is, category 3 above, excluding categories 1 and 2.

For estimating the housing unit coverage of participating areas, a further complication relates to overlapping areas of jurisdictions. Specifically, for counties participating in LUCA, there is no information on whether they reviewed all of the addresses in the county, only those addresses not covered by the constituent places and minor civil divisions, or only some of the addresses not covered by constituent places and minor civil divisions. Case studies conducted by the Working Group on LUCA (2001:Ch.4) indicated that county participation varied in the extent of address coverage.

Because of ambiguities about the nature and extent of LUCA participation, the panel cautions against using the broader definition in Owens (2003) to estimate the percentage of LUCA participants or the percentage housing unit coverage of participating jurisdictions.

census net undercount rate was a strong predictor that a jurisdiction would participate fully. Case studies also identified instances in which a vigorous coordination effort by a state or regional government facilitated participation by local jurisdictions (Working Group on LUCA, 2001:Ch.4).

The governments that participated in LUCA appeared to cover a higher proportion of the nation's housing stock than the proportion of participating governments to eligible governments would suggest. From preliminary data, places that participated fully in LUCA 98 accounted for 67 percent of the 1990 housing stock in eligible places, even though they included only 48 percent of eligible places. (It is not possible with the available data to construct reliable estimates of housing coverage for all fully participating governments, given such problems as lack of information about the portions of a county covered by county review [see Box 4.3], nor is it possible to construct reliable estimates of housing coverage for the two programs [LUCA 98 and LUCA 99] combined.) Even though coverage was higher for housing than for governments, which would be expected given the greater propensity of larger-size areas to participate, substantial portions of the MAF were not accorded local review.

A recent evaluation of the LUCA 98 program (Owens, 2003) estimated that fully participating jurisdictions (counties, places, minor civil divisions) submitted 6.3 million additional addresses (6.5 percent of their housing stock), of which 5.1 million addresses were not already on the MAF. The census enumeration process identified 3.1 million of these added addresses as valid occupied or vacant housing units, but only 506,000 valid housing unit addresses (0.4 percent of the final census MAF) could definitely be attributed to LUCA 98 only and not also to some other address source. Of the addresses provided uniquely by LUCA 98, 62 percent came from local governments in six states—New York (31 percent, principally New York City); Illinois (9 percent); California (7 percent); Georgia (6 percent); Michigan (5 percent); and Florida (4 percent) (Owens, 2003:App.K).

Problems with determining sources of addresses because of overlapping operations and hard-to-interpret source codes on the MAF records complicate the assessment of the unique contribution of any one source. Consequently, the figure of 506,000 housing units contributed uniquely by LUCA 98 may be an underestimate. Indeed, Vitrano et al. (2003b:Table 13) estimate that 659,000 housing units were contributed uniquely by LUCA 98. The same study reports that another 512,000 housing units were contributed both by LUCA 98 and the block canvass, another 396,000 housing units were contributed by both LUCA 98 and the September 1998 DSF, and 289,000 housing units were contributed uniquely by LUCA 99. A thorough assessment of the contribution of LUCA to MAF is needed, including not only the effects of LUCA on the completeness of the census count in participating areas but also the possible effects on the counts in other areas from not having had a LUCA review.

4–E.4 2000 MAF Development: Summary of Findings

The concept for building the 2000 MAF from many sources, including the previous census address file, was an important innovation that the Census Bureau plans to fully implement for the 2010 census. For 2000 the newness of many MAF-building operations led to significant problems of implementation. In turn, these problems not only contributed to errors in the MAF but also complicated a full evaluation. The variables and codes originally included on the MAF records made it difficult to trace address sources. To facilitate analysis, Census Bureau evaluation staff subsequently developed an "original source" code for each address record and used it to identify sources that contributed the highest numbers of addresses to the final 2000 MAF (Vitrano et al., 2003b). However, they could not always identify a single source or even any source for some addresses. In addition, there are no data with which to determine the contribution of each source to correct versus erroneous addresses on the final MAF. Other problems in the MAF coding, such as overstatement of the number of units at a basic street address in some instances and omission of this information for non-city-style addresses, hampered evaluation (see Vitrano et al., 2003a:5). Finally, very little analysis has been done to date of geographic variations in the effectiveness of various sources for building the 2000 MAF.

> **Finding 4.4:** The use of multiple sources to build a Master Address File—a major innovation in 2000—was appropriate in concept but not well executed. Problems included changes in schedules and operations, variabil-

ity in the efforts to update the MAF among local areas, poor integration of the address list for households and group quarters, and difficulties in determining housing unit addresses in multiunit structures. Changes were made to the MAF development: a determination late in the decade that a costly complete block canvass was needed to complete the MAF in mailout/mailback areas and a determination as late as summer 2000 that an ad hoc operation was needed to weed out duplicate addresses. Problems with the MAF contributed to census enumeration errors, including a large number of duplicates.

Finding 4.5: The problems in developing the 2000 Master Address File underscore the need for a thorough evaluation of the contribution of various sources, such as the U.S. Postal Service Delivery Sequence File and the Local Update of Census Addresses Program, to accuracy of MAF addresses. However, overlapping operations and unplanned changes in operations were not well reflected in the coding of address sources on the MAF, making it difficult to evaluate the contribution of each source to the completeness and accuracy of the MAF.

4–E.5 2010 MAF Development: Recommendation

The Census Bureau plans to update the MAF over the 2000 decade. It has embarked on a major effort to reengineer its TIGER geocoding and mapping database to make it more accurate and easier to maintain.[18] For the TIGER reengineering, the Bureau plans to the extent possible to work with local governments, many of which have geographic information systems that are of better quality than TIGER. In addition, although plans are not yet well developed, the Bureau intends to conduct a LUCA Program as part of constructing the 2010 census MAF. We discuss three areas for improvement in the 2010 MAF development process below.

[18]TIGER—Topologically Integrated Geographic Encoding and Reference System—is used to assign MAF addresses to the proper geographic areas; see National Research Council (2003a) for an assessment of the MAF/TIGER reengineering project.

Housing Unit Addresses

The Bureau must recognize that obtaining an accurate address for a structure (single-family home, apartment building) is not sufficient for a complete, accurate MAF. It is also critical to develop accurate information on the number of housing units that exist within structures. Several reports, including the housing unit coverage study (Barrett et al., 2001, 2003) and the case studies prepared by the Working Group on LUCA (2001:Ch.4), document the problems in developing complete, correct address lists for small multiunit structures, which often have a single mail drop. For example, in neighborhoods with a stock of single-family or two-family residences and heavy inmigration since the last census, additional, hard-to-identify housing units may have been carved out of an apparently stable housing stock. Conversely, some structures with two or three units may have been turned into bigger single-family residences (e.g., as the result of gentrification). Indeed, in areas in which it is hard to keep track of housing units, as opposed to structures, it may be necessary to consider another enumeration procedure, such as update/enumerate, in place of mailout/mailback, in order to obtain an accurate count of households and people.

LUCA Redesign

The Bureau needs to redesign its cooperative programs with state and local governments for MAF/TIGER work. The LUCA effort in the late 1990s was one-sided. The Bureau offered each local government the opportunity to participate but did not actively encourage participation or provide financial or other support beyond training materials. Also, to protect confidentiality, the Bureau required participating localities to destroy all MAF-related materials after the review was completed. Consequently, localities could not benefit from the efforts they made, which were often substantial, to update the MAF. Because of TIGER inaccuracies, localities with better geographic information systems had to force their data to fit the TIGER database before they could begin their local review.

We believe that a successful federal-state cooperative program involving the level of effort of LUCA must be a two-way street in which there are direct benefits not only to the Census Bureau but also to participating localities. The Census Bureau could adapt features of other federal-state cooperative statistical programs toward this end. For example, the Bureau should consider paying for a MAF/ TIGER/LUCA coordinator in each state, who would be a focal point for communication with localities.

The Bureau should also give serious consideration to providing localities with updated MAF files, which would not only facilitate continuous updating of the MAF for the Bureau's purposes but would also provide a useful tool for local planning and analysis. An issue for concern would be that sharing of MAF files might violate the confidentiality of individuals—for example, by disclosing overcrowding of housing units in violation of local codes. However, our view is that the confidentiality issues could be resolved; street addresses do not, of themselves, identify information about individual residents or even indicate whether an address is occupied. For structures that may have been divided into multiple housing units illegally, the Bureau could protect confidentiality through the use of data perturbation or masking techniques. For example, the Bureau could provide a street address for the structure and code the number of units as, say, 1 to 3, or 2 to 4, in a way that preserves confidentiality; alternatively, it could randomly assign a value for the number of units in a structure for a small sample of structure addresses. A further protection would be to provide the MAF to localities under a pledge to use it for statistical purposes only, not for enforcement purposes, and not to provide it to others. However, Title 13 of the U.S. Code would probably require amendment similar to the 1994 legislation that authorized LUCA, since U.S. Supreme Court precedent views the MAF as covered under Title 13 confidentiality provisions.[19] There is national benefit in having an accurate address list for statistical uses that can be continuously maintained in a cost-effective manner. This national benefit should challenge the Census Bureau to think creatively about ways to share the MAF with localities while protecting the confidentiality of individual residents.

[19]In *Baldridge v. Shapiro*, 455 U.S. 345 (1982), the U.S. Supreme Court ruled that the Census Bureau's "address list . . . is part of the raw census data intended by Congress to be protected" under the confidentiality provisions of Title 13. Accordingly, the court concluded that the bureau's address list is not subject to disclosure under the Freedom of Information Act or under the discovery process in civil court proceedings.

Evaluation

The Bureau needs to build a continuous evaluation program into its 2010 MAF/TIGER operations from the outset. That program should specify clear codes for tracing the source(s) of each address to facilitate evaluation. It should also include ongoing quality control and review procedures to determine what is working well and what is not so that corrective actions can be taken in a timely manner. Built into the MAF quality control and evaluation efforts should be a capability to analyze geographic variations in the effectiveness of address sources and in the completeness and accuracy of the list.

> *Recommendation 4.3:* **Because a complete, accurate Master Address File is not only critical for the 2010 census but also important for the 2008 dress rehearsal, the new American Community Survey, and other Census Bureau surveys, the Bureau must develop more effective procedures for updating and correcting the MAF than were used in 2000. Improvements in at least three areas are essential:**
>
> 1. **The Census Bureau must develop procedures for obtaining accurate information to identify housing units within multiunit structures. It is not enough to have an accurate structure address.**
> 2. **To increase the benefit to the Census Bureau from its Local Update of Census Addresses (LUCA) and other partnership programs for development of the MAF and the TIGER geocoding system, the Bureau must redesign the program to benefit state and local governments that participate. In particular, the Bureau should devise ways to provide updated MAF files to participating governments for statistical uses and should consider funding a MAF/TIGER/LUCA coordinator position in each state government.**
> 3. **To support adequate assessment of the MAF for the 2010 census, the Census Bureau must plan evaluations well in advance so that the MAF records can be assigned appropriate ad-**

dress source codes and other useful variables for evaluation.

4–F GROUP QUARTERS ENUMERATION

4–F.1 Strategy

The population that resides in group quarters (e.g., college dormitories, prisons, nursing homes, juvenile institutions, long-term care hospitals and schools, military quarters, group homes, shelters, worker dormitories), and not in individual apartments or homes, is a small fraction of the total population—about 2.7 percent in 1990 and 2000.[20] Individual group quarters often contain large numbers of people, however, so that group quarters residents can represent a significant component of the population of particular small areas (e.g., college students living in dormitories in a college town). The census is the only source of data on group quarters residents in addition to household members, so enumeration of this population, even though small, is important.

Mailout/mailback techniques and the questionnaires used for household enumeration are not appropriate for many group quarters, so there was a separate operational plan for group quarters enumeration in 2000, as in censuses since 1970. This plan used procedures similar to those for previous censuses, with a few exceptions (see Citro, 2000c; U.S. Census Bureau, 1999b). Group quarters residents were asked to fill out individual (one-person) questionnaires called Individual Census Reports; somewhat different questionnaires were used for enumeration of four groups: armed forces personnel in group quarters, military and civilian shipboard residents, people enumerated at soup kitchens and mobile food vans (part of service-based enumeration), and all other group quarters residents.

There was no effort to enumerate the homeless population not encountered at designated locations, which included some nonshel-

[20]In Census Bureau terminology, one or more group quarters make up a special place. Such places (e.g., a university) are administrative units, the individual group quarters (e.g., dormitories) are where people sleep. A structure that houses a group quarters may also include one or more housing units (e.g., the apartment for a resident faculty member in a dormitory).

tered outdoor locations in addition to shelters, soup kitchens, and regularly scheduled mobile food vans. The original design for 2000 included a plan to estimate the total homeless population by using information for people enumerated at shelters and soup kitchens on how many nights each of them used the service. The actual responses to questions on service usage exhibited high nonresponse and response bias, so plans were dropped to include adjusted counts for the homeless in any adjusted population totals (Griffin and Malec, 2001; see also U.S. General Accounting Office, 2003b). Such adjusted counts could not be used for reapportionment totals, in any case, because of the U.S. Supreme Court decision in January 1999 that precluded the use of sampling for census counts for reapportionment.

4–F.2 Implementation Problems

For reasons that are not clear, the design and implementation of enumeration procedures for group quarters residents experienced numerous problems (see Jonas, 2002, 2003b):

- There was no definition of a group quarters as distinct from a large household of unrelated individuals. In 1980 and 1990 any residential unit with 10 or more unrelated people was tabulated as a group quarters; no such rule was set for 2000. Moreover, some group quarters (e.g., halfway-homes) are located in structures that appear to be housing units. Such ambiguities contributed to problems in developing an accurate, nonoverlapping MAF for housing unit and group quarters addresses.

- The development of an inventory of special places was handled by the Census Bureau's Population Division staff, and special places addresses were not integrated with the main MAF (maintained by the Bureau's Geography Division) until November 1999. Consequently, the Special Places LUCA operation was delayed by 18 months from the original schedule, and many localities could not participate given demands on their resources for census outreach and other activities close to Census Day.

- A number of group quarters, including long-established prisons and college dormitories, were assigned incorrect geographic codes. For example, a prison would be geocoded to a neighboring town or county. While not affecting the total population count, these errors significantly affected counts for some small geographic areas. Such errors account for a large fraction of the population counts that local jurisdictions challenged in the Census Bureau's Count Question Resolution Program.[21]

- There was no system, such as a preprinted group quarters identification code, for tracking individual questionnaires from group quarters residents. Instead, a total count of questionnaires received was recorded on a control sheet for each group quarters at several steps in the process, and discrepancies and errors crept into the processing as a result. The lack of a good tracking system also impeded evaluation.

- A small but undetermined number of group quarters questionnaires were never returned to a local census office and never included in the census. Also, some questionnaires were assigned to a different group quarters after receipt by a local office.

- In May 2000 the National Processing Center at Jeffersonville, Indiana—which was solely responsible for capturing the information on group questionnaires—reported that many questionnaires were not properly associated with a "control sheet" and therefore did not have a group quarters identification number on them. A team of census headquarters staff reviewed an estimated 700,000 group quarters questionnaires to resolve this problem (no official records were kept of this special operation).

- In July 2000 two special telephone operations were implemented to follow up group quarters with no recorded

[21]This program occurred mainly in 2001. It permitted local governments to challenge their census population counts for geocoding errors and other errors that could be well documented. The Census Bureau reviewed such challenges and issued letters to local jurisdictions when their population count was increased (or decreased) after review. Localities could cite the letters for such uses as obtaining federal program funds; however, the revised counts could not be used for reapportionment or redistricting.

population (presumably refusals) and group quarters for which the number of data-captured questionnaires fell far below the population count obtained in advance visits to special places conducted in February–March 2000. Population counts were obtained for these group quarters, and the results used to impute group quarters residents as needed. More than 200,000 group quarters records (almost 3 percent of the group quarters population) were wholly imputed as a consequence of the telephone follow-up and the reconciliation of multiple population counts on the group quarters control sheets.

- Some group quarters residents were mailed a housing unit questionnaire. If they returned it and the address was matched to a group quarters address, they were added to the appropriate group quarters count, but there was no provision to unduplicate such enumerations with enumerations obtained through the group quarters enumeration procedure. From a clerical review of a sample of cases in selected types of group quarters (excluding prisons, military bases, and service-based facilities such as soup kitchens), an estimated 56,000 group quarters enumerations were duplicates of other enumerations within the same group quarters.

- Residents at some types of group quarters (e.g., soup kitchens, group homes, worker dormitories) could declare a "usual home elsewhere" and be counted at that address if the address could be verified. This procedure was not implemented as designed, and the result was that a net of 150,000 people were moved erroneously from the group quarters to the housing unit universe. Of these, 31,000 people were erroneously omitted from the census altogether.

- A higher-than-expected proportion of group quarters enumerations were obtained from administrative records of the special place. Of the 83 percent of enumerations for which enumerators indicated the source of data, 59 percent were filled out from administrative data, 30 percent were filled out by the resident, and 12 percent were filled out by an enumerator interviewing the resident. Types of group quarters with high percentages of enumerations obtained from administrative data included nursing homes, hospitals, group homes, and prisons. These types of group quarters had especially high rates of missing data for long-form items (see Section 7–D).

- Because of poor results in the 1990 Post-Enumeration Survey Program, a deliberate decision was made to exclude group quarters residents from the 2000 A.C.E. Program (see Section 5–D.1 for the basis of this decision). As a consequence, there was no way to assess omissions of group quarters residents or the full extent of erroneous enumerations for this population.

4–F.3 Group Quarters: Summary of Findings and Recommendations

Overall, we conclude that the procedures for enumerating group quarters residents and processing the information collected from them were not well controlled or carefully executed. There is evidence of errors of omission, duplication, and miscoding by geography and type of group quarters, although there are no data with which to conduct a definitive evaluation, partly due to the lack of procedures to track individual enumerations. The extent of imputation required for group quarters residents was high, not only in terms of the number of whole-person imputations required, but also in terms of item imputations, particularly for long-form-sample items (see Section 7–D).

> **Finding 4.6:** The enumeration of people in the 2000 census who resided in group quarters, such as prisons, nursing homes, college dormitories, group homes, and others, resulted in poor data quality for this growing population. In particular, missing data rates, especially for long-form-sample items, were much higher for group quarters residents than for household members in 2000 and considerably higher than the missing data rates for group quarters residents in 1990 (see Finding 7.3). Problems and deficiencies in the enumeration that undoubtedly contributed to poor data quality included: the lack of well-defined concepts of types of living arrangements to count as group quarters; failure to integrate the development of the group quarters address list with the devel-

opment of the Master Address File; failure to plan effectively for the use of administrative records in enumerating group quarters residents; errors in assigning group quarters to the correct geographic areas; and poorly controlled tracking and case management for group quarters. In addition, there was no program to evaluate the completeness of population coverage in group quarters.

Enumeration procedures for group quarters residents need rethinking and redesign from top to bottom if the 2010 census is to improve on the poor performance in 2000. Questionnaire content and design also need to be rethought to obtain higher quality data (see Section 7–D), including the possibility of asking residents for alternate addresses to facilitate unduplication (e.g., home addresses for college students and prisoners). Given the increase in population of people in some types of group quarters and the lack of other data sources for group quarters residents, improvement in the enumeration of group quarters should be a high-priority goal for the 2010 census. Tracking systems should be built into the enumeration not only to facilitate a high-quality operation but also to support subsequent evaluation.

Recommendation 4.4: **The Census Bureau must thoroughly evaluate and completely redesign the processes related to group quarters populations for the 2010 census, adapting the design as needed for different types of group quarters. This effort should include consideration of clearer definitions for group quarters, redesign of questionnaires and data content as appropriate, and improvement of the address listing, enumeration, and coverage evaluation processes for group quarters.**

CHAPTER 5

Coverage Evaluation: Methods and Background

A KEY GOAL OF THE 2000 CENSUS was to reduce the net undercount of the population and the differences in net undercount rates between historically less-well-counted groups (minorities, children, renters) and others (non-Hispanic whites, adults, owners) compared with the results in 1990. Achieving this goal would also probably reduce differences in net undercount rates among states and local areas. For the 1990 census, the Post-Enumeration Survey (PES) estimated a national net undercount of about 1.6 percent, or a shortfall of 4.0 million people. It also estimated a difference of 3.9 percentage points between the net undercount rate for blacks (4.6 percent) and that for non-Hispanic whites (0.7 percent), and a difference of 4.3 percentage points between the net undercount rate for Hispanics (5.0 percent) and that for non-Hispanic whites.[1]

[1]These rates are for the household and noninstitutionalized group quarters population, which was the universe included in the 1990 PES; they are from an August 1992 revision of the original PES estimates (Hogan, 2001a:Table 2a; see also Hogan, 1993:1054).

For the 2000 census the most recent revision of the Accuracy and Coverage Evaluation (A.C.E.)—Revision II—estimated a slight national net *overcount* of about 0.5 percent of the household population, or 1.3 million extra people, which is the first estimated net overcount in census history. The A.C.E. Revision II also estimated a difference of 2.9 percentage points between the net undercount rate for blacks (1.8 percent) and the net overcount rate for non-Hispanic whites (1.1 percent) and a difference of 1.8 percentage points between the net undercount rate for Hispanics (0.7 percent) and the net overcount rate for non-Hispanic whites (U.S. Census Bureau, 2003c:Table 1).

The final Revision II A.C.E. estimates appear to tell a story of success in reducing the net undercount and differences in net undercount rates in 2000 compared with 1990. Despite the extensive, imaginative, and high-quality work by Census Bureau staff to develop the Revision II estimates, however, it is difficult to draw clear-cut conclusions from them about undercounts and overcounts in 2000 because of limitations in the available data for reestimation. The evaluations that fed into the Revision II estimates provided more information than previously available about the numbers and sources of erroneous census enumerations and, similarly, more information with which to determine the residency status of the independent A.C.E. sample. They provided little new information, however, about the numbers and sources of census omissions. Because of changes in estimation methodology, the revised 2000 estimates are not comparable with the 1990 PES results.

To set the stage for our assessment of population coverage in the two censuses in Chapter 6, we first describe the two major coverage evaluation methods—dual-systems estimation (5–A) and demographic analysis (5–B). We then review the history of coverage evaluation for 1990 and 2000 (5–C and 5–D, respectively), including decisions about the possible use of coverage estimates to adjust census counts for measured net undercount. In June 1991 the Census Bureau director recommended that PES-based population estimates be used to adjust the 1990 census counts, but the secretary of commerce decided against adjustment. In December 1992, the Bureau director decided that revised PES estimates would not be used to adjust intercensal population estimates. In 2000 the Census Bureau planned to use A.C.E.-based population estimates to adjust the census counts, but on three separate occasions—March 2001, October 2001, and March 2003—the Bureau recommended against such adjustment, and the secretary of commerce accepted the recommendations.

5–A DUAL-SYSTEMS ESTIMATION WITH THE A.C.E.

The 2000 A.C.E., like its predecessors, the 1990 PES and the 1980 Post-Enumeration Program (PEP), was designed to estimate the population of the United States and major population groups by *dual-systems estimation* (DSE). This method is closely related to a widely used statistical methodology known as capture-recapture, which was first developed for estimating wildlife populations. The methodology requires adaptation for the census context, as described in Fienberg (2000) and Hogan (1992, 2000a,b).

The basic concept of dual-systems estimation is determining how many people counted in one survey are also validly counted in a second, independent survey. In the census context, the initial survey is called the *E-sample*, which consists of the census enumerations in the A.C.E. sample blocks.[2] The subsequent independent survey is the *P-sample*, which consists of people living at addresses listed independently of the census address list in the sample of census blocks. Not every census enumeration is correct; some are erroneous (e.g., a duplicate), so the process also involves estimating how many of the records in the E-sample represent correct enumerations. This is done by visiting E-sample people who fail to match P-sample records to determine for each individual whether he or she was enumerated in the census despite being missed in the P-sample, or whether the person was enumerated in error, such as a duplicate or a fictitious enumeration.

In general terms, the P-sample and the E-sample are used to estimate two components of the formula for calculating the DSE for each of several hundred population groups, called *poststrata*. Poststrata are defined by categorizing people on the basis of such variables as age, sex, race, ethnicity, owner or renter, region, and others; about 400 poststrata were used for the original A.C.E. To the extent possible, the intention is to develop poststrata that group

[2]The E-sample does not include every census enumeration in the A.C.E. blocks, for such reasons as subsampling of large blocks (see Appendix E.1).

people for whom coverage probabilities are as similar as possible within the group and as different as possible from other groups.

The two components estimated for each poststratum are the proportion of the population correctly included in the census, which is estimated by the P-sample match rate, and the proportion of the census records that were correctly included, which is estimated by the E-sample correct enumeration rate:

- The match rate is the weighted estimate, M, of P-sample persons who match with E-sample or other census persons, divided by the weighted estimate, P, of all valid P-sample persons (including matches and nonmatches).

- The correct enumeration rate is the weighted estimate, CE, of E-sample persons who were correctly enumerated in the census (including matches and correct nonmatches), divided by the weighted estimate, E, of all E-sample persons (including correct and erroneous enumerations).

These components are applied in a formula to calculate DSE, the dual-systems estimate for the total population of a poststratum:

$$DSE = (C - II)\left(\frac{CE}{E}\right)\left(\frac{P}{M}\right), \tag{5.1}$$

where:

- $C - II$ is the census count, C, *minus* people requiring imputation and late additions to the census count, II, who are excluded from the E-sample because they cannot be matched to the P-sample;[3]

- CE/E is the weighted correct enumeration rate from the E-sample; and

- P/M is the inverse of the weighted match rate from the P-sample.

[3] II is a Census Bureau term that originally stood for "insufficient information for matching." Its meaning has evolved, and it now covers late additions to the census and whole-person imputations.

For any poststratum, the net undercount rate (*UR*) is

$$UR = \frac{DSE - C}{DSE},$$

and the coverage correction factor (*CCF*) is

$$CCF = \frac{DSE}{C}$$

where *C* is the census count, *including* people requiring imputation and late additions to the count (*II*s).

The basic assumption underlying the calculation of the DSE can be stated as follows: Given independence of the P-sample survey from the census, the estimated proportion of P-sample people in a poststratum who match to the census (*M*/*P*) is a good estimate of the estimated proportion of all people in the poststratum who were correctly enumerated in the census (*CE*/*DSE*). Independence means that the event of being enumerated in the census does not affect the probability of being enumerated in the P-sample (see National Research Council, 1999b:79–80).

Solving for *DSE* in the following equation,

$$\frac{M}{P} = \frac{\left[(C - II)\frac{CE}{E}\right]}{DSE},$$

gives Equation 5.1 above.

Six points are worth noting about dual-systems estimation in the census context. First, the DSE formula (Equation 5.1) includes a factor for *II*s; that is, census enumerations that either lacked sufficient information or were added too late to be included in the A.C.E. matching. These enumerations must be assessed to fully understand census coverage estimates (see Section 6–C.1). The total number of *II*s in 2000 was about 8.2 million people, including 5.8 million whole-person imputations and 2.4 million people whose records were temporarily removed from the census file as part of the special operation to reduce duplicate Master Address File (MAF) addresses in summer 2000 and reinstated too late to be included in the A.C.E. processing (see Section 4–E). There were no truly late enumerations in 2000. The total number of *II*s in 1990 was much smaller—about 2.2 million people, including 1.9 million whole-person imputations and

0.3 million people who were enumerated in coverage improvement programs too late for PES processing.

Second, there is no assumption that the P-sample must be more complete than the E-sample for DSE to work; it is expected that the P-sample will miss some people who were correctly enumerated in the census, and vice versa. What is important is that the information obtained in the P-sample that is needed to determine a match or valid nonmatch be of high quality and satisfy the assumption of independence.

Third, a key assumption in the calculation of the DSE in the census context is that the procedures used to define who is in and who is not in the census are balanced. The E-sample is used to determine how many census enumerations are correctly in the census according to specified criteria (e.g., a college student living in a dormitory should be enumerated at the college and not at his or her parental home). For the DSE model to work, the same criteria must be applied to determine how many P-sample people match to correct census enumerations (whether or not they are in the E-sample). Failure to apply the same criteria will create an error of balancing.

An important dimension of balancing involves geographic correctness. For each person, there is a defined area where he or she should have been enumerated (this is the block cluster in the A.C.E.). In searching for a match for a person in the P-sample, it is important to search all the census enumerations that are in the correct area and only those enumerations in the correct area. Geographic balancing error occurs when the actual search area for P-sample matches is larger or smaller than that used in the E-sample to determine correct enumerations.

Fourth, the DSE is sample based. Consequently, it is important to estimate not only the DSE itself, but also to accurately estimate the variance in the DSE due to sampling error and other sources of variation. In addition, the use of sampling limits the number of individual population groups (poststrata) for which reliable coverage estimates can be developed.

Fifth, if DSE results are to be used to adjust the census for net undercount for geographic areas that are smaller than those used in the poststratification, the process would involve applying the coverage correction factors to the population counted in each geographic area for which adjusted counts are desired, separately for each poststratum. This procedure assumes that the match rate, correct enumeration rate, and other rates involved in the dual-systems computation for a poststratum apply at all lower geographic levels or, alternatively, that the true coverage correction factor for a poststratum applies at all lower geographic levels. This assumption—known as the synthetic assumption—is strong.[4]

Finally, the design of the A.C.E., similar to the design of the 1990 PES and 1980 PEP, was focused on developing good estimates of net undercount. Given the clustered sampling and other features, the A.C.E. design was not well suited for estimating gross coverage errors or categorizing them by type. For example, the A.C.E. was not designed to identify duplicate census enumerations that involve a household with two or more residences in locations outside the search area for matches (e.g., a household with summer and winter residences in different states). In principle, the A.C.E. process would identify half of such duplicate enumerations as correct and half as erroneous (under the category of "other residence"). However, the balancing assumption might not hold, and, in any case, it would not be possible to identify duplicate enumerations as distinct from other kinds of erroneous enumerations in the "other residence" category. This problem became evident when the original March 2001 A.C.E. coverage estimates were subjected to additional evaluation in summer 2001 (see Section 5–D).

In the preceding discussion of dual-systems estimation, we have referred to several assumptions underlying the method. Statistical assumptions need not be precisely true for a method to be useful and, in fact, such assumptions are rarely precisely true. Determining the sensitivity of results to mild failures of the assumption is an important part of evaluation of a method and of research and development leading to improved methods.

[4]The synthetic procedure also assumes that imputed *II*s are assigned correctly to poststrata (i.e., that the imputation accurately imputed their basic characteristics) and that late addition *II*s resulting from the special unduplication process are correctly assigned to geographic areas (see Zhao, 2003).

5–B DEMOGRAPHIC ANALYSIS

Beginning in the 1940s, Census Bureau and university researchers have used demographic analysis (DA) methods to develop population estimates for every census year since 1880. Estimates are developed for the total U.S. population and a limited number of groups defined by age, sex, and race (two categories—black and all other). The method uses aggregate data from administrative records, supplemented by census and survey results, for estimation. Since 1970, estimates have been constructed using separate procedures and data sets for people under age 65 and age 65 and older (see Robinson, 2000, 2001a:App. A; National Research Council, 1985:133–139, 148–151).

Estimates of the population under age 65 are constructed for single-year birth cohorts by sex and race. The procedure uses the number of births as the starting point (e.g., estimates of people age 60 in 2000 begin with the number of births in 1940), subtracts deaths, and adds estimates of net immigration in each year to the estimation year. In practice, the DA estimate for a census year (2000) usually begins with the DA estimate for the previous census year (1990), updated with estimates of births, deaths, and net immigration in the decade between the two years.

Birth and death data are from vital statistics records collected by the states in a common format and forwarded to the federal government. (All states and the District of Columbia have been part of the vital statistics registration system since 1932, except that Alaska and Hawaii joined the system later.) Corrections are made for underregistration of births, using the results of birth registration studies adjusted by interpolation and extrapolation (the latest study covered 1964–1968). Administrative data from the Immigration and Naturalization Service are used to estimate legal immigration, but they must be adjusted to fill in gaps for illegal (undocumented) immigration and emigration of legal and illegal residents (tourists and other short-term visitors are not included).

For the population age 65 and over, estimates are constructed from Medicare enrollment statistics. The Medicare data are adjusted for those not enrolled, who were estimated to be 3.7 percent of people age 65 and over in 2000.

Historically, DA estimates have been viewed as a "gold standard" against which to assess the completeness of census coverage when, in fact, the administrative data sources used in the estimates are subject to error, and DA estimates have periodically been revised. For example, upon observing unusually high net undercount rates for black males ages 45–54 in 1990 that had persisted for this cohort since 1960, an investigation determined that the correction for underregistration of births for this group of blacks—and hence their estimated net undercount—was too high (Robinson et al., 1993:1064).

Since 1980, the main area of uncertainty in DA estimates has involved the immigration component, especially the number of undocumented immigrants. The problem of illegal immigrants first became evident when the initial demographic analysis total population estimate for April 1980 was lower than the population enumerated in the census (226.0 and 226.5 million, respectively—National Research Council, 1985:148). Revised DA estimates, which included an allowance for illegal immigrants and incorporated other changes, gave a total population estimate for April 1980 of 229.4 million (National Research Council, 1985:Table 2.1).

Other components of DA estimates, such as emigrants and some categories of legal immigrants, also add to the margin of error. There are no direct, comparative measures for evaluating the net immigration component, especially the undocumented component. At present, a "residual" process is used to estimate the number of undocumented immigrants; that is, an estimate of the expected number of foreign-born people legally residing in the country is derived from reported data on legal immigration, and this figure is compared with the number of foreign-born people reported in the census long-form sample (which included about 17–18 million households in 1990 and 2000) or, more recently, in the Current Population Survey (CPS, which includes about 50,000 households each month). The difference between the two is an estimate of the number of undocumented immigrants included in the census (or CPS), although that estimate is subject to error. The computations are carried out in some detail by country (or region) of birth and year of entry, which is believed to add to the validity to the estimates. Data on country of birth and year of immigration are now included regularly in the CPS so that the computations can be carried out more frequently, perhaps adding some stability in the estimates over time.

5–C COVERAGE EVALUATION AND ADJUSTMENT IN 1990

5–C.1 Preparing for 1990

Building on the experience with DSE-based coverage evaluation methodology in the 1980 census and extensive research and tests conducted in the mid-1980s (see Anderson and Fienberg, 1999:Ch.5), the Census Bureau developed plans for a 1990 Post-Enumeration Survey of 300,000 households. In spring 1987 professional advisory groups and a National Research Council Panel on Decennial Census Methodology endorsed the Bureau's plans (National Research Council, 1987), which included the possibility that DSE estimates from the PES might be used to adjust the census results for measured net undercount.

In October 1987 the secretary of commerce cancelled the large PES and announced plans for a PES of 150,000 households, which would be used for coverage evaluation only. After a year of controversy, in November 1988, the State and City of New York and a coalition of local governments filed suit in federal court in New York charging that the secretary's decisions to cancel the large PES and ban a possible adjustment were arbitrary and capricious. In July 1989 when the adjustment lawsuit trial was scheduled to begin, the parties reached a partial settlement. The terms of the settlement included that a 150,000-household PES would be conducted, that the results could be used for adjustment, that a Special Secretarial Advisory Panel of experts would be appointed to advise the secretary (four chosen by the secretary and four by the plaintiffs), and that the secretary would reach a decision on adjustment by July 15, 1991. These provisions ensured that the PES operations and results would receive intense scrutiny.

5–C.2 PES Design and Operations

The final design for the 1990 PES included an area-probability sample of 5,290 census block clusters (comprising one or, in some cases, more than one block) for which an independent address listing was conducted. The P-sample cases from the independent list and the E-sample cases of census enumerations in the PES block clusters totaled about 165,000 households each (slightly larger than the originally planned size of 150,000 households). The two samples also included people living in most types of noninstitutional group quarters. To produce more reliable coverage estimates for minorities, extensive oversampling was used, which resulted in large variation in the PES sampling weights.

Beginning in July 1990, P-sample cases were interviewed using paper-and-pencil questionnaires. A procedure called PES-B was used in gathering information, in which interviewers listed the current residents and ascertained where they lived (at the P-sample address or another address) on Census Day, April 1, 1990. (See Marks [1978], who also describes a PES-A procedure in which each P-sample address is visited to find out who lived there on Census Day. For this procedure, information must be sought from current residents about people who moved away after Census Day.)

Upon completion of interviewing, the P-sample cases were matched with the E-sample cases, using computerized matching software to process as many cases as possible, followed by clerical matching when the computer algorithm could not determine a match. Matches for P-sample people who moved to the P-sample address after Census Day were sought by searching the census records in the area of their residence on Census Day. If a P-sample case did not match any E-sample case in the PES block cluster, then one, or sometimes two, rings of surrounding blocks were searched for P-sample matches. One or two rings were also searched for E-sample correct enumerations that had been identified as geocoding errors—that is, their addresses were incorrectly assigned to the block cluster.

Following matching operations, unmatched and unresolved P-sample and E-sample cases were followed up in the field to obtain more information to try to determine their status. After using the information from field follow-up to determine a final match and enumeration status for as many P-sample and E-sample cases as possible, statistical imputation techniques were used to assign values for missing characteristics and for unresolved match status (P-sample) and unresolved enumeration status (E-sample). All of these operations were completed by December 1990.

In spring 1991, Bureau statisticians used the PES data to develop DSE-based population estimates for the 1990 census. The P-sample and E-sample cases were categorized into 1,392 prespecified poststrata for which coverage correction factors were calculated. It turned out that, because of the reduction in sample size from 300,000 to 165,000 households, the poststrata coverage correction factors and their variance estimates exhibited greater variability than anticipated. Hence, a last-minute decision was made to "presmooth" the variance estimates. The new variance estimates were then used in a planned procedure to reduce the sampling variability of the coverage correction factors. In this procedure, the factors were smoothed by shrinking them toward a predicted value from a multiple regression model (see Anderson and Fienberg, 1999:Ch.5). The final step was to apply the smoothed coverage correction factors for individual poststrata to the census counts, block by block, to develop adjusted population counts for the United States, states, and small areas.

5–C.3 July 1991 Adjustment Decision

The Census Bureau released preliminary PES results on April 18, 1991, which showed a national net undercount of 2.1 percent, and a difference in net undercount rates between blacks and nonblacks of 3.1 percentage points. These estimates were similar to the demographic analysis estimates of a 1.8 percent net undercount overall and a difference in net undercount rates between blacks and nonblacks of 4.4 percentage points (Anderson and Fienberg, 1999:Tables 6.2 and 6.3). Census Bureau staff prepared 19 reports in the P-series evaluating many facets of the PES operations and results (see Anderson and Fienberg, 1999:App.E). The reports were made available in draft form to members of the Special Secretarial Advisory Panel and within the Bureau. In June the four plaintiff-appointed members of the advisory panel recommended that the PES results be used to adjust the census counts, while the four members appointed by the Commerce Department recommended against adjustment. Also in June, an internal Census Bureau Undercount Steering Committee voted seven to two in favor of adjustment, and the director recommended adjustment to the secretary. On July 15, 1991, the secretary of commerce decided against adjustment and released several volumes of supporting documents.

5–C.4 December 1992 Adjustment Decision

The secretary of commerce in July 1991 announced his intention to consider the use of PES estimates to improve the widely used postcensal population estimates that are developed by using birth, death, and immigration records to update the census counts (see Chapter 2). A decision was anticipated before the next release of estimates, scheduled for July 1, 1992, but the decision was subsequently deferred until December.

An internal Census Bureau Committee on Adjustment of Postcensal Estimates (CAPE) conducted additional evaluations of the PES estimates. Its August 1992 report (Bureau of the Census, 1992) documented a computer processing error in the PES estimation that was detected in late 1991. The error increased the estimate of erroneous enumerations which, through the operation of the DSE formula (see Equation 5.1 above), decreased the DSE estimate of the population and consequently reduced the national net undercount by 0.4 percentage points—from 2.1 to 1.7 percent. The CAPE report also presented the results of a reestimation that used 357 poststrata instead of the original 1,392 strata in response to criticisms about the presmoothing and smoothing models used for the original estimation. The report also presented the results of additional refinements and corrections to the PES estimates. The final revised 1990 PES estimates put the national net undercount at 1.6 percent and the difference in net undercount rates between blacks and nonblacks at 3.2 percentage points. In December the Census Bureau director decided not to use the revised PES estimates to adjust the postcensal population estimates, although, at the request of the Bureau of Labor Statistics and other statistical agencies, the Census Bureau used unpublished adjusted national estimates for age, race, ethnicity, and sex groups to correct major household surveys (e.g., the CPS) for undercoverage relative to the census. The secretary of commerce did not assert a role in these decisions.

5–C.5 1993–1996 Court Decisions

Following the July 15, 1991, decision by the secretary of commerce against adjustment of the 1990 census data, the New York plaintiffs renewed their lawsuit. In April 1993 a federal judge ruled on narrow grounds that the secretary's decision was not arbitrary or capricious. This ruling was vacated by a panel of the Second Circuit Court in August 1994, which required further consideration of adjustment on the merits. The states of Oklahoma and Wisconsin appealed to the U.S. Supreme Court, which in March 1996 reversed the Circuit Court decision, allowing the unadjusted census counts to remain the official 1990 census numbers (*Wisconsin v. City of New York*, 517 U.S. 1, 1996).

5–D COVERAGE EVALUATION AND ADJUSTMENT IN 2000

We described in Chapter 3 how the evidence of an increased net undercount in the 1990 census (which resulted despite a 33 percent increase in costs compared with 1980) fueled efforts by the Census Bureau to develop a more cost-effective design for the 2000 census. The original 2000 design called for a large 700,000-household survey and matching operation conducted on a schedule that would permit integrating DSE-based population estimates into the census counts in time to produce adjusted state totals for reapportionment of Congress by December 31, 2000. The Integrated Coverage Measurement (ICM) sample was designed to produce reliable estimates for individual states, so that each state's coverage correction factor would be estimated from the sample for that state and not require the use of sample from other states.

The U.S. Supreme Court precluded the use of sample-based estimation for the reapportionment counts, so the Census Bureau planned the 2000 Accuracy and Coverage Evaluation Program with a 300,000-household survey and completion of DSE-based estimates in time to produce adjusted small-area census counts for legislative redistricting by the mandated deadline of April 1, 2001. The smaller sample size for the A.C.E. was made possible because there was no longer a need to produce direct state estimates; instead, estimates could borrow strength across states. In a May 1999 letter report to the Census Bureau director, our panel supported this decision, although it urged the retention of a minimum sample size in each state that could be used to evaluate the assumption that state effects on coverage correction factors are less important than the effects of such characteristics as age, race, sex, and housing tenure (see

Appendix A.4.a). Controversy over the adjustment issue was even more intense for 2000 than had been the case for 1990.

5–D.1 A.C.E. Design and Operations

Although the 2000 A.C.E. was broadly similar to the 1990 PES, the two programs differed in a number of respects (see Hogan, 2000b), and Census Bureau staff expected that changes implemented for the A.C.E. would produce superior results. This section briefly summarizes the major differences (see Appendix E for a description of A.C.E. operations).

Universe

The A.C.E. universe excluded people living in institutions, college dormitories, and other group quarters; the PES universe included most noninstitutional group quarters. The Census Bureau decided to limit the A.C.E. to the household population because of its experience in the 1990 PES in which rates of unresolved match status were much higher for group quarters residents than for household members in the PES because of much higher rates of short-term mobility for people in group quarters (e.g., college students taking and returning from spring break, shelter residents moving from one shelter to another, migrant worker dormitory residents moving from one farm to another). However, this decision necessitated an unrealistic assumption of perfect coverage of the group quarters population.

Sample Size and Design

The 2000 A.C.E. was twice the sample size of the 1990 PES—300,000 households in about 11,000 block clusters compared with 165,000 households in about 5,000 block clusters. Because of its larger overall sample size, the A.C.E. could produce reliable direct estimates for minorities and other groups with less oversampling than was used in the PES. The A.C.E. design also improved on several other features of the PES sample design. Consequently, the A.C.E. weights varied less than the PES weights, which contributed to reducing the variance of the A.C.E. estimates (see Appendix E.1).

Initial Housing Unit Match

The A.C.E. included a new operation to match P-sample and January 2000 MAF housing units prior to interviewing. The purpose of the match was to facilitate such operations as large block subsampling, telephone interviewing, and matching. Although the P-sample and census address lists were linked, independence was maintained because no changes were carried over from one list to the other as a consequence of the match.

P-Sample Interviewing Technology and Schedule

The A.C.E. used computer-assisted telephone and personal interviewing (CATI/CAPI) to facilitate the accuracy of the information collected and the speed of data capture and processing. The PES used paper-and-pencil methods throughout. The first wave of A.C.E. interviewing began in late April 2000, using the telephone to contact mail return households that provided a telephone number on their census questionnaire and for which there was a clear city-style address. Fully 29 percent of the P-sample household interviews were obtained by telephone. The second wave of interviewing was in person; it began in mid-June and was completed in August, well ahead of the timing achieved in 1990.

Matching Technology

The A.C.E. clerical matching operation, which followed a computer match, was conducted by clerks examining computerized P-sample responses and questionnaire images for census cases in the sampled block clusters. The technology was designed to be user-friendly. Because of complete computerization of the operation, all matching could be done at one location, instead of seven as in 1990.

Treatment of Movers

A major change from 1990 was the treatment of movers. The goal of the 1990 PES was to visit each P-sample address and find out where the current residents usually lived as of Census Day (PES-B). The original design for Integrated Coverage Measurement for 2000 ruled out PES-B because of the plan to use sampling for nonresponse follow-up (see Section 3–B.3). This plan meant that some people in the P-sample who moved to their present residence after Census Day might be falsely coded as not matching to a census enumeration. This could happen if they did not mail back a questionnaire and their Census Day address was not included in the sample for nonresponse follow-up. The decision not to use PES-B was carried over to the A.C.E.

The 2000 A.C.E. had two objectives: (1) to find out who lived at each P-sample address on Census Day and determine whether they were enumerated or missed in the census at that address and (2) to find out who lived at each P-sample address as of the A.C.E. interview day. This procedure is called PES-C; it results in obtaining information not only for nonmovers and inmovers (interview day residents not resident on Census Day), but also for outmovers (Census Day residents not resident on interview day).

The PES-C procedure involved estimating the P-sample match rate for movers by matching outmovers. At the same time, for most poststrata, the A.C.E. estimated the *number* of matched movers by applying the *outmover match rate* to *inmovers*. The underlying assumption is that inmovers would be more completely reported than outmovers. The advantage of PES-C is that a searching operation for the Census Day residence of inmovers is not required. The drawback is that the quality of the information collected to use in the matching for outmovers may be impaired because their information is always supplied by other people. The decision to use the PES-C procedure added complexity to the A.C.E. estimation. Tests of Integrated Coverage Measurement in 1995 and 1996 used PES-A, in which the goal is to find out who lived at each P-sample residence on Census Day and determine whether they were enumerated or missed in the census. PES-C is similar to PES-A except that PES-C applies the match rate for outmovers to the number of inmovers, instead of using both the match rate and number of outmovers in the DSE estimation (see Liu et al., 2001:1).

Targeted Extended Search Procedure

Another important change from the 1990 PES concerned the targeted extended search procedure for searching surrounding blocks if a search in the sampled block cluster did not turn up a match for a

P-sample household and to find out if misgeocoded E-sample cases were located nearby. Such searching could reduce the variance and bias of the DSE estimates. For efficiency reasons, it was decided in the A.C.E. to target the extended search on selected P-sample and E-sample cases and to conduct it for a subsample of the A.C.E. blocks, instead of conducting searches for every block as in 1990.

Definition of Poststrata

The 448 poststrata in the original A.C.E. design (reduced to 416 for developing the original DSE estimates) were similar to the 357 poststrata that were implemented in the reestimation of the 1990 PES in 1992 (see Section 5–C.4). The 2000 poststrata (see Appendix E, Table E.3) included two additional race/ethnicity domains, one for American Indians and Alaska Natives not living on reservations and another for Native Hawaiian and Other Pacific Islanders (who had been combined with Asians in 1990). The 2000 poststrata also categorized non-Hispanic whites and other races, non-Hispanic blacks, and Hispanics by mail return rate (two categories—high and low—calculated separately for each group by housing tenure). Region was dropped as a stratifier except for people in the non-Hispanic white and other race category who owned their homes. Subsequently for the A.C.E. Revision II reestimation in 2002, extensive changes were made to the E-sample poststratification (see Section 6–B.4).

5–D.2 March 2001 Adjustment Decision

In fall 2000, Clinton commerce secretary Norman Mineta delegated authority for the decision on adjustment of the 2000 census to the director of the Census Bureau (65 *Federal Register* 195, 59713–59716). In the first days of the Bush administration, commerce secretary Donald Evans rescinded the earlier regulation, returning the authority to the secretary (66 *Federal Register* 37, 11231–11233). The administration also accepted the resignation of director Kenneth Prewitt.

Both versions of the rule on the authority for deciding on adjustment formally established within the Census Bureau an Executive Steering Committee for A.C.E. Policy (ESCAP), which was made up of senior Bureau staff, to recommend to the director whether to release adjusted or unadjusted block data for redistricting in March 2001. After reviewing the original A.C.E. results and an extensive series of evaluations, ESCAP recommended that unadjusted counts from the enumeration process should be the official data for legislative redistricting. The acting director accepted this recommendation and transmitted it to the secretary of commerce on March 1, 2001. The secretary, after consulting with several outside experts, adopted the Bureau's recommendation on March 6, and unadjusted census redistricting data (the P.L. 94-171 files—see Chapter 2) were released state-by-state during March.

The Bureau's decision against adjustment was surprising to many in light of the evolution of the census design over the decade of the 1990s (see Chapter 3) and particularly the more prominent role of the A.C.E. relative to the analogous 1990 PES. Also, public statements by Census Bureau officials before the census was completed stressed the limitations on the ability to count everyone in the nation through field procedures and the likelihood that statistical adjustment would improve the estimates of the population for important purposes. For example, Census Bureau director Kenneth Prewitt wrote: "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" (Prewitt, 2000:2).

Census Bureau's Decision Process

The Census Bureau reached its conclusion not to adjust after carefully following the decision process it had specified, which was publicly explained at an open workshop held by our panel in October 2000 (see National Research Council, 2001g). All of the evaluations that the Bureau proposed to conduct were completed and reviewed by ESCAP. The ESCAP decision report and the completed A.C.E. evaluations are available at http://www.census.gov/dmd/www/EscapRep.html [1/10/04].

Given the time constraints, these evaluations could not be exhaustive, but they included detailed assessments of A.C.E. operations, supplemented by more limited assessments of census operations and comparisons of adjusted and unadjusted census counts for different levels of geography (Hogan, 2000c). The A.C.E. evaluations covered rates of noninterviews in the independent P-sample and missing data in the P-sample and E-sample; quality control of the matching process; the extent of imputation required for unresolved residence, match, and enumeration status; inconsistent assignment of sample cases to poststrata in the two samples; and variance due to sampling and imputation error in the DSE estimates. The census evaluations covered mail return rates; quality assurance of enumerators' field work; results of unduplication operations; and extent of missing data. Comparisons with 1990 census data were included when feasible. It was hoped that these assessments, which largely addressed how well operations were performed, would provide sufficient information to conclude that adjusted counts did or did not improve the counts from the census process. In addition, the Census Bureau planned to take account of population estimates from demographic analysis, which historically had provided a comparison standard for the census.

What, then, were the reasons for the decision not to adjust? An important reason cited by the ESCAP report (Executive Steering Committee for A.C.E. Policy, 2001c) was the inconsistencies between the population estimates from the census, the A.C.E., and demographic analysis; those inconsistencies could not be resolved or explained with the available evaluation data within the time available for the decision.

The A.C.E. estimated as of March 2001 that the overall net undercount was only 1.2 percent of the population in 2000 (see Table 5.1)—a decrease from the revised PES estimate of a 1.6 percent net undercount in 1990. However, demographic analysis suggested that the undercount had been reduced even more than was estimated by the A.C.E. The Census Bureau's initial demographic analysis estimate (as of January 2001) indicated that the 2000 census resulted in a slight (0.7 percent) net *overcount* of the population and that the A.C.E. overstated the population by even more. Even when the Bureau adjusted the demographic analysis estimate upward in March 2001 to allow for a larger number of undocumented immigrants than were part of the base estimate, the alternate demographic analysis estimate of the 2000 net undercount was only 0.3 percent of the population (Table 5.1).

Table 5.1 Alternative Estimates of the Population and the Percentage Net Undercount, April 2000 (Original March 2001 A.C.E., Base DA, Alternate DA)

Source	Population (in millions)	Estimate of Percentage Net Undercount
Census Count	281.4	—
A.C.E. Estimate (Original, March 2001)	284.7	1.15
Base Demographic Analysis[a]	279.6	−0.65
Alternate Demographic Analysis[b]	282.3	0.32

NOTES: All estimates include the household population (273.6 million people), the group quarters population (about 7.3 million people), and people enumerated in special operations not included in the A.C.E. (e.g., the remote Alaska enumeration). The percentage net undercount is calculated as the population estimate minus the census count divided by the estimate. Minus sign (−) indicates a net overcount of the population.

[a] The base demographic analysis estimate was made available in January 2001; it assumes about 6.0 million undocumented immigrants living in the United States under age 65 in April 2000—3.3 million from the 1990 DA estimate plus a net increase during the 1990s of 2.7 million.

[b] The alternate demographic analysis estimate was made available in March 2001; it assumes a net increase in illegal immigration during the 1990s of 5.4 million for a total of 8.7 million undocumented immigrants as of April 2000.

SOURCE: Robinson (2001a:Table 3).

In addition to the discrepancies between the A.C.E., demographic analysis, and census results, the ESCAP report cited several areas of concern about A.C.E. operations that might have affected dual-systems estimation as reasons for recommending against adjustment for redistricting purposes. It questioned the level of balancing error that may have occurred in the targeted extended search procedure. (Balancing error occurs when different criteria, such as different areas of search, are used in processing the P-sample and E-sample.) It also questioned the level of synthetic error that may have occurred for the poststrata DSE estimates. (Synthetic error occurs when the people included in a poststratum—who are intended to have the same likelihood of being included in the census or the A.C.E.—are in fact not sufficiently similar in this respect.) The report also considered the late additions to the census and cases of people who required whole-person (or whole-household) imputation. Neither of these groups could be included in the A.C.E. There were substantially more such people in 2000 than in 1990, but the report concluded they probably did not affect the DSE estimates. Census Bureau staff also expressed concern about the relatively low numbers of certain types of erroneous enumerations measured in the A.C.E. compared with the 1990 PES—specifically, the A.C.E. found only about two-fifths the number of duplicates and "other residence" erroneous enumerations as were measured in the PES.

The Census Bureau had always planned a longer-term evaluation program, in addition to the short-term evaluations that were feasible to carry out before March 1, 2001. The Bureau proposed to expedite several evaluations on the longer-term agenda and carry out additional evaluations to help reach a second decision by October 15 on whether to recommend adjustment of census population estimates for such purposes as allocation of federal program funds (Executive Steering Committee for A.C.E. Policy, 2001a).

Panel's Assessment

The panel, in its second letter report issued November 2000 (National Research Council, 2000b:2, reproduced in Appendix A), commented on the Census Bureau's evaluation process for the March 2001 decision as follows:

> The planned analyses appear to cover all of the evaluations that can reasonably be expected to be completed within the time available. Furthermore, they appear to be sufficiently comprehensive that they will likely provide support for a reasonably confident decision on adjustment in March.
>
> However, since the numbers themselves, which are, of course, critical to the evaluation process, are not yet available, it is not possible at this time to comment on what the adjustment decision should be nor to conclude definitively that the planned short-term evaluations will be adequate to support the decision.

As it turned out, the Bureau concluded that the evaluation studies did *not* provide sufficient information to decide that adjusted counts would be clearly preferable to unadjusted counts for redistricting. Although not mentioned by the Census Bureau, reaching a conclusion on this point is more difficult when the adjustments to be made for population groups are generally small. A small (or zero) net undercount for the population as a whole is not a reason for or against adjustment because net undercounts can mask sizable gross errors of omissions and erroneously included enumerations. The issue is how the balance between these components of error differs among population groups and geographic areas, resulting in different net undercount rates.

The panel did not necessarily agree with the weight that the Bureau gave to each factor in its decision: specifically, the panel concluded that demographic analysis estimates were sufficiently uncertain that they should not be used as a definitive standard for evaluation. Nonetheless, the panel agreed that the Bureau followed a reasonable process. It also concluded that the Bureau's decision not to recommend adjusting the census data in March was justified, given that additional evaluations of the quality of the A.C.E.—and of the census itself—were needed to resolve the Bureau's concerns.

5–D.3 October 2001 Adjustment Decision

On October 17, 2001, the Census Bureau recommended that unadjusted data from the 2000 census enumeration process should be used for the allocation of federal funds and other purposes. The Bureau released extensive evaluation materials to accompany its decision; the Bureau's decision document and supporting materials are archived at http://www.census.gov/dmd/www/EscapRep2.html [1/10/04]. Completed evaluations of the A.C.E. covered studies of person duplication, the accuracy of information on which E-sample cases were classified as correct or erroneous enumerations, the accuracy of information on which P-sample cases were classified as matches or nonmatches, the targeted extended search, classifying movers in the P-sample, and imputation for unresolved status. The Census Bureau also released revised demographic analysis estimates for 2000 and 1990 with accompanying documentation (Robinson, 2001b).

Census Bureau's Basis for Its Decision

The additional evaluations of the A.C.E. conducted by Bureau staff in summer 2001—while not complete—made it clear that there were considerably more errors of overcounting (erroneous enumerations) in the census than were originally estimated by the A.C.E.

The evaluations suggested that because the A.C.E. did not identify a substantial number of these overcounting errors (mostly duplicates), the use of the original A.C.E. data to adjust the census could lead to overstating the population. Therefore the Census Bureau recommended against adjustment of the 2000 complete-count and long-form-sample data used for fund allocation and other purposes (Executive Steering Committee for A.C.E. Policy, 2001b).

Two principal evaluations of the E-sample identified problems with the classification of erroneous enumerations in the A.C.E. The first was the Evaluation Follow-up Study (EFU), in which field staff, using a more detailed interview than in the original A.C.E., revisited a subsample of the E-sample housing units in one-fifth of the A.C.E. block clusters to determine who should have been counted there. The EFU subsample was about 70,000 people; it was subsequently reduced to 17,500 people for a detailed clerical review that focused on unresolved cases (Adams and Krejsa, 2001; Krejsa and Raglin, 2001). The second evaluation was the Person Duplication Studies, which included computer matching of E-sample records by name and date of birth to census enumerations nationwide. This matching could be conducted because optical character recognition technology used by the Bureau for the first time in 2000 to process the questionnaires made it feasible to include names on all of the computerized sample records. The results of the EFU clerical review and the Person Duplication Studies were combined to develop an estimate of 2.9 million duplicates and other erroneous enumerations in the census that were not measured in the original A.C.E. (Fay, 2001). Such cases included college students who were counted both at their college dormitory and at their parents' household; prisoners who were counted both at prison and at their family's residence; children in joint custody who were counted in the homes of both parents; and people with more than one house, such as those who lived part of the year in the South or West and the rest of the year in the North or Midwest.

Estimated Effects of Unmeasured Erroneous Enumerations on Net Undercount

The results of Fay (2001) were used by Thompson et al. (2001) to construct revised preliminary estimates of the 2000 net undercount.

Unlike the original A.C.E., the revised figures were not built up from estimates for individual poststrata, but were constructed at an aggregate level for the total population and three race/ethnicity groups—non-Hispanic blacks, Hispanics, and all other.[5] The calculations were based on an assumption that a factor for duplicates not detected by the computer matching applied equally to all race/ethnicity groups. They were based on other simplifying assumptions as well, such as that P-sample errors were not likely to affect the dual-systems estimate of the population.

Thompson et al. (2001:1) termed the revised estimates an "early approximation" of the likely effects on the estimated net undercount that might result from a corrected A.C.E. These estimates showed a reduction in the estimated net undercount of the total population in 2000 from 1.2 percent (March estimate) to 0.1 percent (October estimate) and a narrowing of the differences in net undercount rates for blacks and Hispanics compared with all others (Table 5.2).

Revised Demographic Analysis Estimates

The revised demographic analysis estimate of the total net undercount in 2000 (see Table 5.3) was virtually the same as the revised A.C.E. estimate—0.1 percent of the population. The revisions to the demographic analysis estimates incorporated additional information for estimating net immigration (particularly illegal immigration) from the 2000 census itself (the long-form sample) and the Census 2000 Supplementary Survey. It also reflected new assumptions about the extent of undercount of legal immigrants and the completeness of birth registration. Some of these changes increased the DA population estimates; others decreased them (see Robinson, 2001b).

[5]On April 4, 2002, the Census Bureau released revised preliminary estimates of net undercount for seven race/ethnic groups (Mule, 2002b): American Indian and Alaska Native, black (not Hispanic), Hispanic, Asian and Native Hawaiian and Other Pacific Islander (not Hispanic) together, Asian (not Hispanic), Native Hawaiian and Other Pacific Islander (not Hispanic), and white or some other race (not Hispanic). These estimates were derived in the same manner as the estimates in Thompson et al. (2001).

Table 5.2 Alternative Survey-Based Estimates of Percentage Net Undercount of the Population, April 2000 (Original A.C.E. and Preliminary Revised A.C.E.) and April 1990 (Revised PES) (standard error percents in parentheses)

	2000 Estimates		
Category	Original A.C.E. (March 2001)[a]	Revised Early Approximation (October 2001)[b]	Revised 1990 PES Estimate[c]
Total Population	1.18 (0.13)	0.06 (0.18)	1.61 (0.20)
Black, non-Hispanic	2.17 (0.35)	0.78 (0.45)	4.57 (0.55)
Hispanic	2.85 (0.38)	1.25 (0.54)	4.99 (0.82)
All Other	0.73 (—)	−0.28 (—)	0.68 (—)

NOTES: Net undercount rates calculated as the estimate—from the A.C.E. or the PES—minus the census count divided by the estimate. The census count of the population in 2000 was 281.4 million; the census count in 1990 was 248.7 million. —, not available.

[a] Data from Thompson et al. (2001:Table 1, col. 1); see also U.S. Census Bureau (2003c:Table 1). Includes household population. Race/ethnicity defined according to the domain specifications for the A.C.E. (see Appendix E).

[b] Data from Thompson et al. (2001:Table 1, col. 3); see also U.S. Census Bureau (2003c:Table 1). Includes household population. Race/ethnicity defined according to the domain specifications for the A.C.E. (see Appendix E). Takes the A.C.E. estimates of percentage net undercount and subtracts adjustments estimated by Fay (2001:Table 9) for additional unmeasured erroneous enumerations, including an assumption that computer matching was 75.7 percent efficient in identifying duplicates.

[c] Data from Hogan (2001a:Table 2b); see also Hogan (1993:Table 3). Includes household and noninstitutional group quarters population. Race/ethnicity definitions are not strictly comparable with 2000; "all other" is white and some other race, not Hispanic. The "revised" estimate is that developed by the Committee on Adjustment of Postcensal Estimates in August 1992 (published estimates from the committee—see Thompson [1992]—are for the total population and so may differ from the estimates shown).

Panel's Assessment

The revised A.C.E. estimates released in October 2001 were based on preliminary analysis of small subsets of the A.C.E. data and incorporated a number of simplifying assumptions. They could not serve as the basis of an adjustment. The original A.C.E. estimates could not serve as the basis of an adjustment either, as they were most likely to overstate the population. It should be noted, however, that the evaluation work focused on identifying additional census erroneous enumerations not measured in the A.C.E. There was no similarly intensive effort to identify additional omissions not

Table 5.3 Alternative Demographic Analysis Estimates of Percentage Net Undercount of the Population, April 2000 (Base, Alternate, and Revised DA) and April 1990 (Base and Revised DA)

	2000 Estimates			1990 Estimates[a]	
Category	Base[b,c] (Jan. 2001)	Alternate[b,d] (Mar. 2001)	Revised[b,e] (Oct. 2001)	Base (1991)	Revised (Oct. 2001)
Total Population	−0.65	0.32	0.12	1.85	1.65
Black	2.80	3.51	2.78	5.68	5.52
All Other	−1.19	−0.17	−0.29	1.29	1.08

NOTES: Net undercount rates are calculated as the estimate from demographic analysis minus the census count divided by the estimate. The census count of the population in 2000 was 281.4 million; the census count in 1990 was 248.7 million. Minus sign (−) indicates a net overcount of the population.

[a] Data from Robinson (2001b:Table 2). Includes household and group quarters population. Base is the estimate developed following the 1990 census; "revised" is the October 2001 estimate, with revisions to such components as births.

[b] Data from Robinson (2001a:Table 6; 2001b:Table 2). Includes household and group quarters population. The estimates by race are an average of estimates calculated using two different tabulations of the census. "All other" includes Hispanics not classified as black.

[c] "Base" is the original January 2001 estimate, including an allowance for 6 million illegal immigrants—3.3 million from the 1990 demographic analysis estimate and a net increase of 2.7 million over the decade, extrapolated from estimates that mainly reflect changes between 1992 and 1996.

[d] "Alternate" is the March 2001 estimate, including an allowance for 8.7 million illegal immigrants. This estimate was developed as an illustrative alternative to the base estimate when it became apparent that the latter probably underestimated illegal immigration. The alternate estimate reflects an assumed doubling of the net increase in illegal immigrants in the 1990s—from 2.7 to 5.4 million.

[e] "Revised" is the October 2001 estimate, which revises several components, including births and legal and illegal immigration.

measured in the A.C.E. An attempt to account for such omissions through the use of sex ratios from demographic analysis was part of the Revision II estimation (see Section 6–B.5).

From the work done for the October 2001 A.C.E. estimates, our panel concluded in its third letter report, issued November 26, 2001 (reprinted in National Research Council, 2001a:3–14), that the Census Bureau's decision not to adjust the census at that time was justified. The panel urged the Census Bureau to complete the research necessary to develop reliable revised estimates of the net coverage errors in the census, particularly for population groups, in order to determine whether their use would improve the Bureau's population estimates that are regularly produced during the postcensal period and are used for many purposes. The Bureau agreed to reestimate the A.C.E. by the end of 2002 and announce a decision about possible adjustment of postcensal estimates shortly thereafter. We describe and evaluate that work, which produced the final Revision II A.C.E. estimates of the population, in Chapter 6.

The panel also commented in its third letter report on the revised demographic population estimates for 2000 (reprinted in National Research Council, 2001a:11):

> We commend the Bureau for its work to examine each component of demographic analysis. However, its revised estimates of the immigration component are not independent of the census, and the estimates for births and immigration incorporate assumptions that are based primarily on expert judgment. Such judgments may be reasonable, but they retain sufficient uncertainty so that it is not appropriate to conclude that the revised demographic analysis estimates are more accurate than the census.

In Chapter 6, we comment on the role of demographic analysis and a revised A.C.E.-type coverage evaluation program for assessing coverage of the population in 2010.

CHAPTER 6

The 2000 Coverage Evaluation Program

WE TURN NOW TO THE FINAL REVISION II ESTIMATES of the population from the 2000 Accuracy and Coverage Evaluation (A.C.E.) Program. By early 2003 the Census Bureau had completed extensive reanalyses of the evaluation data sets used for the October 2001 preliminary revised A.C.E. estimates of net undercount. It released a new set of detailed estimates based on the original A.C.E. and evaluation data—the A.C.E. Revision II estimates—at a joint public meeting of our panel and the Panel on Research on Future Census Methods on March 12, 2003 (see http://www.census.gov/dmd/www/Ace2.html [12/22/03]). These estimates showed a small net overcount of 0.5 percent of the total population instead of a net undercount of 1.2 percent as originally estimated from the A.C.E. in March 2001. The latest demographic analysis estimates still showed a net undercount of the population, but it was negligible (0.1 percent) (Robinson, 2001b:Table 2).

At the joint panel meeting Census Bureau officials announced the Bureau's recommendation, accepted by the secretary of commerce, to produce postcensal population estimates on the basis of the official 2000 census results. The A.C.E. Revision II estimates would not be used as the basis for developing estimates throughout the decade of the 2000s. The decision document (U.S. Census Bureau, 2003b:1) stated:

> The Accuracy and Coverage Evaluation (A.C.E.) Revision II methodology represents a dramatic improvement from the previous March 2001 A.C.E. results. However, several technical errors remain, including uncertainty about the adjustment for correlation bias, errors from synthetic estimation, and inconsistencies between demographic analysis estimates and the A.C.E. Revision II estimates of the coverage of children. Given these technical concerns, the Census Bureau has concluded that the A.C.E. Revision II estimates should not be used to change the base for intercensal population estimates.

With the final decision not to adjust population estimates for measured net undercount in the 2000 census behind it, the Census Bureau announced its intention to focus exclusively on planning for the 2010 census. Plans for that census include work on possibly using computerized matching of the type conducted for the October 2001 and March 2003 adjustment decisions to eliminate duplicate enumerations as part of the census process itself. Bureau officials also expressed the view that coverage evaluation could not be completed and evaluated in a sufficiently timely fashion to permit adjustment of the data used for legislative redistricting (Kincannon, 2003).

In this chapter we assess the A.C.E. Revision II estimation methodology and the resulting estimates of population coverage in the 2000 census. We first review key aspects of the original A.C.E. Program (6–A) and then review the data sources, methods, and results of the A.C.E. Revision II effort (6–B). We discuss two kinds of enumerations that had more impact on coverage in 2000 than in 1990: (1) whole-person (including whole-household) imputations and (2) duplicate enumerations in the census and the A.C.E. (6–C). Section 6–D provides an overall summary of what we know and do not know about population coverage in 2000. Section 6–E provides our recommendations for coverage evaluation research and development for 2010.

6–A ORIGINAL A.C.E. DESIGN AND OPERATIONS

An important part of evaluating the Revision II A.C.E. population estimates for 2000 is to consider how well the original A.C.E.

Program was designed and executed to produce the underlying input data. We consider below 10 aspects of the original A.C.E. followed by a summary of findings in Section 6–A.11:

(1) basic design;

(2) conduct and timing;

(3) definition of the P-sample—treatment of movers;

(4) definition of the E-sample—exclusion of "insufficient information" cases;

(5) household noninterviews in the P-sample;

(6) imputation for missing data in the P-sample and E-sample;

(7) accuracy of household residence information in the P-sample and E-sample;

(8) quality of matching;

(9) targeted extended search; and

(10) poststratification.

6–A.1 Basic Design

The design of the 2000 A.C.E. was similar to the 1990 Post-Enumeration Survey (PES). The goal of each program was to provide a basis for estimating two key components of the dual-systems estimation (DSE) formula for measuring net undercount or overcount in the census (see Section 5–A). They are:

(1) the match rate, or the rate at which members of independently surveyed households in a sample of block clusters (the P-sample) matched to census enumerations, calculated separately for population groups (poststrata) and weighted to population totals, and

(2) the correct enumeration rate, or the rate at which census enumerations in the sampled block clusters (the E-sample) were correctly included in the census (including both matched cases and nonmatched correct enumerations), calculated separately for poststrata and weighted to population totals.[1]

[1]The E-sample by design excluded some census cases in the A.C.E. block clusters (see Appendix E.1.e and E.3).

Other things equal, the higher the match rate, the *lower* will be the DSE population estimate and the estimated net undercount in the census. Conversely, the more nonmatches, the higher will be the DSE population estimate and the estimated net undercount. In contrast, the higher the correct enumeration rate, the *higher* will be the DSE population estimate and the estimated net undercount. Conversely, the more erroneous enumerations, the lower will be the DSE population estimate and the estimated net undercount (for how this result obtains, refer to Equation 5.1 in Section 5–A).

The A.C.E. and PES design and estimation focused on estimating the net undercount and not on estimating the numbers or types of gross errors of erroneous enumerations (overcounts) or of gross omissions. There are not widely accepted definitions of components of gross error, even though such errors are critically important to analyze in order to identify ways to improve census operations. Some types of gross errors depend on the level of geographic aggregation. For example, assigning a census household to the wrong small geographic area (geocoding error) is an erroneous enumeration for that area (and an omission for the correct area), but it is not an erroneous enumeration (or an omission) for larger areas. Also, the original A.C.E. design, similar to the PES, did not permit identifying duplicate census enumerations as such outside a ring or two of blocks surrounding a sampled block cluster. On balance, about one-half of duplicate enumerations involving different geographic areas should be classified as an "other residence" type of erroneous enumeration at one of the two addresses because the person should have been counted only once, but this balancing may not be achieved in practice.

Several aspects of the original A.C.E. design were modified from the PES design in order to improve the timeliness and reduce the variance and bias of the results (see Section 5–D.1). Some of these changes were clearly improvements. In particular, the larger sample size (300,000 households in the A.C.E. compared with 165,000 households in the PES) and the reduction in variation of sampling rates considerably reduced the variance of the original A.C.E. estimates compared with the PES estimates (see Starsinic et al., 2001). The coefficient of variation for the originally estimated coverage correction factor of 1.012 for the total population in 2000 was 0.14 percent, a reduction of 30 percent from the comparable coefficient of variation in 1990.[2] (The 0.14 percent coefficient of variation translates into a standard error of about 0.4 million people in the original DSE total household population estimate of 277.5 million.) The coefficients of variation for the originally estimated coverage correction factors for Hispanics and non-Hispanic blacks were 0.38 and 0.40 percent, respectively, in 2000, reduced from 0.82 and 0.55 percent, respectively, in 1990 (Davis, 2001:Tables E-1, F-1). However, some poststrata had coefficients of variation as high as 6 percent in 2000, which translates into a large confidence interval around the estimate of the net undercount for these poststrata and for any geographic areas in which they are a large proportion of the population.

Three other improvements in the A.C.E. design deserve mention. First, an initial housing unit match of the independent P-sample address listing with the Master Address File (MAF) facilitated subsequent subsampling, interviewing, and matching operations. Second, the use of computer-assisted interviewing—by telephone in the first wave—facilitated timeliness of the P-sample data, which had the positive effect of reducing the percentage of movers compared with the 1990 PES (see Section 6–A.3). Third, improved matching technology and centralization of matching operations probably contributed to a higher quality of matching than achieved in 1990 (see Section 6–A.8).

Another innovation—the decision to target the search for matches and correct enumerations in surrounding blocks more narrowly in the A.C.E. than in the PES—was originally suspected of having contributed balancing errors to the original (March 2001) DSE estimates, but subsequent evaluation allayed that concern (see Section 6–A.9). The treatment of movers was more complex in the A.C.E. than in the 1990 PES, but, primarily because there were proportionately fewer movers in the A.C.E. compared with the PES, on balance, movers had no more effect on the dual-systems estimates for 2000 than on those for 1990 (see Section 6–A.3).

In retrospect, the decision to exclude the group quarters population from the A.C.E. universe (see Section 5–D.1) was unfortunate, as it precluded the development of coverage estimates for group

[2]The coefficient of variation (CV) is the standard error of an estimate as a percentage of the estimate. The coverage correction factor, which would be used in implementing an adjustment, is the dual-systems estimate divided by the census count (including whole-person imputations and late additions); see Section 5–A.

quarters residents, who appear to have been poorly enumerated (see Section 4–F). Finally, a flaw in the planning for not only the A.C.E., but also the census itself, was the failure to anticipate the extent to which certain groups of people (e.g., college students, prisoners, people with two homes) would be duplicated in more than one census record (see Section 6–A.7).

6–A.2 Conduct and Timing

Overall, the original A.C.E. was well executed in terms of timely and well-controlled address listing, P-sample interviewing, matching, follow-up, and original estimation. Although the sample size was twice as large as that fielded in 1990, the A.C.E. was carried out on schedule and with only minor problems that necessitated rearrangement or modification of operations after they had been specified. Mostly, such modifications involved accommodation to changes in the MAF that occurred in the course of the census. For example, the targeted extended search (TES) procedures had to be modified to handle deletions from and additions to the MAF that were made after the determination of the TES housing unit inventory (Navarro and Olson, 2001:11).

Some procedures proved more useful than had been expected. In particular, the use of the telephone (see Appendix E.2) enabled P-sample interviewing to begin April 24, 2000, whereas P-sample interviewing for the PES did not begin until June 25, 1990. All A.C.E. processes, from sampling through estimation, were carried out according to well-documented specifications, with quality control procedures (e.g., reviews of the work of clerical matchers and field staff) implemented at appropriate junctures.

6–A.3 Defining the P-Sample: Treatment of Movers

The A.C.E. P-sample, partly because of design decisions made for the previously planned Integrated Coverage Measurement Program (see Section 5–D.1), included three groups of people and not two as in the 1990 PES. The three groups were: nonmovers who lived in a P-sample housing unit on Census Day and on the A.C.E. interview day; outmovers who lived in a P-sample housing unit on Census Day but had left by the A.C.E. interview day; and inmovers who moved into a P-sample housing unit between Census Day and the A.C.E. interview day. In the dual-systems estimation for each poststratum (population group), the number of matched movers was calculated by applying the estimated match rate for outmovers to the weighted number of inmovers.[3] This procedure, called PES-C, assumed that inmovers would be more completely reported than outmovers. The Bureau also anticipated that it would be easier to ascertain the Census Day residence status of outmovers than to search nationwide for the Census Day residence of inmovers, as was done in the 1990 PES using the PES-B procedure.

An analysis of movers in the A.C.E. P-sample conducted by the Census Bureau in summer 2001 supported the assumption of more complete reporting of inmovers: the total weighted number of outmovers was only two-thirds (0.66) the total weighted number of inmovers (the outmover/inmover ratio varied for broad population groups from less than 0.51 to more than 0.76—Liu et al., 2001:App.A). A subsequent evaluation found little effect on the dual-systems population estimates of using inmovers to estimate the number of movers (Keathley, 2002).

Noninterview and missing data rates were substantially higher for outmovers compared with inmovers and nonmovers (see Sections 6–A.5 and 6–A.6), so one might have expected to see an increase in variance of the dual-systems estimates from using outmovers to estimate match rates from the P-sample. Yet Liu et al. (2001) found that movers had no more and probably less of an effect on the dual-systems estimates in 2000 than in 1990. A primary reason for this result is that the percentage of movers among the total population was lower in the A.C.E. than in the PES—using the number of inmovers in the numerator, the A.C.E. mover rate was 5.1 percent, compared with a mover rate of 7.8 percent in the PES. In turn, the lower A.C.E. mover rate resulted from the 2-month head start that was achieved by telephone interviewing in the A.C.E. (see Section 6–A.2). The mover rate for A.C.E. cases interviewed after June 25, 2000, was comparable to the PES mover rate (8.2 and 7.8 percent, respectively); the mover rate for A.C.E. cases interviewed before June 25, 2000, was only 2.1 percent (Liu et al., 2001:5).

[3]For 63 poststrata with fewer than 10 outmovers, the weighted number of outmovers was used instead.

6–A.4 Defining the E-Sample: Exclusion of "Insufficient Information" Cases

Dual-systems estimation in the census context requires that census enumerations be excluded from the E-sample when they have insufficient information for matching and follow-up (so-called *II*s—see Section 5–A). The 2000 census had almost four times as many *II*s as the 1990 census—8.2 million, or 2.9 percent of the household population, compared with 2.2 million or 0.9 percent of the population. In 2000 5.8 million people fell into the *II* category because they were whole-person imputations (types 1–5, as described in Section 4–D); another 2.4 million people were *II*s because their records were not available in time for the matching process. These people were not in fact enumerated late; rather, they represented records that were temporarily deleted and subsequently reinstated on the census file as part of the special MAF unduplication process in summer–fall 2000 (see Section 4–E). In 1990 only 1.9 million whole-person imputations and 0.3 million late additions from coverage improvement programs fell into the *II* category.

Because the phenomenon of reinstated cases in the 2000 census was new and the number of such cases was large, the Bureau investigated the possible effects of their exclusion from the E-sample on the dual-systems estimate. Hogan (2001b) demonstrated conceptually that excluding the reinstated people would have little effect so long as they were a small percentage of census correct enumerations or their A.C.E. coverage rate (ratio of matches to all correct enumerations) was similar to the E-sample coverage rate. To provide empirical evidence, a clerical matching study was conducted in summer 2001 of reinstated people whose census records fell into an evaluation sample of one-fifth of the A.C.E. block clusters (Raglin, 2001). This study found that 53 percent of the reinstated records in the evaluation sample duplicated another census record (and, hence, had no effect on the DSE), 25 percent matched to the P-sample, and 22 percent were unresolved (such a large percentage resulted from the infeasibility of follow-up to obtain additional information). Using a range of correct enumeration rates for the unresolved cases, the analysis demonstrated that the exclusion of reinstated records from the E-sample had a very small effect on the DSE for the total population (less than one-tenth of 1 percent). Moreover, because total and matched reinstated cases were distributed in roughly the same proportions among age, sex, race/ethnicity, and housing tenure groups, their exclusion from the E-sample had similar (negligible) effects on the DSE estimates for major poststrata.

Nonetheless, the large number of *II*s in 2000 cannot be ignored in understanding patterns of net undercount. Although reinstated cases accounted for roughly the same proportion of each major poststratum group (about 1 percent) in 2000, whole-person imputations accounted for higher proportions of historically undercounted groups, such as minorities, renters, and children, than of historically better counted groups. We consider the role of whole-person imputations in helping to account for the measured reduction in net undercount rate differences among major population groups in 2000 from 1990 in Section 6–C.1.

6–A.5 Household Noninterviews in the P-Sample

The P-sample survey is used to estimate the match rate component of the dual-systems estimation formula. A small bias in the match rate can have a disproportionately large effect on the estimated net undercount (or overcount) because coverage error is typically so small relative to the total population (1–2 percent or less). To minimize variance and bias in the estimated match rate, it is essential that the A.C.E. successfully interview almost all P-sample households and use appropriate weighting adjustments to account for noninterviewed households.

Interview/Noninterview Rates

Overall, the A.C.E. obtained interviews from 98.9 percent of households that were occupied on the day of interview. This figure compares favorably with the 98.4 percent interview rate for the 1990 PES.[4] However, the percentage of occupied households as of Census Day that were successfully interviewed in A.C.E. was somewhat lower—97 percent, meaning that a weighting adjustment had to account for the remaining 3 percent of noninterviewed households.

[4]These percentages are unweighted; they are about the same as weighted percentages for 2000. Weighted percentages are not available for 1990 (see Cantwell et al., 2001).

The lower interview rate for Census Day households was due largely to the difficulty of finding a respondent for housing units in the P-sample that were entirely occupied by people who moved out between the time of the census and the A.C.E. interview (outmovers). Such units were often vacant, and it was not always possible to interview a neighbor or landlord who was knowledgeable about the Census Day residents. The interview rate for outmover households was 81.4 percent. Such households comprised 4 percent of Census Day occupied households in the P-sample.

Noninterview Weighting Adjustments

Two weighting adjustments—one for the A.C.E. interview day and one for Census Day—were calculated so that interviewed households would represent all households that should have been interviewed. Each of the two weighting adjustments was calculated separately for households by type (single-family unit, apartment, other) within block cluster.

For Census Day, what could have been a relatively large noninterview adjustment for outmover households in a block cluster was spread over all interviewed Census Day households in the cluster for each of the three housing types. Consequently, adjustments to the weights for interviewed households were quite low, which had the benefit of minimizing the increase in the variance of A.C.E. estimates due to differences among weights: 52 percent of the weights were not adjusted at all because all occupied households in the adjustment cell were interviewed; for another 45 percent of households, the weighting adjustment was between 1.0 and 1.2 (Cantwell et al., 2001:Table 2).

Evaluation

Although the P-sample household noninterview adjustments were small, a sensitivity analysis determined that alternative weighting adjustments could have a considerable effect on the estimated value of the DSE for the national household population. Three alternative noninterview adjustments were tested: assigning weights on the basis of characteristics other than those used in the original A.C.E. estimation, assigning weights only to late-arriving P-sample interviews, and replicating the weights for missing interviews from nearby households. All three alternatives produced weighted total household population estimates that were higher than the original (March 2001) DSE estimate (Keathley et al., 2001:Table 4). Two of the alternative estimates exceeded the original estimate by 0.5–0.6 million people, which translates into an added 0.2 percentage points of net undercount on a total household population of 277.2 million. The differences between these two estimates and the original estimate also exceeded the standard error of the original estimate, which was 0.4 million people.

6–A.6 Missing and Unresolved Data in the P-Sample and E-Sample

Missing and unresolved person data can bias the estimated P-sample match rate, the estimated E-sample correct enumeration rate, or both rates. Imputation procedures used to fill in missing values can also add bias and variance, so achieving high-quality P-sample and E-sample data is critical for dual-systems estimation.

Missing Characteristics Needed for Poststratification

Overall rates of missing characteristics data in the P-sample and E-sample were low, ranging between 0.2 and 3.6 percent for age, sex, race, Hispanic origin, and housing tenure. Missing data rates for most characteristics were somewhat higher for the E-sample than for the P-sample. Missing data rates for the 2000 A.C.E. showed no systematic difference (up or down) from the 1990 PES; see Table 6.1.

As would be expected, missing data rates in the P-sample were higher for proxy interviews, in which someone outside the household supplied information, than for interviews with household members; see Table 6.2. By mover status, missing data rates were much higher for outmovers than for nonmovers and inmovers, which is not surprising given that 73.3 percent of interviews for outmovers were obtained from proxies, compared with only 2.9 and 4.8 percent of proxy interviews for nonmovers and inmovers, respectively. Even "nonproxy" interviews for outmovers could have been from household members who did not know the outmover.

Table 6.1 Missing Data Rates for Characteristics, 2000 A.C.E. and 1990 PES P-Sample and E-Sample (weighted)

	Percentage of People with Imputed Characteristics			
	2000 A.C.E.		1990 PES	
Characteristic	P-Sample	E-Sample	P-Sample	E-Sample
Age	2.4	2.9	0.7	2.4
Sex	1.7	0.2	0.5	1.0
Race	1.4	3.2	2.5	11.8
Hispanic Origin	2.3	3.4	—	—
Housing Tenure	1.9	3.6	2.3	2.5
Any of Above	5.4	10.4	—	—

NOTES: Accuracy and Coverage Evaluation (A.C.E.) E-sample imputations were obtained from the imputations performed on the census records; Post-Enumeration Survey (PES) E-sample imputations were performed specifically for the E-sample. A.C.E. E-sample "edits" (e.g., assigning age on the basis of the person's date of birth, or assigning sex from first name) are not counted as imputations here. The base for the A.C.E. P-sample imputation rates includes nonmovers, inmovers, and outmovers, including people who were subsequently removed from the sample as nonresidents on Census Day. Excluded from the base for the A.C.E. P-sample and E-sample imputation rates are people eligible for the targeted extended search who were not selected for the targeted extended search sample and who were treated as noninterviews in the final weighting. —, not available.

SOURCE: Cantwell et al. (2001:Tables 3b, 3c).

Table 6.2 Percentage of 2000 A.C.E. P-Sample People with Imputed Characteristics, by Proxy Interview and Mover Status (weighted)

	Percentage of People with Imputed Characteristics				
Characteristic	Household Interview	Proxy Interview	Nonmover	Inmover	Outmover
Age	2.1	7.9	2.3	2.3	6.0
Sex	1.5	4.2	1.7	0.4	3.4
Race	1.0	8.7	1.2	1.3	8.0
Hispanic Origin	1.8	11.0	2.1	0.8	9.0
Housing Tenure	1.7	5.2	1.9	0.4	2.4
Any of Above	4.4	21.9	5.0	3.7	17.4
Percent of Total P-Sample	94.3	5.7	91.7	4.8	3.4

NOTES: See notes to Table 6.1.

SOURCE: Cantwell et al. (2001:Table 3b).

P-sample imputations were performed separately for individual missing characteristics after all matching and follow-up had been completed. For example, tenure on the P-sample was imputed by using tenure from the previous household of the same type (e.g., single-family home) with tenure reported, while race and ethnicity were imputed when possible from the distribution of race and ethnicity of other household members or from the distribution of race and ethnicity of the previous household with these characteristics reported (see Cantwell et al., 2001). Imputations for missing characteristics in the E-sample records were obtained from those on the census data file (see Section 7–B). Because the overall rates of missing data were low, the imputation procedures had little effect on the distribution of individual characteristics (Cantwell et al., 2001:24–26). However, given the somewhat different procedures for the P-sample and the E-sample, imputation could misclassify people by poststrata and contribute to inconsistent poststrata classification for matching P-sample and E-sample cases (see Section 6–A.10).

P-sample Unresolved Residence Status

The use of the PES-C procedure to define the P-sample for estimating match rates (see Sections 5–D.1 and 6–A.3) made it necessary to impute a residence status probability to P-sample nonmover and outmover cases whose status as Census Day residents in an A.C.E. block cluster was not resolved after matching and follow-up. (Mover status was assigned before follow-up, which explains why "nonmovers" and "outmovers" could be reclassified as Census Day nonresidents or as having unresolved residence status.) On a weighted basis, unresolved residence cases accounted for 2.2 percent of all the cases considered for inclusion in the Census Day P-sample. Outmovers accounted for 29 percent of P-sample cases with unresolved residence status, although they were less than 4 percent of the total P-sample (Cantwell et al., 2001:Table 5b).

The imputation procedure assigned an average residence probability to each unresolved case taken from one of 32 cells defined by owner/renter, non-Hispanic white/other, and before follow-up match status (eight categories). After imputation the percentage of Census Day residents among the original Census Day P-sample dropped slightly from 98.2 percent of resolved cases to 97.9 percent of all cases, because the imputation procedure assigned lower residence probabilities to unresolved cases (77.4 percent overall).[5]

P-sample Unresolved Match Status

The weighted percentage of P-sample cases with unresolved match status was only 1.2 percent. The denominator for the percentage is P-sample nonmovers and outmovers who were confirmed Census Day residents or had unresolved residence status; confirmed non-Census Day residents were dropped from the P-sample at this point. This percentage compares favorably with the 1.8 percent of cases with unresolved match status in the 1990 PES. Very little was known about the A.C.E. P-sample people with unresolved match status; 98 percent of them lacked enough reported data for matching (i.e., they lacked a valid name or at least two characteristics or both).

The imputation procedure assigned an average residence probability to each unresolved case taken from one of 16 cells defined by resolved/unresolved residence status, nonmover/outmover status, housing unit a match/not a match, and person had one or more characteristics imputed/no characteristics imputed (Cantwell et al., 2001:Table 9). After imputation, the percentage of matches dropped slightly, from 91.7 percent of resolved cases (matches and nonmatches) to 91.6 percent of all cases because the imputation procedure assigned lower match status probabilities to unresolved cases (84.3 percent overall).

E-sample Unresolved Enumeration Status

The weighted percentage of E-sample cases with unresolved enumeration status was 2.6 percent, slightly higher than the comparable 2.3 percent for the 1990 PES. Most of the unresolved cases (89.4 percent) were nonmatches for which field follow-up did not resolve whether their enumeration was correct or erroneous.

The imputation procedure assigned an average correct enumeration probability to each unresolved case taken from one of 64 cells defined by 13 categories of before-follow-up match status (3 of which were tabulated separately for non-Hispanic whites and

[5]This figure is a correction from the original number in Cantwell et al., 2001:Table 8).

others), housing unit a match/not a match, and person had one or more characteristics imputed/no characteristics imputed (Cantwell et al., 2001:Table 10). After imputation, the percentage of correct enumerations dropped slightly, from 95.5 percent of resolved cases (correct and erroneous enumerations) to 95.3 percent of all cases because the imputation procedure assigned lower correct enumeration probabilities to unresolved cases (76.2 percent overall).

Evaluation

A sensitivity analysis determined that alternative procedures for imputing P-sample residence and match status probabilities and E-sample correct enumeration status probabilities could have a considerable effect on the estimated value of the DSE for the national household population, particularly when combined with alternative procedures for making P-sample household noninterview adjustments (see Section 6–A.5).[6] One of the alternative imputation procedures substituted multivariate logistic regressions for the average cell values used in the original A.C.E. Another procedure, which assumed that unresolved cases differed significantly from resolved cases (what is termed nonignorable missingness), used 1990 PES data to develop alternative (lower) probabilities of residence, match, and correct enumeration status. These probabilities are illustrative; there is no evidence for their reasonableness compared with the probabilities used in the original A.C.E.

The results of the sensitivity analysis demonstrate the difference that alternative procedures could make. Thus, for all 128 combinations of noninterview adjustment and imputation procedures, about one-third of them differed from the average DSE population estimate by more than plus or minus 0.7 million household members; the remaining two-thirds differed by less than this amount (Keathley et al., 2001:2, as revised in Kearney, 2002:5).

6–A.7 Accuracy of Household Residence Information

For dual-systems estimation to produce highly accurate population estimates, it is critical not only for there to be very low household

[6]A sensitivity analysis was not conducted for imputations for missing demographic characteristics.

and item nonresponse rates in the P-sample and E-sample, but also for household composition to be accurately reported. Two factors are critical to accurate reporting: first, the ability of the A.C.E. questionnaires and interviewing procedures for P-sample interviewing and follow-up of nonmatched P-sample and E-sample cases to elicit the Census Day residence status of household members; and, second, the willingness of respondents to answer the questions as they were intended.

As an example, a household in the census that was part of the E-sample may have claimed a college student or an institutionalized family member as a household member even though the person was enumerated in his or her group quarters according to census residence rules. The result would be a duplicate census enumeration. In the case when the household was missed by the P-sample, the matching and follow-up process should have identified the nonduplicated E-sample household residents as correct (nonmatched) enumerations and the duplicated college student as having been erroneously enumerated at the household address in the census. If, however, the household persisted in claiming the student as a household member, then the A.C.E. would incorrectly classify him or her as a correct (nonmatched) enumeration, thereby overstating the correct enumeration rate and the DSE estimate of the population. This example and one other involving undetected duplicates of census enumerations with nonmatched P-sample cases are described, along with their effects, in Box 6.1.

The A.C.E. questionnaires were improved over the PES questionnaires, and computer-assisted interviewing ensured that interviewers asked all of the questions as they were written. However, Census Bureau staff worried that the A.C.E. interviewing might not have ascertained Census Day household membership accurately in many cases because the original A.C.E. estimated only 4.6 million duplicate and "other residence" erroneous enumerations whereas the PES estimated 10.7 million of these types of erroneous enumerations (see Anderson and Fienberg, 2001:Table 2). Indeed, the Evaluation Follow-Up and Person Duplication Studies conducted by the Census Bureau in summer 2001 provided evidence that the A.C.E. failed to detect numerous instances in which census respondents listed one or more residents who should not have been counted as part of the household on Census Day. Consequently, the original

Box 6.1 Alternative Treatment of Duplicate Census Enumerations, Two Examples

(1) **Census Household-to-Group Quarters Duplication; Household in E-Sample, not P-Sample**: College student is enumerated in group quarters (college dormitory) and by parents at home.

Proper treatment in the A.C.E. when parents' household is not in the P-sample:

E-sample follow-up of nonmatched household should classify the parents as correct enumerations and the student as an "erroneous enumeration, other residence" (i.e., should have been enumerated at the college location only). In this instance, the A.C.E. would not label the college student as a duplicate because it would not know of the group quarters enumeration; the label would be "other residence," meaning that the person should have been enumerated at the group quarters. Regardless, the A.C.E. would correctly classify the enumeration as erroneous.

Erroneous treatment:

Household persists in claiming the student in E-sample follow-up, so all three household members are classified as correct (nonmatched) enumerations.

Effect of erroneous treatment on DSE:

Extra "correct" enumeration raises the correct enumeration rate, which (incorrectly) raises the DSE estimate of the population and net undercount.

(2) **P-Sample Resident Nonmover Household-to-Census Household Duplication Outside A.C.E. Search Area:** P-sample household duplicates census enumeration outside its block cluster and ring of surrounding blocks.

Proper treatment in the A.C.E.:

The P-sample interview should have reclassified the household as comprising inmovers (and hence not eligible for estimating the match rate) or dropped it from the sample as having been wrongly assigned to an A.C.E. block cluster.

Erroneous treatment:

The household is retained in the P-sample as a nonmover resident and used to contribute to the numerator as well as the denominator of the match rate depending on whether it matches a census enumeration inside the A.C.E. search area.

Effect of erroneous treatment on the DSE:

Depends on match status of erroneously retained P-sample resident nonmover cases. If predominantly matches, their inclusion (incorrectly) raises the match rate and lowers the DSE estimate of the population and net undercount.

(March 2001) A.C.E. underestimated duplicate enumerations in the census and correspondingly overestimated correct enumerations. It also overestimated P-sample Census Day residents, particularly nonmatches. The net effect (assuming an accurate estimate of omissions) was to overstate the correct enumeration rate, understate the match rate, and overstate the DSE estimate of the population by about 6.3 million people. Correcting this overstatement (before an adjustment for undercounting of men relative to women) would have produced an estimated net *overcount* of the population of 3 million people or 1.1 percent of the household population (see U.S. Census Bureau, 2003c:Table 12; see also Sections 6–B and 6–C.2).

6–A.8 Quality of Matching

The A.C.E. (and PES) involved a two-stage matching process. The first stage of matching occurred after P-sample interviewing; it began with a computer match followed by a clerical review of possible matches and nonmatches in order to establish an initial match status (P-sample) or enumeration status (E-sample) for as many cases as possible. The second stage occurred after follow-up of specified nonmatched and unresolved cases to try to resolve their status using additional information from the follow-up interviews. The accuracy of the matching can be no better than the accuracy of the underlying data about household composition and residence (as discussed in Section 6–A.7). Assuming accurate information, the question is the quality of the matching itself.

Examination of data from the original (production) A.C.E. matching provides indicators that the quality was high. Specifically, initial match status codes were rarely overturned at a subsequent stage of matching: clerks confirmed a high percentage (93 percent) of computer-designated possible matches as matches; technicians and analysts who reviewed clerical matches rarely overturned the clerks' decisions, and field follow-up most often confirmed the before-follow-up match code or left the case unresolved.[7] Because of the dependent nature of the production matching, however, such

[7]From tabulations by panel staff of the P-Sample and E-Sample Person Dual-System Estimation Output Files, provided to the panel February 16, 2001 (see U.S. Census Bureau, 2001b). Tabulations weighted using TESFINWT.

indicators do not answer the question of whether the final match status codes were correct.

In summer 2001, Census Bureau analysts completed a Matching Error Study to evaluate the quality of the A.C.E. matching criteria and procedures (Bean, 2001). The Matching Error Study involved an independent rematch in December 2000 by highly trained matching staff (technicians and analysts) of all of the P-sample and E-sample cases in one-fifth of the A.C.E. block clusters (2,259 clusters). The Matching Error Study used the original A.C.E. data on household composition and residence and not any data from evaluation studies, so that it could measure the extent of matching error only, not confounded with measurement error. The study assumed that agreement of the original (production) and rematch codes, or agreement of an analyst in conflicting cases with either the production or the rematch code would produce match codes as close to truth as was possible. The study also assumed that the production matching and the evaluation rematching were independent—Matching Error Study rematchers did not review clusters that they worked on during the original A.C.E. and did not have access to the original match codes. Bean (2001:3) notes some minor ways in which independence could have been compromised.

A comparison of the results of the Matching Error Studies for the A.C.E. (Bean, 2001) and the PES (Davis and Biemer, 1991a,b) provides evidence of improved matching quality in the A.C.E. over the PES. For the four final P-sample match codes (match, nonmatch, remove from the P-sample, and unresolved), the A.C.E. matching error study estimated only a 0.4–0.5 percent gross difference rate compared with a 1.5 percent gross difference rate for the PES.[8] The net difference rate was also reduced in the A.C.E. from the PES (0.4 and 0.9 percent, respectively).[9] Gross and net difference rates for classification of E-sample cases (correct enumeration, erroneous enumeration, unresolved) were also substantially reduced in the A.C.E. from the PES (0.5–0.6 percent and 2.3 percent gross difference rates, respectively; 0.2 and 1.1 percent net difference rates, respec-

[8]The gross difference rate is the proportion of cases whose match codes differ in the production and the rematch.

[9]The net difference rate is the sum of the absolute differences between the production and rematch totals for all four match codes divided by the population total.

tively). A measure of proportionate error for matches and correct enumerations for 16 aggregated poststratum groups showed smaller errors and less variation in the degree of error among groups in the A.C.E. than in the PES (Bean, 2001:Tables 5b, 5c).

Despite the improved quality of matching in the A.C.E. compared with the PES, matching error still affected the original DSE population estimates for 2000. It significantly decreased the original P-sample match rates for the nation as a whole and for 2 of 16 aggregated poststratum groups (minority and nonminority renters in large or medium metropolitan mailback areas with high return rates). It did not, however (assuming correct information on residence), significantly affect the original E-sample correct enumeration rates for the nation or any of the 16 groups. The effect of matching error on the ratio of the match rate to the correct enumeration rate resulted in an overstatement of the 2000 DSE total population estimate by about 0.5 million people (Bean, 2001:20), which amounts to an overstatement of the net undercount of about 0.2 percentage points.

6–A.9 Targeted Extended Search

The TES operation in the A.C.E. was designed to reduce the variance and bias in match and correct enumeration rates that could result from geocoding errors (i.e., assignment of addresses to the wrong block) in the census or in the P-sample address listing. In a sample of block clusters for which there was reason to expect geocoding errors (2,177 of 6,414 such clusters), the clerical search for matches of P-sample and census enumerations and for correct E-sample enumerations was extended to one ring of blocks surrounding the A.C.E. block cluster. Sampling was designed to make the search more efficient than in 1990, as was targeting the search in some instances to particular blocks (see Appendix E.3.b).

For the P-sample, only people in households that did not match an E-sample address (4.7 percent of total P-sample cases that went through matching) were searched in the sampled block clusters. On the E-sample side, only people in households identified as geocoding errors (3 percent of total E-sample cases) were searched in the sampled block clusters. Weights were assigned to the TES persons in the sampled block clusters to adjust for the sampling. Correspondingly, persons who would have been eligible for TES but were not in a sampled block cluster were assigned a zero weight.

The TES had the desired effect of reducing the variance of the DSE estimates for poststrata. The reduction in the average and median coefficient of variation (the standard error of an estimate as a percentage of the estimate) for poststrata was 22 percent, similar to an average reduction of 20 percent for the nationwide extended search operation in 1990 (Navarro and Olson, 2001:7).

The TES operation in the A.C.E. was methodologically more complex than the corresponding operation in the 1990 PES. At the time of the decision not to use the original DSE estimates of the population to adjust census data for redistricting in March 2001, the Census Bureau cited concerns that the TES may have been unbalanced, thereby introducing bias into the DSE. Suggestive of an imbalance, which could occur if the P-sample and E-sample search areas were not defined consistently for the TES, was the larger increase in the P-sample match rate (3.8 percentage points) compared with the E-sample correct enumeration rate (2.9 percentage points) (Navarro and Olson, 2001:Table 1). Such an imbalance may also have occurred in 1990, when the extended search increased the P-sample match rate by 4.1 percentage points and the E-sample correct enumeration rate by 2.3 percent. A follow-up study to the 1990 census was not able to determine whether balancing error had occurred (Bateman, 1991).

A subsequent evaluation, which used data from two TES follow-up studies that rechecked certain kinds of housing units (Adams and Liu, 2001:i), determined that the larger increase in the TES in the P-sample match rate compared with the E-sample correct enumeration rate was due to P-sample geocoding errors and E-sample classification errors that did not affect the DSE. P-sample geocoding errors were the primary explanation; they occurred when P-sample address listers mistakenly assigned addresses from surrounding blocks to A.C.E. block clusters. When the original A.C.E. clerk did not find matches for these cases in the A.C.E. block cluster because there were no corresponding census addresses, then a search for matches in the surrounding ring was likely to be successful. If the TES had not been conducted, these matches would have been missed, resulting in an underestimate of the P-sample match rate and an overestimate of the DSE population estimate and the net undercount.

The TES evaluation study did find about 246,000 P-sample nonmatches and about 195,000 E-sample correct enumerations located beyond the surrounding blocks for a search. These cases should have been treated as nonresidents and geocoding errors, respectively; their inclusion as nonmatched residents and correct enumerations resulted in a slight overestimate in the DSE population estimate.

6–A.10 Poststratification

Poststratification is an important aspect of dual-systems estimation. Because research suggests that the probabilities of being included in the census or in the P-sample vary by individual characteristics, it is important to classify P-sample and E-sample cases into groups or strata for which coverage probabilities are as similar as possible within the group and as different as possible from other groups. The DSE then is performed stratum by stratum.

Counterbalancing the need for finely defined poststrata are two considerations: each poststratum must have sufficient sample size for reliable estimates, and the characteristics used to define the poststrata should be consistently measured between the P-sample and the E-sample. As an example, a census respondent whose household was in the E-sample may have reported age 30 for a household member when a different respondent for the same household in the P-sample reported the person to be age 29. The matched person, then, would contribute to the P-sample match rate for the 18- to 29-year-old poststrata and to the E-sample correct enumeration rate for the 30- to 49-year-old poststrata. Misclassification can be consequential if the proportions misclassified are large and if the coverage probabilities vary greatly for the affected poststrata. Finally, a consideration for the Census Bureau for 2000 was the need to define the poststrata in advance and to specify categories for which direct estimates could be developed without the complications of the modeling that was used in 1990.

Taking all these issues into account, the Census Bureau specified 448 poststrata in advance of the original A.C.E. (collapsed to 416 in the estimation, see Table E.3 in Appendix E). The somewhat larger number of A.C.E. poststrata, compared with the 357 poststrata used for the revised 1990 PES estimation, was made possible by the larger

A.C.E. sample size.[10] On the face of it, the original A.C.E. poststratification seemed reasonable in terms of using characteristics (age, sex, race, ethnicity, housing tenure, mail return rate) that historically have related to coverage probability differences. There was some inconsistency of classification by poststrata between the P-sample and E-sample in the A.C.E., although whether the level of inconsistency was higher or lower than in 1990 cannot be determined because of the unavailability of data for 1990 matched cases. Overall, 4.7 percent of A.C.E. matched cases (unweighted) were inconsistently classified as owner or renter; 5.1 percent were inconsistently classified among age and sex groups, and 3.9 percent were inconsistently classified among race/ethnic domains (Farber, 2001a:Table 1).

Among race/ethnicity domains, inconsistent cases as a percentage of E-sample matches showed wide variation, ranging from 1.5 percent for American Indians and Alaska Natives on reservations, to 18.3 percent for Native Hawaiians and Pacific Islanders, to 35.7 percent for American Indians and Alaska Natives off reservations. The latter inconsistency rate is very high. The major factor is that a large number of people (relative to the Native American population) who identified themselves as non-Hispanic white or other race in one sample identified themselves as American Indian or Alaska Native off reservations in the other sample (see Section 8–C.2). The effect was to lower the coverage correction factor for the latter group below what it would have been had there been no inconsistency. However, the coverage correction factor would have been lower yet for American Indians and Alaska Natives off reservations if they had been merged with the non-Hispanic white and other races domain. The reverse flow of American Indians and Alaska Natives identifying themselves as non-Hispanic whites or other races had virtually no effect on the coverage correction factor for the latter group, given its very large proportion of the population.

[10]The original 1990 PES estimation used 1,392 poststrata together with a composite estimation procedure to smooth the resulting DSEs. The much smaller revised set of 1990 poststrata were developed by analyzing census results that had become available (e.g., mail return rates, imputation rates, crowding) to determine which characteristics that could be used for poststratification best explained variations in those results. The Census Bureau also analyzed 1990 data to determine the original A.C.E. poststratification (Haines, 1999a,b).

6–A.11 Original A.C.E.: Summary of Findings

From the extensive evaluations of the original A.C.E. data conducted by Census Bureau staff, we draw several conclusions. First, coverage evaluation using the dual-systems estimation method is a highly complex effort in the census context. Many things have to go very well, and errors need to be very small, particularly as the quantity being estimated—net undercount—is so small relative to the population. Second, as a major data collection, data processing, and estimation operation, the A.C.E. in fact went extremely well. It was conducted in a timely, controlled manner. Evaluation studies of matching error, the targeted extended search, poststratification inconsistency, and treatment of movers in the A.C.E. found generally improved performance over the PES and only small biasing effects on the DSE estimates—at least for national totals and aggregates of poststrata. Percentages of household noninterviews and cases with unresolved residence, match, or enumeration status were small (although evaluation found significant effects of plausible alternative reweighting and imputation methods on the DSE estimates).

Dwarfing all of these generally positive outcomes for the original A.C.E., however, were the findings from the Evaluation Follow-up and Person Duplication Studies of substantial underestimation of duplicate and other census erroneous enumerations and overestimation of P-sample residents, particularly nonmatched cases (see Sections 6–B.1 and 6–B.2). There was no counterpart of the Person Duplication Studies for the 1990 PES.

> **Finding 6.1:** The 2000 Accuracy and Coverage Evaluation (A.C.E.) Program operations were conducted according to clearly specified and carefully controlled procedures and directed by a very able and experienced staff. In many respects, the A.C.E. was an improvement over the 1990 Post-Enumeration Survey, achieving such successes as high response rates to the P-sample survey, low missing data rates, improved quality of matching, low percentage of movers due to more timely interviewing, and substantial reductions in the sampling variance of coverage correction factors for the total population and important population groups. However, inaccurate

reporting of household residence in the A.C.E. (which also occurred in the census itself) led to substantial underestimation of duplicate enumerations in 2000 in the original (March 2001) A.C.E. estimates.

6–B A.C.E. REVISION II ESTIMATION DATA, METHODS, AND RESULTS

The A.C.E. Revision II process began in spring 2002 and was completed in early 2003. It had five major goals, which we discuss in turn: (6–B.1) improve estimates of erroneous census enumerations; (6–B.2) improve estimates of census omissions; (6–B.3) develop new models for missing data; (6–B.4) enhance the estimation poststratification; and (6–B.5) consider adjustment for correlation bias (phrased simply, the assumption that some groups are disproportionately missed in both the census and the independent P-sample survey). We consider the combined results of these efforts in Section 6–B.6 and the findings from analyses of error in the Revision II estimates in Section 6–B.7. The Revision II work did not collect new data beyond the additional data collected in evaluation studies conducted in 2001. It used a combination of the original A.C.E. data and evaluation data to develop the revised dual-systems estimates of the population and coverage error (see Table 6.3).

6–B.1 Reestimation of Erroneous Census Enumerations

A major concern of Census Bureau staff in reviewing the original March 2001 A.C.E. estimates was the smaller number of duplicates and "other residence" erroneous enumerations identified in the A.C.E. compared with the 1990 PES. To develop better estimates of erroneous enumerations, the Revision II analysts relied on two major studies: the Evaluation Follow-up Study and the Person Duplication Studies.

Evaluation Follow-Up, E-Sample

The Evaluation Follow-Up (EFU) Study was planned as one of the longer term A.C.E. evaluations (i.e., evaluations that would be completed after March 2001); it resembled a similar study conducted

Table 6.3 Data Sources and Evaluations Used in A.C.E. Revision II

Program	Sample Size	Use
Decennial Census	Not applicable	Source of *C* (total census enumerations) and *II* (whole-person imputations and reinstated cases) in dual-systems estimation (DSE) formula; Source of E-sample of about 700,000 people in 11,000 block clusters; Conducted spring–summer 2000
Accuracy and Coverage Evaluation (A.C.E.)		
P-Sample Interview	About 700,000 people in 11,000 block clusters	Information on nonmovers, inmovers, and outmovers (as of Census Day) at households on independent address list to use in first-stage matching; Conducted April–August 2000
Follow-up Interview (also known as Person Follow-Up, or PFU)	Nonmatched E-sample cases; Selected nonmatched P-sample cases; in 11,000 block clusters	Additional information to facilitate second-stage matching; conducted fall 2000
Matching Error Study (MES)	About 170,000 P-sample and E-sample people in 2,259 block clusters	Rematch by highly trained staff of A.C.E. subsample using original A.C.E. information to estimate matching error; Conducted December 2000.
Evaluation Follow-up (EFU)	About 70,000 E-sample cases; About 52,000 P-sample cases; in 2,259 block clusters	Additional information on residency collected in January–February 2001; Highly trained staff rematched cases using the EFU information in summer 2001

EFU Reanalysis (also known as PFU/EFU Review)	About 17,500 E-sample cases in 2,259 block clusters	Most experienced matchers used PFU (Person Follow-Up) and EFU data to determine enumeration status; Conducted summer 2001
Revision II EFU Reanalysis (also known as Recoding Operation)	About 77,000 E-sample cases; About 61,000 P-sample cases; in 2,259 block clusters	EFU Reanalysis recoding extended to full EFU evaluation samples (plus other cases, e.g., with insufficient information for matching); computer recoding used for about one-half of cases; results used to correct for measurement error in cases not linked outside the search area; Conducted summer 2002 (MES results also used for P-sample correction)
Person Duplication Studies	Full E-sample (700,000 cases); Full P-sample (700,000 cases)	Matched to census enumerations by name and birthdate nationwide; Conducted summer 2001
Further Study of Person Duplication (FSPD)	Full E-sample (700,000 cases); Full P-sample (700,000 cases)	Refinement of Person Duplication Study matching; results used to correct for duplicate E-sample cases linked to census enumerations outside A.C.E. search area; and to correct nonmover residence status for P-sample cases linked to census enumerations outside A.C.E. search area; Conducted summer 2002.

NOTES: Adapted from Kostanich (2003b:Chart 1).

for the 1990 PES (see West, 1991). In the EFU, interviewers revisited a subsample of the E-sample housing units in one-fifth of the A.C.E. block clusters in January–February 2001, obtaining data for about 70,000 people on their Census Day residence (about 10 percent of the total E-sample). The EFU subsample included E-sample cases in the evaluation block clusters who were followed up in the original A.C.E. (mostly nonmatches) and a sample of E-sample cases in the evaluation block clusters who were not followed-up in the original A.C.E. (mostly matches). The EFU also interviewed households containing about 52,000 P-sample cases in the evaluation clusters; see Section 6–B.2. The EFU interview asked detailed questions about other residences for the entire household, while the original A.C.E. follow-up interview conducted after the first stage of matching focused on specific nonmatched individuals. Experienced matchers used the information from the EFU interview to determine the match status of the EFU cases.

The rematching estimated that the A.C.E. should have classified an additional 2.8 million people as erroneous and not correct census enumerations, most often because the person lived elsewhere on Census Day. The EFU also estimated that the A.C.E. should have classified an additional 0.9 million people as correct and not erroneous census enumerations. On balance, the EFU estimated that the A.C.E. failed to measure 1.9 million erroneous census enumerations. The EFU did not resolve the status of an estimated 4.6 million census enumerations, or 1.7 percent of the total E-sample (by comparison, the unresolved rate estimated from the original A.C.E. for the E-sample cases in the EFU was 2.6 percent).[11] Population groups that exhibited the highest percentages of classification errors included people ages 18–29 and nonrelatives of the household head (Krejsa and Raglin, 2001:i–ii).

Because the EFU estimate of 1.9 million (net) unmeasured erroneous census enumerations in the A.C.E. seemed high, a subset of the EFU sample (about 17,500 cases) was reanalyzed by Census Bureau staff with the most extensive experience in matching, using the information from the original A.C.E. follow-up and the EFU interview. The result was an estimate that, on balance, the A.C.E. had failed to measure about 1.5 million erroneous census enumerations. (This es-

[11] All numbers and percentages are weighted to the household population.

timate was produced independently of the original EFU estimate of 1.9 million net unmeasured erroneous enumerations.) However, the reanalysis could not resolve the enumeration status of an estimated 15 million cases, including some cases in which the original A.C.E. and EFU enumeration status codes (correct or erroneous) conflicted (Adams and Krejsa, 2001:i).

The reanalysis examined the source of E-sample classification errors. Of cases that changed from correct to erroneous in the reanalysis, the three largest groups were: (1) people who duplicated a correct enumeration in a group quarters (34 percent, made up of 17 percent college dormitory, 8 percent nursing home, 8 percent other); (2) people who duplicated a correct enumeration at another residence (23 percent, made up of 7 percent other home, 4 percent joint custody, 5 percent visiting, 3 percent other home for work, 4 percent other type of second home); and (3) movers who had been incorrectly coded as residents (16 percent) (Adams and Krejsa, 2001:Table 3). Of cases that changed from erroneous to correct, the three largest groups were: (1) people who had been miscoded in the original A.C.E. (16 percent); (2) movers who had been incorrectly coded as nonresidents (14 percent); and (3) correctly enumerated people for whom a duplicate enumeration at another residence or a group quarters residence had been incorrectly accepted (14 percent) (Adams and Krejsa, 2001:Table 4).

The information from the reanalysis of 17,500 E-sample cases was used in developing the October 2001 preliminary revised estimates of the population and net undercount (see Section 5–D.3). For the Revision II reestimation, the reanalysis was extended to the full 70,000 EFU sample plus 7,000 cases not included in the EFU. To make best use of the limited time available for the full reanalysis, computer recoding was tested and ultimately used for 39,000 cases; the rest were recoded by technicians and analysts; a special review was conducted to resolve conflicting cases.

The A.C.E. Revision II E-sample reanalysis estimated that the A.C.E. should have classified an additional 2.7 million people as erroneous and not correct census enumerations. It also estimated that the A.C.E. should have classified 0.7 million people as correct and not erroneous enumerations. On balance, the Revision II reanalysis estimated that the A.C.E. failed to measure 2 million erroneous census enumerations (slightly higher than the EFU estimate). The

Revision II reanalysis did not resolve the status of an estimated 6.4 million census enumerations (2.4 percent of the total weighted E-sample; see Krejsa and Adams, 2002:Table 6). The unresolved cases were imputed an enumeration status probability (see Section 6–B.3).

Person Duplication Studies, E-Sample

A new study was implemented in summer 2001 to measure duplicate enumerations in the census and provide the basis for determining how many such duplications were not detected in the A.C.E. In a first stage, all E-sample cases were processed by computer, searching for an exact match by first and last name and month and year of birth among all nonimputed census enumerations nationwide (including group quarters and reinstated cases). In a second stage, members of pairs of households for which an exact match had been identified were statistically matched to identify additional duplicate enumerations.[12]

This computer matching study estimated that 1.9 million census household enumerations (weighted) duplicated another household enumeration within the A.C.E. search area (block cluster and surrounding ring); another 2.7 million census household enumerations duplicated another household enumeration outside the A.C.E. search area; and 0.7 million census household enumerations duplicated a group quarters enumeration (Mule, 2001:Table 6). In addition, an estimated 2.9 million census household enumerations duplicated other enumerations in housing units that were deleted from the census as part of the special summer 2000 effort to reduce duplicate enumerations from duplicate MAF addresses. Only 260,000 of these links were outside the A.C.E. search area.

For the Revision II reestimation, a Further Study of Person Duplication was conducted, in which some refinements were made to the computer matching methodology, including greater use of statistical matching and computing a probability for establishing a duplicate link instead of using a model weight approach. Statistical matching was used when two or more duplicates were detected within a

[12]See Mule (2001) for details of the matching, which was more complex than indicated in the text. Such matching could not be conducted for the 1990 PES because the 1990 census ascertained only year of birth and did not capture names for cases not in the PES E-sample.

household; exact matching had to be relied on when there was a link for only one person to another household or to a group quarters. No efficiency correction was made, in contrast to the October 2001 preliminary revised estimates when Fay (2001) increased the estimate of duplicate enumerations from the Person Duplication Studies by making an "efficiency" adjustment to allow for the likelihood that the computer matching did not detect all the duplicates that a computer and clerical matching process would have done. For the basis of the decision not to make an efficiency correction in Revision II, see U.S. Census Bureau (2003c:52).

In total, the Further Study of Person Duplication estimated 5.8 million duplicates in the 2000 census, an increase of 0.5 million over the original Person Duplication Studies. This number included 2.5 million duplicates involving another household in the A.C.E. search area; 2.7 million duplicates involving another household outside the A.C.E. search area; and 0.6 million duplicates involving group quarters, of which 0.5 million were outside the A.C.E. search area (Mule, 2002a:Table 2). The estimated percentages of duplicate household-to-household enumerations among total census household enumerations were higher for non-Hispanic blacks, Hispanics, and American Indians than for other race/ethnicity groups, and higher for children under 18 and young adults ages 18–29 than for older people. The estimated percentages of duplicate household-to-group quarters enumerations were higher for non-Hispanic blacks and Asians than for other race/ethnicity groups and higher for young men and women ages 18–29 than for children or older adults (Mule, 2002a:Tables F1, F3, F5, F7). We discuss in greater detail the characteristics of duplicate enumerations in Section 6–C.2.

The Further Study of Person Duplication was evaluated by two different studies. One study compared the Further Study of Person Duplication results with duplicates detected using the Census Bureau's database of administrative records (the Census and Administrative Records Duplication Study—see Bean and Bauder, 2002). The other study used the Bureau's elite matching team to clerically review samples of duplicates detected by the Further Study of Person Duplication statistical matching and by the administrative records review outside the A.C.E. search area (the Clerical Review of Census Duplicates study—see Byrne et al., 2002). The clerical review concluded that the Further Study of Person Duplication was more effective than the administrative records review in finding duplicates that were geographically close. Conversely, the administrative records review identified more duplicates that were geographically distant, but many of them were questionable.

With regard to the accuracy of the Further Study of Person Duplication, the clerical review agreed 95 percent of the time when the Further Study of Person Duplication established an E-sample duplicate link outside the A.C.E. search area and 94 percent of the time when the Further Study of Person Duplication concluded that an apparent link was not in fact a duplication. For the 1.2 million additional duplicates found by the administrative records review, but not the Further Study of Person Duplication, the clerical review agreed with the administrative records review 37 percent of the time, disagreed 47 percent of the time, and was undecided about the rest (U.S. Census Bureau, 2003c:38–39). Overall, these evaluations indicate that the Revision II estimates of census duplicates, even though higher than the preliminary revised estimates, were still an underestimate of duplication in the census.

Estimating Correct Census Enumerations in Revision II

For the Revision II estimates of correct census enumerations to include in the DSE formula, the Census Bureau used both the EFU and the Further Study of Person Duplication results. It used the Further Study of Person Duplication directly to estimate correct enumerations from among the E-sample cases that duplicated another census household or group quarters enumeration outside the A.C.E. search area (including a small number of links to deleted housing unit enumerations). It used the EFU to estimate factors to correct the measurement error in the original A.C.E. estimate of correct enumerations among the remaining E-sample cases. However, because the EFU was a subset of the full E-sample (see Table 6.3), the correction factors were calculated only for a small number of aggregate poststrata and not for the full set of E-sample poststrata (see Kostanich, 2003b:Table 4).

For the estimate from the Further Study of Person Duplication of correct enumerations among E-sample cases linked to enumerations outside the A.C.E. search area, the Census Bureau had to assign probabilities of being a duplicate or not—the Further Study of Person Duplication by itself could not determine which of two linked enumerations was correct and which the duplicate. In some cases, a decision rule was used to assign enumeration status as 0 (erroneous) or 1 (correct). Thus, for E-sample links that were already coded as erroneous in the A.C.E., this code was accepted; for E-sample links with group quarters residents, the group quarters enumeration was assumed to be correct and the E-sample case erroneous; for E-sample links of people age 18 and over who were listed as a child of the householder in one source and not so in the other source, the enumeration for "not a child of" was accepted as correct. This rule handled adults living independently in housing units who were also listed at the parents' residence, such as college students living off campus. For other duplicate links, a correct enumeration probability was assigned so that the weighted number of correct enumerations would be one-half the total weighted number of duplicate links. Probabilities were assigned separately within 18 categories defined by race/ethnicity (blacks, Hispanics, others), housing tenure, and type of linkage situation (entire household, children under age 18, all other links).

The effect of the Revision II adjustments for measurement error and undetected census duplications on the estimated number of census erroneous enumerations was substantial: the estimated number of erroneous enumerations increased from 12.5 million in the original A.C.E. to 17.2 million in Revision II.[13] Correspondingly, the estimated correct enumeration rate decreased from 95.3 percent in the original A.C.E. to 93.5 percent in Revision II, which had the effect of lowering the DSE estimate of the population and reducing the net undercount (see Section 6–B.6).

6–B.2 Reestimation of Census Omissions

The preliminary revised (October 2001) A.C.E. estimates did not take account of possible errors in the P-sample. The Revision II estimation did take account of measurement errors in the P-sample, using results from a reanalysis of the Evaluation Follow-Up Study and from the Further Study of Person Duplication (see Section 6–

[13]From tabulations by panel staff (see National Research Council, 2001a:Table 7-5) of the Full Sample Poststratum-Level Summary File, provided to the panel June 20, 2003 (see Haines, 2002:5).

B.1), as well as results from the Matching Error Study (see Section 6–A.8).

Evaluation Follow-Up Study, P-Sample

The original Evaluation Follow-Up Study completed in summer 2001 included a sample of 52,000 P-sample cases in addition to the sample of 70,000 E-sample cases discussed in Section 6–B.1 (see Raglin and Krejsa, 2001). For the preliminary revised October 2001 population estimates, no effort was made to reexamine the P-sample component of the EFU. For Revision II, all 52,000 P-sample cases were reanalyzed by a combination of computer matching and clerical review by the Bureau's most experienced matchers using information from the EFU interview and the original A.C.E. follow-up interview.[14] The purpose of the original EFU and the reanalysis for P-sample cases was to determine their residence status on Census Day; if they were not Census Day residents then they did not belong in the P-sample for estimating the match rate component of the DSE formula (see Section 6–A.3).

After a special review to minimize the number of conflicting cases, the Revision II P-sample reanalysis estimated that the A.C.E. should have classified 2.5 million Census Day residents as nonresidents; conversely, it should have classified 0.3 million nonresidents as residents. The net difference is 2.2 million people who should have been excluded from the P-sample match rate component because they were not Census Day residents. (Classification errors also occurred for inmovers who were used to estimate the number of movers in the DSE formula for most poststrata. On net, the reanalysis estimated a decrease in the estimated number of inmovers from 14.1 million to 13.3 million.) The Revision II reanalysis did not resolve the status of an estimated 7 million P-sample cases (compared with 5.8 million unresolved P-sample cases in the original A.C.E.) (Krejsa and Adams, 2002:Table 11). The unresolved cases were imputed a residence status probability (see Section 6–B.3).

[14]See the description of the Revision II reanalysis operation in Section 6–B.1; the P-sample component of the reanalysis added another 9,000 cases to the 52,000 examined in the original EFU.

Further Study of Person Duplication, P-Sample

The original Person Duplication Studies implemented in summer 2001 matched all of the P-sample cases, as well as all of the E-sample cases, to all nonimputed census enumerations nationwide. The objective for the P-sample was to determine how many cases originally classified as Census Day residents matched census enumerations outside the A.C.E. search area and, hence, might not be part of the P-sample. The preliminary revised October 2001 population estimates did not use the P-sample results. For Revision II, the Further Study of Person Duplication was implemented for the P-sample as well as the E-sample (see Section 6–B.1). The Further Study of Person Duplication results for P-sample cases who had been resident nonmovers in the original A.C.E. were used in the Revision II estimation as summarized below.

For original P-sample resident nonmovers, the Further Study of Person Duplication found links for 5.4 million cases to census enumerations outside the A.C.E. search area. Of these, 5 million involved links of household members to members of other households, and 0.4 million involved links of household members to group quarters residents (Mule, 2002a:Table 5). Of these linked P-sample cases, 2.7 million had been originally coded as P-sample nonmatches and the other 2.7 million had been originally coded as P-sample matches (implying that the matched census enumerations might also be duplicates).

The estimated percentages of linked P-sample household-to-census household cases among total original P-sample resident nonmovers were higher for non-Hispanic blacks, Hispanics, Native Hawaiians and Other Pacific Islanders, and American Indians than for other race/ethnicity groups, and higher for children under 18 and young adults ages 18–29 than for older adults. The estimated percentages of linked P-sample household-to-census group quarters cases were higher for non-Hispanic blacks than for other race/ethnicity groups and higher for young men and women ages 18–29 than for children or older adults (Mule, 2002a:Tables G1, G3, G5, G7).

The Further Study of Person Duplication P-sample component was evaluated by the Census and Administrative Records Study and the Clerical Review of Census Duplicates Study (see Section 6–B.1), which generally supported the accuracy of the Further Study of Person Duplication links for P-sample nonmover residents (U.S. Census Bureau, 2003c:39–41). The clerical review agreed 96 percent of the time when the Further Study of Person Duplication established a P-sample duplicate link with a census enumeration outside the A.C.E. search area, but only 66 percent of the time when the Further Study of Person Duplication concluded that an apparent link was not in fact a duplication. For the 2.3 million additional duplicates found by the administrative records review, but not the Further Study of Person Duplication, the clerical review agreed with the administrative records review 29 percent of the time, disagreed 56 percent of the time, and was undecided about the rest (U.S. Census Bureau, 2003c:39–41). Overall, these evaluations indicate that the Revision II estimates of P-sample nonmover resident cases that duplicated a census enumeration outside the A.C.E. search area, while substantial in number, still probably underestimated cases that should have been classified as nonresident and dropped from the P-sample.

Estimating P-Sample Residents and Matches in Revision II

The use of evaluation study results for the P-sample in the Revision II estimates of the population was complex. It involved adding separate terms to the DSE formula for total and matched nonmovers not linked outside the search area, total and matched nonmovers linked outside the search area, total and matched outmovers, and total inmovers with a duplication adjustment (Kostanich, 2003b:15). We briefly summarize the main elements in the Revision II estimation.

The correction for nonmovers linked to census enumerations outside the search area was performed using the Further Study of Person Duplication P-sample component in a manner similar to that described for the E-sample in Section 6–B.1. The purpose of the correction was to remove some of the linked cases from the P-sample on the assumption that some of them were not truly residents in the A.C.E. search area on Census Day. Unlike the E-sample, however, there was no obvious assumption on which to base the correction—it might be that one-half of the linked P-sample cases were residents and one-half nonresidents, but it might be that the proportion of residents was not one-half. By default, the Revision II estimation assigned residence probabilities to each P-sample case linked to a census enumeration outside the A.C.E. search area in the same manner as the E-sample duplicates as described in Section 6–B.1 (see also U.S. Census Bureau, 2003c:58). The reduction in nonmover residents was proportionately greater for nonmatched cases than for matched cases, which had the result of raising the P-sample match rate and lowering the DSE and net undercount overall.

The correction for measurement error for P-sample outmovers and nonmover cases that did not match to a census enumeration outside the search area was based on the results of the reanalysis of the Evaluation Follow-up Study and the Matching Error Study. The EFU reanalysis results were used to adjust residence probabilities, and the Matching Error Study results were used to correct for false matches and false nonmatches. As was the case for the E-sample, the P-sample correction factors from the EFU reanalysis and the Matching Error Study were calculated only for a small number of aggregate poststrata and not for the full set of P-sample poststrata.

On balance, the effect of the adjustments to the P-sample was to raise the match rate slightly, from 91.59 to 91.76. In turn, this increase lowered the DSE population estimate and net undercount estimates (see Section 6–B.6).

6–B.3 New Models for Missing Data

In A.C.E. Revision II, as in the original A.C.E., not all E-sample cases could be clearly assigned a status as correct or erroneous, and not all P-sample cases could be clearly assigned a status as a resident or nonresident, or, if a resident, as a match or nonmatch. Enumeration, residence, and match statuses had to be imputed to these unresolved cases. Evaluation of the original imputation procedures, which used only variables that were available after the initial matching and before follow-up to define imputation cells, indicated that they contributed significant variability, and possibly bias, to the original DSE population estimates (see Section 6–A.6). For Revision II, new imputation procedures were devised; they used variables from the evaluation follow-up data, such as whether the respondent provided an alternate Census Day address (Beaghen and Sands, 2002; Kostanich, 2003a:Ch.4). New procedures were also devised to adjust P-sample Census Day weights for households determined in the Evaluation Follow-Up reanalysis to be noninterviews and to impute residence and match status to the small number of E-sample and P-sample cases with conflicting information between the original A.C.E. follow-up and the Evaluation Follow-Up. The new procedures were estimated to have reduced the variance from imputation in the Revision II DSE estimates by 60 percent compared with the original A.C.E. procedures (Kearney, 2002:5).

6–B.4 Refined Poststratification

The Revision II effort followed the advice of many statisticians to reassess the definition of the poststrata (population groups) for which DSE population estimates were calculated, by examining the usefulness of variables obtained in the A.C.E. itself.[15] For the P-sample, for which considerable research had been done prior to 2000, the poststrata were retained largely unchanged from the original A.C.E. (see Section 6–A.10). The only change was to split people ages 0–17 into two groups: ages 0–9 and ages 10–17.

For the E-sample, the poststrata were substantially revised (U.S. Census Bureau, 2003c:18). For each of the seven race/ethnicity domains, separate E-sample poststrata were created for proxy responses, which were cases obtained by enumerators from landlords or neighbors. Generally, these cases had low correct enumeration rates (average 60 percent).[16] For nonproxy cases for non-Hispanic whites, the revised E-sample poststratification dropped region, metropolitan population size and type of enumeration area, and mail return rate of census tract. In their place it used household size for heads of nuclear families and all others by type of census return (early mail return, late mail return, early nonmail return, late nonmail return). Housing tenure (owner/renter) was retained. For non-Hispanic blacks and Hispanics, the stratification was the same as for non-Hispanic whites except that it did not use household size. For Native Hawaiians and Other Pacific Islanders, non-Hispanic

[15]Several participants at a workshop in fall 1999 (National Research Council, 2001e) urged the Bureau to use modeling techniques to develop poststrata on the basis of the A.C.E. data themselves. The model would assess the best predictors of coverage, but the Bureau decided such an approach was not feasible for the original estimation.

[16]Tabulations by panel staff from the Full E Sample Post-Stratum-Level Summary File, provided to the panel June 20, 2003 (see Haines, 2002:5).

Asians, and American Indians off reservations, the stratification dropped housing tenure and used household relationship by the four categories of type of census return. For American Indians on reservations, the stratification used only household relationship (head of nuclear family, other). The revised poststrata captured more of the variation in correct enumeration rates than the original poststrata.

On net, the new poststratification had little effect on the national net undercount. However, for some small geographic areas with large numbers of proxy responses, the new poststratification produced large estimated net overcounts of their population (5 percent or more—see U.S. Census Bureau, 2003c:Table 10). While such sizeable net overcount estimates could be accurate, they could have resulted instead from the lack of comparability of the Revision II E-sample and P-sample poststrata. Consider the situation in which proxy respondents (landlords or neighbors) tended not only to report people who should not have been counted (erroneous enumerations), but also to omit people who should have been counted (omissions). In such a case, the Revision II estimation would have calculated a high erroneous enumeration rate, based on the proxy poststrata, but it would not have calculated a comparably high nonmatch rate because there were no P-sample poststrata comparable to those for proxy respondents in the E-sample—all P-sample poststrata included household respondents as well as proxies. The result would have been an underestimate of the DSE population. Nationally, this underestimation would have little effect because only 3 percent of household members were enumerated by proxies. However, the effects could have been substantial for some subnational areas (see Section 6–D.4).

6–B.5 Adjustment for Correlation Bias

The last major goal of the Revision II A.C.E. work was to consider a further adjustment for census omissions among selected population groups that were assumed to be missed disproportionately in both the P-sample and the census; such groups contribute disproportionately to the people who are not included in the DSE estimate of the population. This phenomenon is referred to as "correlation bias." The term correlation bias has been used in several ways in research and evaluation of dual-systems estimation (see National Research Council, 1999b:Ch.4). Here, we use the term loosely to refer to the phenomenon in which some groups are more apt than other groups to be missed in both systems (P-sample and census).

For several censuses, comparisons of demographic analysis results with those from dual-systems estimation have lent support to a hypothesis that adult black men (and nonblack men to a much lesser extent) contribute disproportionately compared with black (nonblack) women to the group of people not included in the DSE estimate of the population. This hypothesis is based on the finding of higher estimated net undercount rates from demographic analysis compared with dual-systems estimation for black men relative to black women, together with the finding that sex ratios (men per 100 women) appear reasonable in demographic analysis but are low in both the DSE and the census, particularly for blacks (see Table 6.4).

Census Bureau analysts previously conducted research on the possibilities of using the sex ratios from demographic analysis to reduce correlation bias in DSE population estimates (see Bell, 1993). However, not until the 2000 A.C.E. Revision II work did the Bureau seriously contemplate making such an adjustment to the DSE estimates.

A motivation for using a sex ratio adjustment for correlation bias emerged from the results of all the revisions incorporated into the A.C.E. Revision II population estimates summarized in Sections 6–B.1 through 6–B.4. The corrections for duplication of E-sample enumerations with census enumerations outside the A.C.E. search area and other measurement error in the E-sample decreased the estimated correct enumeration rate, while the correction for duplication of P-sample nonmover resident cases with census enumerations outside the search area increased the estimated match rate. The consequence was to reduce the DSE population estimate by 6.3 million people from the original March 2001 estimate, leaving an estimated net *overcount* of 3 million people, or 1.1 percent of the household population, instead of an estimated net *undercount* of 1.2 percent or 3.3 million people. Furthermore, several race domains, including American Indians on reservations, non-Hispanic blacks, non-Hispanic Asians, and non-Hispanic whites, went from an estimated net undercount in March 2001 to an estimated net overcount in Revision II (Mule, 2003:Table 2, cumulative column for the row

Table 6.4 Sex Ratios (Men per 100 Women) from the Census, Demographic Analysis (DA), Accuracy and Coverage Evaluation (A.C.E.) Revision II, and Post-Enumeration Survey (PES), 1990 and 2000

	1990			2000		
Race/Age Group	DA	PES	Census	Revised DA[a]	A.C.E. Revision II[b]	Census
Black						
Total	95.2	90.4	89.6	95.1	90.8	90.6
0–17 years	102.4	102.4	102.4	—	—	—
0–9	—	—	—	102.7	103.1	103.1
10–17	—	—	—	102.7	103.4	103.4
18–29 years	99.3	92.1	94.0	100.2	94.0	93.9
30–49 years	95.9	89.0	86.2	96.9	88.9	88.5
50 or more years	78.3	72.1	71.5	77.2	73.4	73.4
Nonblack						
Total	97.2	96.5	95.9	98.1	97.6	97.1
0–17 years	105.2	105.5	105.5	—	—	—
0–9	—	—	—	104.8	105.2	105.2
10–17	—	—	—	105.5	105.9	106.0
18–29 years	104.9	104.6	103.8	106.7	107.1	105.3
30–49 years	102.0	100.3	99.6	102.3	100.7	100.6
50 or more years	80.8	79.9	79.4	84.2	83.3	83.1

NOTE: —, not available.

[a] "Revised" demographic analysis estimates are those from October 2001, incorporating changes to births and legal and illegal immigration; see Table 5.3.

[b] A.C.E. Revision II estimates are before adjustment for correlation bias.

SOURCE: Robinson (2001a:Table 8); Robinson and Adlakha (2002:Table 4).

labeled "P-sample coding corrections"). The Bureau was concerned that the DSE without a correlation bias adjustment could move the estimates further from the truth for groups that were truly undercounted but were estimated to be overcounted (U.S. Census Bureau, 2003c:50).

Hence, the Bureau made an adjustment for correlation bias, based on sex ratios from demographic analysis. Assuming that the DSE estimates for females were correct, the P-sample match rates were recomputed for black males ages 18–29, 30–49, and 50 and over, and all other males ages 30–49 and 50 and over (U.S. Census Bureau, 2003c:5). The resulting match rate estimates are termed "census inclusion rates"; they are lower for the relevant groups than the originally estimated match rates (see "Match Rates" in Section 6–B.6). No correlation bias adjustment was made for children under age 18 or for nonblack men ages 18–29 because the sex ratios for these groups exceeded 100 by similar amounts in all three sources: demographic analysis, the A.C.E., and the census (see Table 6.4). Because of the limitations of demographic analysis it was not possible to determine if adjustments for correlation bias for particular groups of nonblack men age 30 and over (e.g., Asians or Hispanics) were warranted.

6–B.6 Putting It All Together: The A.C.E. Revision II Estimates

Four tables present results from the A.C.E. Revision II effort for major population groups. They are: a table of correct enumeration rates (Table 6.5); a table of match rates (Table 6.6), including rates before and after the correlation bias adjustment described in Section 6–B.5; a table of net undercount rates (Table 6.7); and a table of the percentage contribution of each major change in the A.C.E. Revision II estimation method to the Revision II net undercount rates (Table 6.8).

Correct Enumeration Rates

Correct enumeration rates show marked changes comparing the A.C.E. Revision II estimates and the original March 2001 A.C.E. estimates (Table 6.5). The direction of the change is always toward a lower rate—2 percentage points for most groups and 5 percentage points for American Indians on reservations. Most often, the Revision II rate is below the corresponding PES rate, which is not surprising because there was no equivalent to the Further Study of Person Duplication in the PES, and the results of the Evaluation Follow-Up that was conducted were not used to adjust the PES. It is quite possible that a nationwide matching study would have turned up more erroneous enumerations than the PES estimated, although whether the increase would have been as great as in the A.C.E. cannot be assessed.

Match Rates

Match rates show very small changes comparing the A.C.E. Revision II estimates and the original March 2001 A.C.E. estimates (Ta-

Table 6.5 Correct Enumeration Rates Estimated from the E-Sample (percents), 2000 A.C.E. and 1990 PES, by Race/Ethnicity Domain and Housing Tenure (weighted)

Domain and Tenure Group	Original A.C.E. (March 2001)	A.C.E. Revision II (March 2003)	1990 PES[a]
American Indian/Alaska Native on Reservation			
Owner	95.65	90.95	91.54[b]
Renter	96.15	91.23	
American Indian/Alaska Native Off Reservation			
Owner	94.56	92.42	—
Renter	93.16	91.09	—
Hispanic Origin			
Owner	96.25	94.33	95.56
Renter	92.79	90.88	90.58
Black (Non-Hispanic)			
Owner	94.25	92.28	92.84
Renter	91.16	88.96	89.19
Native Hawaiian/Other Pacific Islander			
Owner	93.79	91.86	—
Renter	92.33	90.67	—
Asian (Non-Hispanic)[c]			
Owner	95.84	93.70	93.13
Renter	92.45	91.41	92.22
White and Other Races (Non-Hispanic)			
Owner	96.70	95.10	95.84
Renter	93.20	91.25	92.61
Total	95.28	93.48	94.27

NOTES: Correct enumeration rates are correct enumerations divided by the sum of correct and erroneous enumerations; —, not estimated. See Appendix E (Table E.3) for definitions of race/ethnicity domains.

[a] Revision II and PES rates are not comparable because there was no equivalent to the Further Study of Person Duplication in 1990 and the results of the 1990 Evaluation Follow-Up were not used to adjust the estimated rates.
[b] Total; not available by tenure.
[c] 1990 correct enumeration rates include Pacific Islanders.

SOURCES: Original A.C.E. and PES correct enumeration rates from Davis (2001:Tables E-2, F-1, F-2); Revision II A.C.E. correct enumeration rates by race/ethnicity and housing tenure group from table provided by the U.S. Census Bureau to the panel, May 22, 2003; total Revision II correct enumeration rate from Fenstermaker (2002:Table 9).

Table 6.6 Match Rates and Census Inclusion Rates Estimated from the P-Sample (percents), 2000 A.C.E. and 1990 PES, by Race/Ethnicity Domain and Housing Tenure (weighted)

Domain and Tenure Group	Original A.C.E. (March 2001)	Revision II A.C.E. (March 2003)		
	Match Rate	Match Rate	Census Inclusion Rate	1990 PES Match Rate[a]
American Indian/Alaska Native on Reservation				
Owner	85.43	86.38	86.13	78.13[b]
Renter	87.08	87.34	87.14	
American Indian/Alaska Native off Reservation				
Owner	90.19	90.86	90.54	—
Renter	84.65	84.48	84.25	—
Hispanic Origin				
Owner	90.79	91.25	90.96	92.81
Renter	84.48	84.57	84.34	82.45
Black (Non-Hispanic)				
Owner	90.14	90.56	88.27	89.65
Renter	83.67	83.88	82.03	82.28
Native Hawaiian/Other Pacific Islander				
Owner	87.36	87.46	87.15	—
Renter	82.39	83.49	83.27	—
Asian (Non-Hispanic)[c]				
Owner	92.34	92.66	92.32	93.71
Renter	87.33	87.37	87.07	84.36
White and Other Races (Non-Hispanic)				
Owner	94.60	95.02	94.63	95.64
Renter	88.37	88.43	88.14	88.62
Total	91.59	91.76	91.19	92.22

NOTES: Match rates are matches divided by the sum of matches and nonmatches; census inclusion rates are match rates adjusted for correlation bias for adult males using sex ratios from demographic analysis (see U.S. Census Bureau, 2003c:49–52); —, not estimated. See Appendix E (Table E.3) for definitions of race/ethnicity domains.

[a] Revision II and PES rates are not comparable because there was no equivalent to the Further Study of Person Duplication in 1990, and the results of the 1990 Evaluation Follow-Up and Matching Error Studies were not used to adjust the estimated match rates.
[b] Total; not available by tenure.
[c] 1990 match rates include Pacific Islanders.

SOURCES: Original A.C.E. and PES match rates from Davis (2001:Tables E-2, F-1, F-2); Revision II A.C.E. match rates and census inclusion rates for race/ethnicity and housing tenure groups from table provided by the U.S. Census Bureau to the panel, May 22, 2003; total Revision II rates from Fenstermaker (2002:Table 9).

Table 6.7 Estimated Net Undercount Rates for Major Groups (percents), Original 2000 A.C.E. (March 2001), Revision II A.C.E. (March 2003), and 1990 PES (standard error percents in parentheses)

	Original A.C.E.		Revision II A.C.E.		1990 PES	
Major Group	Estimate	(S.E.)	Estimate	(S.E.)	Estimate	(S.E.)
Total Population	1.18	(0.13)	−0.49	(0.20)	1.61	(0.20)
Race/Ethnicity Domain						
American Indian/Alaska Native on Reservation	4.74	(1.20)	−0.88	(1.53)	12.22	(5.29)
American Indian/Alaska Native off Reservation	3.28	(1.33)	0.62	(1.35)	—	—
Hispanic Origin	2.85	(0.38)	0.71	(0.44)	4.99	(0.82)
Black (Non-Hispanic)	2.17	(0.35)	1.84	(0.43)	4.57	(0.55)
Native Hawaiian or Other Pacific Islander	4.60	(2.77)	2.12	(2.73)	—	—
Asian (Non-Hispanic)[a]	0.96	(0.64)	−0.75	(0.68)	2.36	(1.39)
White and Other Races (Non-Hispanic)	0.67	(0.14)	−1.13	(0.20)	0.68	(0.22)
Age and Sex						
Under 10 years[b]	1.54	(0.19)	−0.46	(0.33)	3.18	(0.29)
10-17 years	1.54	(0.19)	−1.32	(0.41)	3.18	(0.29)
18-29 years						
Male	3.77	(0.32)	1.12	(0.63)	3.30	(0.54)
Female	2.23	(0.29)	−1.39	(0.52)	2.83	(0.47
30-49 years						
Male	1.86	(0.19)	2.01	(0.25)	1.89	(0.32)
Female	0.96	(0.17)	−0.60	(0.25)	0.88	(0.25)
50 years and over						
Male	−0.25	(0.18)	−0.80	(0.27)	−0.59	(0.34)
Female	−0.79	(0.17)	−2.53	(0.27)	−1.24	(0.29)
Housing Tenure						
Owner	0.44	(0.14)	−1.25	(0.20)	0.04	(0.21)
Renter	2.75	(0.26)	1.14	(0.36)	4.51	(0.43)

NOTES: Net undercount is the difference between the estimate (A.C.E. or PES) and the census, divided by the estimate. Minus sign (−) indicates a net overcount. For 2000, total population is the household population; for 1990, it is the household population plus the noninstitutional group quarters population. See Appendix E (Table E.3) for definitions of race/ethnicity domains; —, not estimated.

[a] In 1990 includes Pacific Islanders.
[b] In the original A.C.E. and PES, children ages 0–17 were a single category.

SOURCE: U.S. Census Bureau (2003c:Table 1).

ble 6.6). The direction of the change is usually toward a higher rate. For most groups, the Revision II rate is similar to the corresponding PES rate, even though Revision II estimated a much reduced net undercount (or overcount) compared with the PES (see Table 6.7). The reason for this result has to do with the much larger number of census whole-person imputations and reinstated records that could not be included in the A.C.E. E-sample for matching (see Section 6–C.1).

Census inclusion rates incorporate the sex ratio-based correlation bias adjustments. They are about 2 percentage points below the Revision II match rates for blacks, and only slightly below the match rates for other groups.

Net Undercount Rates

Net undercount rates show substantial changes comparing the A.C.E. Revision II estimates with the original March 2001 A.C.E. estimates (Table 6.7). The national net undercount rate declined from 1.2 percent of the household population to a slight net overcount (0.5 percent) of the population—a swing of 1.7 percentage points. By race and ethnicity, net undercount rates were reduced by 1.7 to 5.6 percentage points for every group except blacks, for whom the reduction was only 0.3 percentage points (from 2.2 percent in the original A.C.E. to 1.8 percent in Revision II). By age and sex, net undercount rates were reduced by 1.6 to 3.6 percentage points for every group except men ages 30–49, for whom the reduction was only 0.2 percentage points, and men age 50 and over, for whom the reduction was only 0.6 percentage points. For homeowners, the net undercount rate decreased by 1.7 percentage points (from 0.4 percent in the original A.C.E. to a net overcount of 1.3 percent in Revision II). For renters, the net undercount rate decreased by 1.6 percentage points.

In regard to differential net undercount among population groups, which is more important than levels of net undercount, there were increases as well as decreases in net undercount rate differences comparing the original A.C.E. with Revision II. The difference in net undercount rates between Hispanics and non-Hispanic whites and other races narrowed somewhat from 2.2

percentage points in the original A.C.E. to 1.8 percentage points in Revision II. However, the difference in net undercount rates between owners and renters increased slightly from 2.3 percentage points in the original A.C.E. to 2.4 percentage points in Revision II. The difference in net undercount rates between non-Hispanic blacks and non-Hispanic whites and other races increased from 1.5 percentage points in the original A.C.E. to 3.0 percentage points in Revision II. By comparison with the 1990 PES, differences in Revision II net undercount rates between population groups were smaller: for example, the differences in rates in the 1990 PES between Hispanics and non-Hispanic whites and other races, between owners and renters, and between non-Hispanic blacks and non-Hispanic whites and other races were 4.3 percentage points, 4.5 percentage points, and 3.9 percentage points, respectively. The important changes in the Revision II estimation methods, however, impair the validity of comparisons with the PES (see Section 6–D.2).

We do not have information with which to examine differences in net undercount rates among states and other geographic areas, either for the original A.C.E. estimates compared with Revision II or for either set of 2000 estimates compared with the 1990 PES. However, given the reduction in net undercount rates and in differential net undercount observed in 2000 for poststrata defined for population groups, it is reasonable to infer that net undercount rates and differential net undercount were also probably reduced for geographic areas. The reason is that poststratification research to date has not strongly supported the use of geographic variables in preference to such variables as age, sex, race, and housing tenure (see, e.g., Griffin and Haines, 2000; Schindler, 2000). In the poststratification for the original 2000 A.C.E. and Revision II P-sample (but not E-sample), a regional classification was used for non-Hispanic whites along with a classification by size of metropolitan statistical area and type of enumeration area for non-Hispanic whites, blacks, and Hispanics. The Integrated Coverage Measurement design would have supported direct state estimates of net undercount, but that design was adopted principally because of concerns from Congress and other stakeholders that state population totals for reapportionment should be estimated from data for the individual state and not borrow data from other states (see Appendix A.4.a).

Components of Change

Specific components of change in the Revision II methods show somewhat different effects for four race and ethnicity groups (Table 6.8). Overall, the new poststratification and adjustment for measurement error in the P-sample had little effect on Revision II net undercount rates compared with the original A.C.E. rates. The correction for E-sample duplications with census enumerations outside the A.C.E. search area had the largest effect of all of the changes for most groups, reducing the net undercount rate by 0.9 percentage points for non-Hispanic whites and other races up to 1.5 percentage points for non-Hispanic blacks. E-sample measurement error corrections also reduced the net undercount rate by about 1 percentage point for each group shown, and the correction for duplications of P-sample nonmover residents with census enumerations outside the A.C.E. search area reduced the net undercount rate by 0.3 to 0.5 percentage points for each group shown. Finally, the correlation bias adjustment had a small effect (less than 0.4 percent increase in the net undercount rate) for Hispanics, non-Hispanic Asians, and non-Hispanic whites and other races, but it increased the net undercount rate for blacks by 2.4 percentage points. Consequently, the correlation bias adjustment largely explains why black net undercount rates did not decline as much as the rates for Hispanics and other groups. The correlation bias adjustment also explains why net undercount rates for men ages 30–49 and age 50 and over did not decline as much as the rates for women, children, and younger men (see Table 6.7).

6–B.7 Assessment of Error in the A.C.E. Revision II Estimates

As part of the immense amount of data analysis and estimation undertaken for the A.C.E. Revision II estimates of net undercount in 2000, Census Bureau analysts conducted evaluations to estimate the possible bias (systematic error) and variance (random error) in the estimates. The evaluations of bias included the construction of 95 percent confidence intervals around the Revision II estimates and comparisons of the relative accuracy of the Revision II estimates and census counts for geographic areas (see Section 6–D.4 for the latter analysis). These evaluations were limited because data that

Table 6.8 Components of Change from the Original A.C.E. Net Undercount Rate to the Revision II Net Undercount Rate for Selected Race/Ethnicity Domains

			Non-Hispanic		
Component of Change	Total	Hispanic	Black	Asian	White
Census Count, Household Population (thousands)	273,578	34,538	33,470	9,960	192,924
Original A.C.E. Net Undercount as Percent of Census	1.19	2.94	2.21	0.97	0.68
Change in Estimated Net Undercount Rate as Percent of Census					
New Post-Stratification	0.01	−0.09	0.16	0.16	<0.01
E-Sample Duplication Corrections	−1.03	−1.07	−1.50	−1.07	−0.93
E-Sample Measurement Error Corrections	−0.89	−0.97	−0.98	−0.92	−0.85
P-Sample Duplication Corrections	−0.40	−0.45	−0.49	−0.26	−0.38
P-Sample Measurement Error Corrections	<0.01	0.07	0.07	0.01	−0.02
Correlation Bias Adjustment	0.62	0.29	2.40	0.36	0.39
Cumulative Change	−1.68	−2.22	−0.34	−1.72	−1.79
Net Undercount as Percent of Census, A.C.E. Revision II	−0.49	0.72	1.88	−0.75	−1.11

NOTES: Net undercount rates differ slightly from those in Table 6.7 because of different denominators (census count in this table; A.C.E. or PES estimate of population group in Table 6.7); minus sign (−) indicates a net overcount or a decrease in the net undercount rate. See Appendix E (Table E.3) for definitions of race/ethnicity domains. Individual component effects are from introducing one change at a time in the estimation methodology. A different ordering of the revisions would result in slightly different component effects. E-sample and P-sample duplication corrections were based on links of cases to census enumerations outside the A.C.E. search area identified in the Further Study of Person Duplication. E-sample and P-sample measurement error corrections were based on the results of the Evaluation Follow-Up reanalysis for cases not linked to census enumerations outside the A.C.E. search area (the P-sample measurement error corrections also include the results of the Matching Error Study).

SOURCE: Adapted from Mule (2003:Tables 1, 2).

were used to estimate bias in the original A.C.E. estimates were used as part of the Revision II estimation. Thus, the bias evaluations did not take account of the following sources of error: response error or coding error in the Revision II determination of P-sample residency or match status or E-sample correct enumeration status, which were based on evaluation samples listed in Table 6.3; response error or coding error in the Revision II determination of P-sample mover status; error in the approach used to estimate the contribution to correct enumerations from E-sample cases with duplicate links; error in demographic analysis sex ratios; or error in the model used to estimate correlation bias from these sex ratios (see U.S. Census Bureau, 2003c:68).

The evaluation to construct 95 percent confidence intervals around the Revision II estimates found a possible bias for some population groups (see Mulry and ZuWallack, 2002). In particular, it appears that the Revision II population estimates for non-Hispanic black owners and renters might be too low (U.S. Census Bureau, 2003c:45). The results of the Census and Administrative Records Duplication Study largely account for this result (see Section 6–B.1).

Looking simply at variance from sampling, imputation, and other sources, the Revision II estimates exhibited only slightly larger standard errors than the original A.C.E. estimates, and, in most cases, the Revision II standard errors were lower than the corresponding standard errors for the 1990 PES estimates (see Table 6.7). This result obtained even though many of the data sources used in the Revision II estimation were subsamples of the original A.C.E. The explanation is that the Revision II estimation used data from the full A.C.E. to estimate some components (e.g., E-sample duplications of census enumerations outside the A.C.E. search area), while, for other components (e.g., corrections for measurement error), the estimation used evaluation subsamples to develop correction factors to apply to the full A.C.E. samples. This strategy produced a complex DSE formula with multiple components for each of the elements in the basic formula (see Kostanich, 2003b:15), but it enabled the Revision II estimation to make use of all the available data and not just small evaluation samples as was done for the October 2001 preliminary revised estimates.

A caveat is in order, however. The various E-sample and P-sample correction factors that were developed from subsamples were computed only for highly aggregated poststrata. The variance calculations had to assume that the aggregate correction factors did not vary for individual poststrata within the larger aggregates. If for no other reason, it is likely that the variance estimates for the Revision II A.C.E. estimates are biased downward.

6–C FACTORS IN COVERAGE

Two factors merit discussion for the role they played in producing an estimated net overcount in the 2000 census. They are computer-based, whole-person census imputations and duplicate census enumerations.

6–C.1 Whole-Person Imputations

The much larger number of whole-person imputations in 2000 (5.8 million) compared with 1990 (1.9 million) helps explain one of the initial puzzles regarding the original A.C.E. estimates of net undercount.[17] The puzzle was that the original A.C.E. correct enumeration and match rates were very similar to the PES rates (see Tables 6.5 and 6.6). Other things equal, these similarities should have produced similar estimates of net undercount. Yet the original A.C.E. estimates showed marked reductions in net undercount rates from 1990 levels for such groups as minorities, renters, and children and a consequent narrowing of differences in net undercount rates between historically less-well-counted and better-counted groups (see Table 6.7).

The explanation lies in those census cases that had to be excluded from the A.C.E. because they were wholly imputed and hence could not be matched or because they were available too late for matching—the *II* term in the DSE formula (see Section 5–A). There were so many more of these cases in 2000 than in 1990 that when they were added back to the census counts for comparison with the DSE population estimates, the result was to lower the net undercount estimates for 2000 compared with 1990.

The *II*s in 2000 included 2.4 million reinstated cases from the special summer 2000 MAF unduplication operation, whereas the

[17] Whole-person imputations include types 1–5 as described in Section 4–D and Box 4.2; see also Section G.4.

*II*s in 1990 included only 0.3 million late enumerations. However, the reinstated cases were distributed in roughly the same proportions across major population groups (see Table 6.9), so that they do not explain the narrowing of coverage differences observed in the A.C.E. compared with the PES. In contrast, the 5.8 million whole-person imputations in 2000 accounted for proportionately more of historically less well-counted groups than of better counted groups, whereas the much smaller number of 1.9 million whole-person imputations in 1990 did not show such large differences among groups (see Table 6.9).

If all of the whole-person imputations were accurately assigned to poststrata and represented people who otherwise would have been omitted from the census, then the puzzle of similarly low P-sample match rates and yet lower net undercount estimates in 2000 compared with 1990 would have a ready explanation. The explanation is that these cases would have matched to the P-sample if they had been enumerated instead of imputed, so that the original A.C.E. would have exhibited higher match rates than the PES and lower net undercount estimates. Instead, because imputation was substituted for additional field work, the original A.C.E. had artificially low match rates, but once the *II*s were added back to the census count for comparison with the DSE estimates, the original A.C.E. had lower net undercount estimates than the PES.

The question is the accuracy of the whole-person imputations in 2000 in terms of their numbers and imputed characteristics needed for poststratification. Of the 5.8 million whole-person imputations, a large number—2.3 million—were imputed in situations when the household size and characteristics of other members were known (type 1 imputations). Many of these were children in large households who could not be reported because of lack of room on the questionnaire. (The 2000 questionnaire had room for six persons, compared with seven in 1990.) Another large group—also about 2.3 million—were people imputed into households believed to be occupied for which household size, but not other information, was available (type 2 imputations).

These two types of whole-person imputations did not alter the numbers of people who were included in the census overall so long as household size was accurately reported. Moreover, because the imputation process used the characteristics of other household members and neighboring households of the same size, respectively, type 1 and 2 imputations may have been quite accurate with regard to their characteristics. Evidence from the 2000 Census Administrative Records Experiment (see Section 4–D.2) suggests that imputations of household size and demographic characteristics may not have been accurate in many individual cases, although whether distributions were biased is not known.

A problematic group with regard to accuracy—amounting to 1.2 million people in 2000 compared with only 54,000 people in 1990—comprised those who were imputed into the census when there was no information about the size of the household, or, in some instances, whether the address was occupied or even a housing unit (imputation types 3–5). An alternative approach could have been to delete all of these addresses from the census; however, such an approach would undoubtedly have underestimated the true number of household residents (particularly when a unit was known to be occupied). The question is whether the numbers (and characteristics) of imputed people were larger (or smaller) than the true numbers and therefore contributed to overestimating (or underestimating) the population at these addresses. The effects of such over- or underestimation on coverage would be quite small for the nation as a whole, but they could be significant for particular geographic areas or population groups.

For example, geographic analysis of types of whole-person imputations for census tracts conducted by panel staff revealed considerable clustering of some imputation types by geographic area.[18] In particular, type 5 imputations, in which status as a housing unit had to be imputed first, followed by imputation of occupancy status, household size, and, finally, household member characteristics, were heavily clustered in rural list/enumerate areas, such as the Adirondacks region of New York State and parts of Arizona and New Mexico. Although there were only 415,000 type 5 imputations nationwide (0.2 percent of the household population), for some counties in list/enumerate areas, type 5 imputations accounted for significant percentages of the population. The address list for list/enumerate areas was developed by enumerators in a single-stage

[18]The analysis used a census tract summary file of whole-person imputations by type, provided to the panel April 4, 2002; see also Kilmer (2002).

Table 6.9 Percentage Distribution of People Requiring Imputation and Reinstated Records in the 2000 Census, and Percentage Distribution of Total People with Insufficient Information in 1990, by Race/Ethnicity Domain and Housing Tenure and by Age/Sex Categories

	Percent of Household Population, 2000			
Panel A Domain and Tenure Group	People Requiring Imputation	Reinstated Records	Total with Insufficient Information	Percent of Household Population with Insufficient Information, 1990[a]
American Indian/Alaska Native on Reservation				
Owner	5.13	0.97	6.00	3.16
Renter	4.74	0.94	5.58	(Total)
American Indian/Alaska Native off Reservation				
Owner	2.36	1.20	3.51	—
Renter	3.00	1.16	4.12	—
Hispanic Origin				
Owner	3.74	0.92	4.61	1.03
Renter	3.99	1.00	4.96	1.56
Black (Non-Hispanic)				
Owner	2.84	1.00	3.81	1.20
Renter	3.95	0.96	4.88	1.89
Native Hawaiian/Pacific Islander				
Owner	3.67	0.87	4.49	—
Renter	3.83	0.92	4.70	—

Asian (Non-Hispanic)				
Owner	2.46	0.69	3.13	0.74
Renter	3.35	0.77	4.10	1.71
White and Other Races (Non-Hispanic)				
Owner	1.24	0.71	1.93	0.46
Renter	2.38	1.12	3.47	1.44
Total Owner	1.66	0.75	2.39	0.56
Total Renter	3.08	1.05	4.10	1.55
Panel B				
Age/Sex Group				
Children Under Age 18	3.11	0.92	4.00	0.82
Men Ages 18–29	2.86	0.82	3.65	1.45
Women Ages 18–29	2.56	1.03	3.46	1.45
Men Ages 30–49	1.77	0.79	2.53	0.76
Women Ages 30–49	1.58	0.81	2.37	0.70
Men Age 50 and Over	1.25	0.81	2.04	0.69
Women Age 50 and Over	1.30	0.80	2.08	0.79
Total	2.11	0.85	2.93	0.90

NOTES: The 2000 total with insufficient information is the unduplicated sum of people requiring imputation and reinstated records to the census; 1990 figures include small numbers of reinstated records to the census from coverage improvement operations. —, not available.

[a] Data exclude American Indians living on reservations; the Asian (non-Hispanic) data for 1990 include Pacific Islanders.

SOURCE: Data for 2000 are from tabulations by panel staff of U.S. Census Bureau, Pre-Collapsed Post-Stratum Summary File (U.S.), provided to the panel February 16, 2001; data for 1990 are from Davis (2001:Tables F.1, F.2).

operation, with a follow-up operation to recheck units classified as vacant. However, some of these addresses may have been temporary recreational lodging (e.g., fishing camps) for which it was difficult to determine their housing status. In addition, the Census Bureau identified some data processing problems that produced larger-than-expected numbers of addresses for which housing status was not known (type 5 imputations) and that probably contributed to overimputation of households for which occupancy status was not known (type 4 imputations; see Section 4–D.2).

6–C.2 Duplicate Census Enumerations

The 2000 census included several operations that were explicitly designed to reduce duplicate enumerations. They included the Primary Selection Algorithm (PSA) and the various operations to unduplicate MAF addresses and associated households. The MAF unduplication operations included a planned operation prior to nonresponse follow-up and the special unplanned operation in summer 2000, which temporarily deleted census records that appeared to duplicate records at another address and then reinstated some of them (see Section 4–E).

The purpose of the PSA was to determine which households and people to include in the census when more than one questionnaire was returned with the *same* MAF identification number (see Appendix C.5.c). Such duplication could occur, for example, when a respondent mailed back a census form after the cutoff date for determining the nonresponse follow-up workload and the enumerator then obtained a second form from the household. In all, 9 percent of census housing units had two returns and 0.4 percent had three or more returns that were eligible for the PSA operation. In most instances, the PSA discarded duplicate returns; less often, the PSA found additional people to assign to a basic return or identified more than one household at an address (Baumgardner et al., 2001:22–27).

Despite these operations, however, the original A.C.E. identified 1.9 million census duplicates (Anderson and Fienberg, 2001:Table 2). The original A.C.E. also identified 2.7 million "other residence" erroneous enumerations (e.g., the person should have been enumerated in group quarters or at another home), many of which were probably duplicates. On the basis of the Evaluation Follow-Up Study and the

Further Study of Person Duplication, A.C.E. Revision II identified an additional 5.2 million duplicates and "other residence" erroneous enumerations in 2000 (see Section 6–B.1), for a total of 9.8 million such enumerations (Mule, 2003:2). Moreover, evaluations of the Further Study of Person Duplication determined that it underestimated duplicate census enumerations.

Because a similar Further Study of Person Duplication cannot be conducted for the 1990 census, it is hard to know how many duplicate enumerations occurred in that census beyond those identified in the PES (4.5 million plus some fraction of 6.2 million "other residence" erroneous enumerations; Anderson and Fienberg, 2001:Table 2). Societal trends, such as more children of divorced parents in joint custody, more people with winter and summer homes or weekday and weekend homes, could mean that duplication was more of a potential and actual problem in 2000 than in 1990, but the panel knows of no evidence on that point.

The Further Study of Person Duplication provided distributions of duplicate enumerations (Table 6.10), which indicate that they occurred disproportionately in 2000 for some historically less-well-counted groups. This result differed from previous censuses in which omissions were more concentrated than duplicates or other erroneous enumerations among hard-to-count groups (e.g., see Ericksen et al., 1991:Table 1, which examined omission and erroneous enumeration rates for census tracts grouped into deciles of mail return rates).

For race/ethnicity domains, the Further Study of Person Duplication estimated higher census household-to-household duplication rates for American Indians on reservations (2.7 percent), blacks (2.4 percent), Hispanics (2.4 percent), and American Indians off reservations (2.3 percent), compared with non-Hispanic whites and other races (1.8 percent). By age and sex, the Further Study of Person Duplication estimated higher census household-to-household duplication rates for population groups under age 30 (rates were 2.3 percent or higher) compared with population groups age 30 and over (rates were 1.8 percent) or lower. The highest rates of census household-to-group quarters duplications among race/ethnicity domains were for blacks (0.4 percent, mostly in college dormitories and prisons) and Asians (0.4 percent, almost entirely in college dormitories). The highest rates of census household-to-group quarters

Table 6.10 Percent Duplicate Enumerations in 2000 Census by Type for Race/Ethnicity Domains and Age/Sex Groups from the Further Study of Person Duplication

	Census Housing Unit to Housing Unit			
Population Group	Within A.C.E. Search Area	Outside A.C.E. Search Area	Total	Census Housing Unit to Group Quarters
Race/Ethnicity Domain				
American Indian/Alaska Native off Reservation	0.91	1.38	2.29	0.20
American Indian/Alaska Native on Reservation	0.97	1.77	2.74	0.21
Hispanic	1.41	1.02	2.43	0.16
Black (Non-Hispanic)	1.27	1.19	2.46	0.36
Native Hawaiian/Other Pacific Islander	0.66	0.96	1.63	0.16
Asian (Non-Hispanic)	1.22	0.85	2.08	0.35
White and Other Races (Non-Hispanic)	0.79	0.97	1.76	0.22
Age and Sex				
Under 10 years	0.96	1.36	2.33	0.03
10-17 years	1.05	1.45	2.50	0.11
18-29 years				
Male	1.03	1.32	2.35	0.88
Female	1.11	1.53	2.64	0.80
30-49 years				
Male	0.89	0.63	1.52	0.17
Female	0.89	0.61	1.51	0.06
50 years and over				
Male	0.88	0.92	1.81	0.15
Female	0.84	0.74	1.58	0.18

NOTE: See Appendix E (Table E.3) for definitions of race/ethnicity domains. Housing unit to housing unit duplications include duplications to reinstated units.

SOURCE: Mule (2002a:Tables F1, F3, F5, F7).

duplications among age/sex groups were for men ages 18–29 (0.9 percent, mostly in college dormitories and prisons) and for women ages 18–29 (0.8 percent, almost entirely in college dormitories).

At least two factors contributed to the large number of duplicates in 2000. One factor—the problem-plagued development of the MAF from new and multiple sources—has already been discussed. The second factor had to do with the "usual residence" rules for the census—each U.S. resident is supposed to be enumerated once at his or her usual residence—and how these rules did or did not match people's living situations. The rules were often not explained on the questionnaires or were unclear. In some instances, respondents did not want to follow the rules as stated. The result was often duplication of enumerations. For example, many college dormitory residents and people in prisons were counted at those locations, according to Census Bureau rules, but they were also reported by their families back home, counter to the instructions on the questionnaire. Some divorced parents with joint custody reported the children at both parents' homes, and people with two houses (e.g., in New York and Florida) were sometimes counted in both locations. (Such double counting probably occurred mainly in follow-up operations, when enumerators at the second house were told that the owners lived there.)

From the Evaluation Follow-Up Study and the Further Study of Person Duplication, it became evident not only that the census residence rules were not always recognized in the census, but that they were not always recognized in the A.C.E., either. Consequently, the original A.C.E. substantially underestimated census duplicates. In turn, corrections for undetected duplications substantially reduced the net undercount, particularly for blacks and young people.

6–D WHAT CAN WE CONCLUDE ABOUT COVERAGE ERROR IN 2000?

Our conclusions about coverage error in the 2000 census address eight issues. They are: (6–D.1) the quality of the A.C.E. Revision II analysis and documentation; (6–D.2) comparability with the 1990 PES; (6–D.3) net coverage error at the national level and for major population groups; (6–D.4) net coverage error for subnational areas; (6–D.5) net coverage error for group quarters residents; (6–D.6) gross coverage errors; (6–D.7) comparisons with demographic analysis; and (6–D.8) the Census Bureau's decision not to use the Revision II estimates to adjust the census counts that form the base for updated population estimates over the decade. Section 6–D.9 summarizes our findings.

6–D.1 A.C.E. Revision II Estimation and Documentation

From the outset of the 2000 Accuracy and Coverage Evaluation Program, the Census Bureau commendably dedicated the staff and resources needed to meet high standards for quality of implementation, evaluation, and documentation. In particular, when it became clear in early 2001 that the original A.C.E. estimates required further evaluation before the results could be considered for adjustment purposes, the Bureau speeded up planned evaluations and mounted additional evaluations to address its concerns. Then, when the results of these evaluations suggested that the original A.C.E. substantially underestimated the number of duplicates and other erroneous census enumerations, the Bureau devoted yet additional resources to a full reestimation of the dual-systems population and net undercount estimates by using the original A.C.E. and evaluation data. All of these efforts were commendable in the highest degree.

The A.C.E. Revision II work exhibited an outstanding level of creativity and productivity devoted to a very complex problem. It provided useful, previously unavailable information about problems of erroneous enumerations in the census and the original A.C.E. from such evaluations as the Further Study of Person Duplication, for which algorithms were developed to permit nationwide matching of A.C.E. cases with census enumerations.

The Revision II work, however, shed no light on additional omissions from the census that the A.C.E. may have missed (beyond what could be inferred from sex ratios for a limited number of population groups for which there were demographic analysis estimates). Evaluation studies provided estimates of measurement error in the P-sample, such as misclassification of match and nonmatch status and of residency status. However, time and resource constraints did not permit investigation of methods to estimate census omissions that the P-sample did not identify. One such method, for example, could be to match administrative records with E-sample and P-sample records.[19]

Part of the Revision II estimation included a commendable effort to estimate bias, as well as variance, in the estimates. However, this effort was necessarily limited, because most of the available evaluation data were used in Revision II directly. The Census Bureau's description of the Revision II methods identified areas of weakness that could not be assessed in the bias evaluations. At the national level for population groups, weaknesses included the uncertain and limited nature of the available data with which to adjust for correlation bias and the uncertainty in the selection of a model with which to determine Census Day residence for P-sample cases that linked to a census enumeration outside the A.C.E. search area (U.S. Census Bureau, 2003c:48). For subnational estimates (see Section 6–D.4), the decision to use separate E-sample and P-sample poststrata could have increased error for some small places.

The identification of potential weaknesses in Revision II exemplifies the Census Bureau's praiseworthy thoroughness of documentation and explanation for every step of the effort. Complete documentation was prepared as well for every component of the original A.C.E. and evaluation studies and for the preliminary revised (October 2001) estimates. Commendably, the Bureau produced the extensive A.C.E. documentation in a timely manner: most documentation was released at the same time as or very shortly after each adjustment decision, in March 2001, October 2001, and March 2003.

6–D.2 Comparability with the 1990 PES

The original A.C.E. was comparable in design and execution in most respects to the 1990 PES. Most changes in the A.C.E. design and operations were intended to facilitate timely, accurate data collection and matching and to reduce variance. The two programs did differ in coverage: the PES covered the household and noninstitutional group quarters population, while the A.C.E. covered the household population; the two programs also used somewhat different defini-

[19]Research related to such a triple-systems estimation was conducted as part of the 2000 Census Administrative Records Experiment (see Bauder and Judson, 2003; Berning and Cook, 2003; Berning, 2003; Heimovitz, 2003).

tions of race/ethnicity domains and a different treatment of movers (PES-B for the PES, PES-C for the A.C.E.).

In contrast, five features of the A.C.E. Revision II methodology significantly impaired comparability with the PES (see Mule, 2003:Table 1):

(1) the use in an expanded DSE formula of results from nationwide matching of the E-sample and the census on name and birthdate to estimate duplicate enumerations outside the A.C.E. search area, which reduced the estimated net undercount by 2.8 million people;

(2) the use of Evaluation Follow-Up Study results to reestimate correct enumerations among E-sample cases not linked to census enumerations outside the A.C.E. search area, which reduced the estimated net undercount by 2.4 million people;

(3) the use of results from nationwide matching of the P-sample and the census to estimate deletions from the P-sample for nonmover resident cases that linked to census enumerations outside the A.C.E. search area, which reduced the estimated net undercount by 1.1 million people;

(4) the use of Evaluation Follow-Up and Matching Error Study results to reestimate residents and matches among P-sample cases not linked to census enumerations outside the A.C.E. search area, which increased the estimated net undercount by 0.01 million people; and

(5) the use of sex ratios from demographic analysis for black men age 18 and over and nonblack men age 30 and over to adjust for correlation bias for these groups, which increased the estimated net undercount by 1.7 million people (0.8 million black men and 0.9 million nonblack men).

It would be possible (although probably not feasible) to reestimate the 1990 PES dual-systems estimates to include the fifth item (adjustment for correlation bias) and to use the results of the PES Evaluation Follow-Up Study to adjust the PES for E-sample and P-sample measurement error. However, the lack of a question on month of birth in the 1990 census and the fact that names were captured only for E-sample cases preclude the possibility of nationwide matching of the E-sample or P-sample to census enumerations in 1990.

One could only speculate about the likely results if all five changes in the A.C.E. Revision II methodology could be implemented for the PES. The correlation bias adjustments might be similar, given that sex ratios were similar for most groups in 1990 and 2000 for demographic analysis, the two censuses, and the A.C.E. and PES (refer to Table 6.4 above). Measurement error corrections might be similar also, given that evaluations estimated data quality advantages for the A.C.E. on some dimensions and for the PES on others (e.g., a smaller percentage of movers and a somewhat higher quality of matching in the A.C.E., but a somewhat higher household interview rate in the PES—see Sections 6–A.3, 6–A.5, and 6–A.8). Whether a Further Study of Person Duplication would result in such a large number of additional duplicate enumerations not detected by the PES as turned out to be the case for the A.C.E. is highly speculative.

6–D.3 Net Coverage Error in 2000, Nation and Poststrata

From the A.C.E. Revision II work, it appears that net undercount rates were lower in 2000 compared with the 1990 rates estimated by the PES and that differences in net undercount rates between historically less-well-counted groups (minorities, renters, and children) and better counted groups were smaller as well. Thus, the A.C.E. Revision II and PES estimates of the national net undercount were a 0.5 percent net overcount and a 1.6 net undercount, respectively. Estimates of differences in net undercount rates between Hispanic and non-Hispanic white domains were 1.8 percentage points in Revision II and 4.3 percentage points in PES; estimates of differences in net undercount rates between blacks and whites were 3 percentage points in Revision II and 3.9 percentage points in PES; and estimates of differences in net undercount rates between owners and renters were 2.4 percentage points in Revision II and 4.5 percentage points in PES. The smaller decline in differential net undercount for blacks and whites compared with the declines for Hispanics and whites and for owners and renters was because of the large size of the correlation bias adjustment for black men, which increased their net undercount estimate.

It is also apparent that whole-person imputations played a large role in explaining the reason why the original A.C.E. estimated lower net undercount rates than the PES despite similar match and correct enumeration rates. In turn, census duplications and other erroneous enumerations not detected by the original A.C.E. played a major role in the further reduction in estimated net undercount rates in Revision II from the original A.C.E.

Beyond that, it is hard to draw definitive conclusions about trends in coverage error from 1990 to 2000 because of the significant differences in the methods for Revision II compared with the PES. Given that the original A.C.E. also estimated lower net undercount rates than the PES and smaller differences in net undercount rates among population groups, we are fairly confident in concluding that net undercount and differences in net undercount rates were, by and large, reduced in 2000 from 1990. We are also fairly confident, despite the considerable reductions in estimated net undercounts from the original A.C.E. to Revision II, that differences in net undercount rates between such groups as minorities and others and owners and renters remain. Beyond these general statements, we cannot be more specific. We do not know the effect on the PES net undercount estimates that would have resulted from implementation of the changes in methods for A.C.E. Revision II.

6–D.4 Coverage Error in 2000, Subnational Areas

Assessment of net undercount rates and differences in net undercount for states and other geographic areas is an important part of census coverage evaluation because of the many uses of the data for small-area analysis. However, it is difficult to estimate error in subnational estimates, which are constructed by a synthetic method. In that procedure, coverage correction factors (the DSE estimate divided by the census count) are developed for individual poststrata; these factors are then multiplied by the population in each poststratum in each census block, the results rounded to whole people, and the rounded population estimates summed by block and, in turn, summed for larger geographic areas. This method makes the strong assumption that coverage probabilities and errors in estimation do not vary markedly across geographic areas within any poststratum.

As discussed in Section 6–B.6, the reduction in net undercount rates and differences in net undercount rates for population groups estimated in 2000 compared with 1990 probably had the effect of lowering net undercount rates and differences in net undercount rates in 2000 for geographic areas as well. There is some evidence on this point, but the limited available comparisons are affected by the significant differences in methodology between the Revision II A.C.E. and the 1990 PES.

For states, it appears that differences in estimated net undercount and overcount rates were smaller in 2000 compared with 1990. Thus, the estimated net undercount (overcount) rates for states in 2000 from A.C.E. Revision II spanned a range of 2.2 percentage points (from Minnesota with the largest net overcount rate of 1.7 percent to Nevada with the largest net undercount rate of 0.5 percent; see Schindler, 2003:Table 1). By comparison, the estimated net undercount rates for states in 1990 from the revised PES spanned a range of 3 percentage points (from Rhode Island with the smallest net undercount rate of 0.1 percent to New Mexico with the largest net undercount rate of 3.1 percent; see Bureau of the Census, 1992:Attachment 4). There were similarities in patterns of coverage error among states between 2000 and 1990 and clustering by region. Most striking, most of the midwestern states were not only in the quartile with the highest estimated net overcount rates in 2000 (exceeding 1.1 percent) but also in the quartile with the lowest estimated net undercount rates in 1990 (smaller than 0.7 percent). Other states with high net overcount rates in 2000 and low net undercount rates in 1990 were in the Northeast region. At the other end of the distribution, a group of nine states in the South and West were not only in the quartile with estimated net undercount rates or the lowest net overcount rates in 2000 (below 0.1 percent net overcount), but also in the quartile with the highest estimated net undercount rates in 1990 (exceeding 2 percent): Georgia, Maryland, Louisiana, Texas, Colorado, Montana, Nevada, New Mexico, and California.

For large counties of 1 million or more population (33 in 2000, 30 in 1990), differences in estimated net undercount and overcount rates were also smaller in 2000 compared with 1990. In 2000, estimates ranged from net undercounts of less than 2 percent to net overcounts of less than 2 percent, with 27 counties having rates between 1 percent net undercount and 1 percent net overcount (U.S.

Census Bureau, 2003c:Table 11). In 1990, the range was wider, from a net undercount of 4.9 percent to a net overcount of 0.8 percent, and only 16 counties were in a range of 2 percentage points (from 1 percent to 3 percent net undercount; Bureau of the Census, 1992:Attachment 12). For large places with 100,000 or more population, the story is similar (compare U.S. Census Bureau, 2003c:Table 10, with Bureau of the Census, 1992:Attachment 11). Whether estimated net overcount and undercount rates were smaller in 2000 than in 1990 for smaller areas is not clear, particularly given the very large net overcounts estimated for some small counties and places, as discussed below. Interested researchers can examine Revision II A.C.E. estimated net overcount and undercount rates for counties and places of all sizes from data files that are available at www.census.gov/dmd/www/ACEREVII_COUNTIES.txt and www.census.gov/dmd/www/ACEREVII_PLACES.txt, with record layouts described in Schindler (2003).[20] However, similarly detailed information is readily available for 1990 only for counties.

The A.C.E. Revision II work included a loss function analysis to assess the relative accuracy of the Revision II estimates and census figures for population levels and shares for counties and places nationwide and within state (see Mulry and ZuWallack, 2002). The loss function analysis used estimates of sampling variance and nonsampling bias and variance to compute the weighted mean square error of the Revision II estimates. The analysis indicated that the Revision II estimates were more accurate than the census for every loss function considered except for places with 100,000 or more people, for which the error appeared to be concentrated in the Revision II estimates for the nine places with 1 million or more people (U.S. Census Bureau, 2003c:42). The Revision II loss function analysis was greatly improved over the analysis conducted for the original A.C.E. estimates, but it, too, did not include all sources of error, excluding such potentially important sources as the errors in the various evaluation studies and in the correlation bias adjustments (see U.S. Census Bureau, 2003c:43–44,68; see also Section 6–D.1).

[20]The data files use census *collection* geography (the definitions and boundaries set before the census) rather than the finished 2000 census *tabulation* geography.

Of particular relevance for subnational estimates, the loss function analysis did not include any estimate of synthetic error. A simulation analysis, using artificial populations, that modeled patterns of coverage variation within poststrata was carried out to assess the effects on the loss function analysis for states, counties, and places of omitting consideration of synthetic error. The results did not in general contradict the loss function results, but the simulation analysis itself had limitations, principally that the variables chosen to construct artificial populations (e.g., people with two or more items imputed) did not correlate highly with estimated gross undercount or overcount (Griffin, 2002).

An examination of the A.C.E. Revision II subnational estimates by population size for counties and places identified high estimated net overcounts for some small counties and places. For example, 863 places with populations of fewer than 1,000 people (of a total of 10,421 such places) had estimated net overcount rates of 5 percent or more (106 had estimated net overcount rates of 10 percent or more). In contrast, places with 1,000 or more people rarely had estimated net overcounts of more than 5 percent, and places with 100,000 or more people had estimated net overcounts of no more than 2 percent (U.S. Census Bureau, 2003c:Table 10; Table 11 provides the same information for counties). A possible explanation would credit the high estimated net overcount rates for some small places to the presence of proportionately larger numbers of proxy census enumerations in these places. Proxy enumeration poststrata had low estimated correct enumeration rates, but there were no corresponding P-sample poststrata, so the use of correct enumeration rates for proxy enumerations and match rates spread over proxy and nonproxy enumerations would overestimate net overcount rates in these places if proxy enumerations also exhibited low match rates.

We note that accuracy of population estimates cannot be expected for very small geographic areas, such as blocks, whether the data are from an adjustment procedure or the census. A principal reason for low-quality block-level census counts is geocoding errors; that is, putting households in the wrong location (e.g., locating an apartment building on the opposite side of the street and hence in a different block from its true location). Research that expanded on the analysis of geocoding errors in the Housing Unit Coverage Study estimated that almost 5 percent of housing units in the 2000 census were geocoded to the wrong block, rising to as much as 11 percent of housing units in multiunit structures with 10 or more units (Vitrano et al., 2003b:76–77; see also Section 4–E.2). Hence, the usefulness of block statistics is not the data themselves, but the facility they provide for the user to aggregate the data to larger areas defined by the user (e.g., congressional districts, local service areas). Because geocoding errors typically involve nearby areas, the data for aggregates of blocks will be more accurate than the data for individual blocks.

6–D.5 Coverage Error in 2000, Group Quarters

We can say virtually nothing about coverage error for group quarters residents—either erroneous enumerations or omissions. Net undercount estimates for residents of noninstitutional group quarters from the 1990 PES were based on very uncertain data because of the difficulty of tracking such populations as college students on spring or summer break. Net undercount estimates for group quarters residents were not available from the A.C.E., which did not include this population in its universe. All indications are that the group quarters enumeration process was poorly controlled (see Section 4–F), so that coverage errors for group quarters residents were probably large.

6–D.6 Gross Coverage Errors

Although coverage correction factors for adjustment are based on estimated net error rates, components of gross error—that is, types of census omissions and erroneous enumerations—are important to measure and analyze to understand the census process and how to improve it. In this regard, higher or lower net undercount does not relate directly to the level of gross errors. There can be a zero net undercount and high rates of gross omissions and gross erroneous enumerations.

However, there is not widespread acceptance of the definition of different types of gross errors (see Section 1–C.1). Moreover, some types of gross errors depend on the level of aggregation or are not clearly identified by type in the design used for the A.C.E. and PES. Many errors identified by the A.C.E. or PES involved the balancing of a nonmatch on the P-sample side against an erroneous enumeration on the E-sample side—for example, when an E-sample case that should match was misgeocoded to another block. These kinds of balancing errors were not errors for such levels of geography as counties, cities, and even census tracts, although they affected error at the block cluster level. Also, the classification of type of gross error in the A.C.E. or PES was not necessarily clean. For example, without the nationwide matching used for A.C.E. Revision II, a duplicate enumeration involving an E-sample unit and a census enumeration in another state (perhaps winter and summer homes) would be classified as an "other residence" erroneous enumeration and not as a duplicate.

Gross errors in the original A.C.E. were smaller in number than gross errors in the PES. The original A.C.E. estimated 28.3 million gross errors of all types, including 12.5 million erroneous census enumerations and 15.8 million census omissions. By comparison, the PES estimated 36.6 million gross errors of all types, including 16.3 million erroneous census enumerations and 20.3 million census omissions (Anderson and Fienberg, 2001:Table 3). The A.C.E. Revision II estimated more gross errors than the original A.C.E. but still fewer than the PES; it estimated 33.1 million gross errors of all types, including 17.2 million erroneous census enumerations and 15.9 million census omissions.[21] Many of these errors, as noted, were not consequential for larger geographic areas.

6–D.7 Comparison with Demographic Analysis

The use of sex ratios from the revised demographic analysis 2000 estimates to adjust net undercount rates for adult men in the A.C.E. Revision II estimation produced net undercount patterns for age and sex groups that more closely resembled the demographic analysis patterns than if no such adjustments had been made (see Table 6.11). The biggest remaining discrepancy was for children ages 0–9 years, for which demographic analysis estimated a sizeable net undercount

[21]Census omissions are estimated as a residual, by adding the net undercount estimate to the estimate of erroneous enumerations, which for Revision II is the original A.C.E. estimate plus 4.7 additional estimated erroneous enumerations (Mule, 2003:2).

compared with a small net overcount estimate from A.C.E. Revision II. No explanation for this discrepancy has been advanced.

Demographic analysis estimates themselves are subject to considerable uncertainty from such sources as the estimates for immigrants (particularly illegal immigrants). The revised demographic analysis estimates released in October 2001 used 2000 census long-form-sample data to estimate components of immigration and so are not independent of the census. In addition, they included revised assumptions for births and coverage of immigrants in the census that are based primarily on expert judgment. For example, expert judgment was used to conclude that registration of births after 1984 should be assumed to be 100 percent complete (see Robinson, 2001b:11). Consequently, the small correction for estimated net underregistration of births after 1984 was omitted from the revised estimation. Such judgments may be reasonable, but they retain sufficient uncertainty so that it is not appropriate to conclude that the revised demographic estimates are necessarily more accurate than the census or the A.C.E. Revision II.

6–D.8 March 2003 Decision Not to Adjust Base for Postcensal Estimates

The Census Bureau decided not to use the A.C.E. Revision II population estimates to adjust the 2000 census results for estimated coverage error and instead to base the Bureau's postcensal small-area population estimates program on the unadjusted census counts. The program includes periodic release of updated estimates for states, counties, places, and school districts (see Citro, 2000e), so an adjustment of the 2000 base counts would need to be made for small areas to be useful for the estimates program. For updated population controls for the major household surveys, such as the Current Population Survey, an adjustment of the 2000 base counts would only need to be made for national-level population groups. However, no such adjustments are currently planned.

The three major factors cited by the Bureau in its decision against adjustment were: the uncertainty about the correlation bias adjustment; the possible errors for small places from the construction of disparate E-sample and P-sample poststrata; and the discrepancies between demographic analysis and A.C.E. Revision II net under-

Table 6.11 Estimated Net Undercount Rates (percents), Original 2000 A.C.E. (March 2001), Revised Demographic Analysis (October 2001), and A.C.E. Revision II (March 2003) by Race, Sex, and Age

Category	Original A.C.E.	Revised Demographic Analysis (March 2001)	A.C.E. Revision II
Total	1.15	0.12	−0.48
Black Male			
All Ages	2.36	5.15	4.19
Under 10 years[a]	2.92	3.26	0.72
10–17 years	2.92	−1.88	−0.59
18–29 years	3.82	5.71	6.14
30–49 years	2.58	9.87	8.29
50 years and over	−0.68	3.87	2.43
Black Female			
All Ages	1.77	0.52	−0.61
Under 10 years[a]	2.95	3.60	0.70
10–17 years	2.95	−1.20	−0.55
18–29 years	3.76	−0.66	0.00
30–49 years	1.26	1.28	−0.40
50 years and over	−0.84	−1.03	−2.51
Nonblack male			
All Ages	1.40	0.21	−0.19
Under 10 years[a]	1.28	2.18	−0.68
10–17 years	1.28	−2.01	−1.46
18–29 years	3.39	−0.63	0.19
30–49 years	1.70	0.63	1.05
50 years and over	−0.20	0.14	−1.10
Nonblack female			
All Ages	0.64	−0.78	−1.41
Under 10 years[a]	1.28	2.59	−0.68
10–17 years	1.28	−1.55	−1.44
18–29 years	1.83	−1.94	−1.54
30–49 years	0.91	−1.01	−0.63
50 years and over	−0.75	−1.18	−2.42

NOTES: Net undercount is the difference between the estimate (A.C.E. or demographic analysis) and the census, divided by the estimate. Minus sign (−) means a net overcount. Population is total population, including household members and group quarters residents.

[a] In the original A.C.E., children ages 0–17 were a single category.

SOURCE: U.S. Census Bureau (2003c:Tables 13, 14); Robinson (2001a:Table 4, Table 7, column labeled A.C.E. Model 1).

count estimates for children. We agree that these are important sources of concern about the reliability of basing adjustments for geographic areas on the A.C.E. Revision II results. We also have other concerns about the Revision II estimates. These concerns do not reflect on the high quality, exceptional effort, and innovative nature of the Revision II estimation, but rather stem primarily from the limitations of the data available for that work.

Generally, we are concerned that the Revision II results are too uncertain to be used with sufficient confidence about their reliability for adjustment of census counts for subnational geographic areas and population groups. We have identified at least six sources of uncertainty.

First, only small samples of the A.C.E. data (see Table 6.3) were available to provide estimates of classification of E-sample cases as erroneous and of P-sample cases as nonresidents in the A.C.E. search area. These subsamples were used to develop correction factors to apply to the full A.C.E. original samples, but these factors were subject to estimation error and were developed for aggregates of poststrata, not the full set of individual E-sample and P-sample poststrata.

Second, it was not possible, in many instances, to determine which of each pair of duplicates involving a census enumeration outside the A.C.E. search area in the E-sample component of the Further Study of Person Duplication was correct and which should not have been counted in the census. (The exceptions, based on Census Bureau residency rules, involved such cases as a duplication between the E-sample and a census group quarters enumeration.) Similarly, it was not possible to determine whether the P-sample case in a pair of duplicates involving a census enumeration outside the A.C.E. search area in the P-sample component of the Further Study of Person Duplication was correct or whether the census enumeration was correct and, consequently, the P-sample case was a nonresident and should be dropped from the sample. For residency status, it was not even clear whether the probability of being a resident or nonresident should be one-half or some other proportion.

Third, the correction for duplicate census enumerations with E-sample cases and with P-sample cases involved use of two separate evaluations—the EFU reanalysis for cases inside the A.C.E. search area and the Further Study of Person Duplication for cases outside the search area. Differences in the methods of these two analyses could introduce biases into the estimates. In addition, there is evidence that the numbers of duplicates outside the A.C.E search area were underestimated in both studies for both the E-sample and the P-sample.

Fourth, the correlation bias adjustments incorporated strong assumptions that are not easily supported. A first assumption was that the DSE estimates for women (and children) were unbiased. The Census Bureau believes that the corrections for other known biases from the duplication studies and other analyses addressed the concern about possible bias in the DSE estimates for women, and the Revision II estimates for women accord reasonably well with the revised demographic analysis estimates (see Table 6.11). However, there remains a sizeable discrepancy between the Revision II estimates for children ages 0–9 and the revised demographic analysis estimates. A second assumption was that the adjustments for black men varied only by age group and not also by such categories as housing tenure, when it is plausible from findings about higher net undercount rates for renters compared with owners that correlation bias affected black male renters differently from black male owners. A third assumption was that the adjustments for nonblack men applied equally to all race/ethnicity groups in that broad category, when it is plausible that correlation bias affected Hispanics and other race groups differently from non-Hispanic whites. Of course, it was the absence of data that precluded the Census Bureau from making correlation bias adjustments for groups other than those defined by age and the simple dichotomy of blacks and all others.

Fifth, the Census Bureau had to pick a particular correlation bias adjustment model, but noted that alternative adjustment models could have been used. The selected model and the alternatives would all have produced estimates that were consistent with the demographic analysis sex ratios and the A.C.E. Revision II data at the national level, but they would have produced different subnational DSE estimates. The loss function analysis does not take account of the potential error from the choice of the correlation bias adjustment model, so we do not know the possible effects of this choice on subnational estimates.

Sixth, the use of different poststrata for the E-sample and the P-sample in the Revision II estimation could have increased the error for some geographic areas. We understand that the revised E-sample poststrata better explained variations in correct enumeration rates compared with the original E-sample poststrata. However, there were no logical counterparts on the P-sample side for some of the E-sample poststrata (including those based on proxy response and type of return), and the use of different poststrata could have introduced bias for some estimates.

Because of these sources of uncertainty, our view is that the Census Bureau's decision not to use the A.C.E. Revision II estimates to adjust the census data that provide the basis for postcensal estimates was justified. A consideration in our agreement with the bureau's decision against adjustment was the finding that estimated net undercount rates and differences in net undercount rates for population groups (and, most probably, subnational areas) were smaller in 2000 than in 1990. The smaller the measured net coverage errors in the census, the smaller would be the effects of an adjustment on important uses of the data. Because the benefits of an adjustment would be less when net coverage errors are small, a high level of confidence is needed that an adjustment would not significantly increase the census errors for some areas and population groups. In our judgment, the A.C.E. Revision II estimates, given the constraints of the available data for correcting the original A.C.E. estimates, are too uncertain for use in this context. We do not intend by this conclusion, however, to set a standard of perfection whereby it would never be possible to carry out an adjustment that improved on the census counts. Indeed, had it been possible to implement the A.C.E. Revision II methodology from the outset on the original A.C.E. data and to make some other improvements in the estimation (see Section 6–E), it is possible that an adjustment of the 2000 census data could have been implemented that was well supported.

6–D.9 Revision II Coverage Evaluation Findings

See also Section 6–A.11.

Finding 6.2: The Census Bureau commendably dedicated resources to the A.C.E. Revision II effort, which completely reestimated net undercount (and overcount) rates for several hundred population groups (poststrata) by using data from the original A.C.E. and several evaluations. The work exhibited high levels of creativity and effort devoted to a complex problem. From innovative use of matching technology and other evaluations, it provided substantial additional information about the numbers and sources of erroneous census enumerations and, similarly, information with which to correct the residency status of the independent A.C.E. sample. It provided little additional information, however, about the numbers and sources of census omissions.

Documentation for the original A.C.E. estimates (March 2001), the preliminary revised estimates (October 2001), and the A.C.E. Revision II estimates (March 2003) was timely, comprehensive, and thorough.

Finding 6.3: We support the Census Bureau's decision not to use the March 2003 Revision II A.C.E. coverage measurement results to adjust the 2000 census base counts for the Bureau's postcensal population estimates program. The Revision II results are too uncertain to be used with sufficient confidence about their reliability for adjustment of census counts for subnational geographic areas and population groups. Sources of uncertainty stem from the small samples of the A.C.E. data that were available to correct components of the original A.C.E. estimates of erroneous enumerations and non-A.C.E. residents and to correct the original estimate of nonmatches and the consequent inability to make these corrections for other than very large population groups; the inability to determine which of each pair of duplicates detected in the A.C.E. evaluations was correct and which should not have been counted in the census or included as an A.C.E. resident; the possible errors in subnational estimates from the choice of one of several alternative correlation bias adjustments to compensate for higher proportions of missing men relative to women; the inability to make correlation bias adjustments for population groups other than blacks and nonblacks; and the possible errors for some small areas from the use of different population groups (post-strata) for estimating erroneous census enumerations and census omissions. In addition, there is a large discrepancy in coverage estimates for children ages 0–9 when comparing demographic analysis estimates with Revision II A.C.E. estimates (2.6 percent undercount and 0.4 percent net overcount, respectively).

Finding 6.4: Demographic analysis helped identify possible coverage problems in the 2000 census and in the A.C.E. at the national level for a limited set of population groups. However, there are sufficient uncertainties in the revised estimates of net immigration (particularly the illegal component) and the revised assumption of completeness of birth registration after 1984, compounded by the difficulties of classifying people by race, so that the revised demographic analysis estimates cannot and should not serve as the definitive standard of evaluation for the 2000 census or the A.C.E.

Finding 6.5: Because of significant differences in methodology for estimating net undercount in the 1990 Post-Enumeration Survey Program and the 2000 Accuracy and Coverage Evaluation Program (Revision II), it is difficult to compare net undercount estimates for the two censuses. Nevertheless, there is sufficient evidence (from comparing the 1990 PES and the original A.C.E.) to conclude that the national net undercount of the household population and net undercount rates for population groups were reduced in 2000 from 1990 and, more important, that differences in net undercount rates between historically less-well-counted groups (minorities, children, renters) and others were reduced as well. From smaller differences in net undercount rates among groups and from analysis of available information for states and large counties and places, it is reasonable to infer that differences in net undercount rates among geographic areas were also probably smaller in 2000 compared with 1990. Despite reduced differences in net undercount rates, some groups (e.g., black men and renters) continued to be undercounted in 2000.

Finding 6.6: Two factors that contributed to the estimated reductions in net undercount rates in 2000 from 1990 were the large numbers of whole-person imputations and duplicate census enumerations, many of which were not identified in the original (March 2001) A.C.E. estimates. Contributing to duplication were problems in developing the Master Address File and respondent confusion about or misinterpretation of census "usual residence" rules, which resulted in duplication of household members with two homes and people who were enumerated at home and in group quarters.

6–E RECOMMENDATIONS FOR COVERAGE EVALUATION IN 2010

6–E.1 An Improved Accuracy and Coverage Evaluation Program

The complexities of the original A.C.E. and Revision II reestimation and the uncertainties about what the Revision II results tell us about net and gross coverage errors in the 2000 census could lead policy makers to question the value of an A.C.E.-type coverage evaluation program for the 2010 census. To the contrary, we recommend that research and development for the 2010 census give priority to improving an A.C.E.-type coverage evaluation mechanism and that it be implemented in 2010.

Without the 2000 original A.C.E. and Revision II estimation, we would not have acquired so much valuable information about strengths and weaknesses of the 2000 census. In particular, differences between the census counts, the original A.C.E., and the original demographic analysis estimates spurred the development of innovative methods for identifying duplicate census enumerations. These differences also motivated a reexamination of assumptions about immigration estimates in demographic analysis.

The plans for the 2010 census include the serious possibility that the matching methods used in the Further Study of Person Duplication would be used as part of the enumeration process, so that duplicate enumerations could be identified, followed up, and eliminated from the census counts in real time (Smith and Whitford, 2003).[22] If these plans reach fruition, then the 2010 census could be expected to have many fewer erroneous enumerations than the 2000 census. Because it is difficult to imagine the development of effective new ways of reducing census omissions, then a reduction in erroneous enumerations could well result in a significant net undercount in the 2010 census and an increase in differential undercoverage among population groups. Without an A.C.E.-type coverage evaluation program, it would not be possible to measure such an undercount or to adjust some or all of the census counts for coverage errors should that be warranted. Demographic analysis, while providing useful information about census coverage at the national level for a few population groups, could not substitute for an A.C.E.

We urge that the 2010 census testing program give priority to research and development for an improved A.C.E.-type coverage evaluation program. We see possibilities for improvements in many areas, such as the estimation of components of gross census error as well as net error, expansion of the search area for erroneous census enumerations and P-sample nonresidents, the inclusion of group quarters residents, better communication to respondents of residence rules (and reasons for them), understanding the effects of *II*s on A.C.E. estimation, the treatment of movers, and the development of poststrata. The optimal size of a 2010 A.C.E. is also a consideration. The survey must be large enough to provide estimates of coverage errors with the level of precision that was targeted for the original (March 2001) A.C.E. estimates for population groups and geographic areas.

With regard to the estimates of erroneous enumerations and P-sample nonresidents outside the traditional search area, the nationwide matching technology developed for the 2000 A.C.E. Revision II would make it possible to incorporate the search for such errors

[22]Some observers may be concerned about privacy issues with regard to the capture of names on the computerized census records and their use for matching on such a large scale. The panel did not consider this issue, but we note that the Census Bureau has always been sensitive to such concerns, and Title 13 of the U.S. Code safeguards the data against disclosure.

as part of the 2010 A.C.E. production process, not waiting for subsequent evaluation. Such cases could be identified and followed up in real time, similar to what is planned for the census itself. Such a procedure could not only significantly reduce errors in the A.C.E., it could also provide valuable information about types of gross errors that has not previously been available.

The nationwide matching technology together with the possible increased use of administrative records for group quarters enumeration, could make it possible to include group quarters residents in the 2010 A.C.E. with an acceptable level of data quality. Administrative records for such group quarters as college dormitories, prisons, nursing homes, and other institutions could provide home addresses for use in the matching to identify duplicate enumerations. With this information, the follow-up to verify a duplicate enumeration of a college student, for example, would simply need to establish that the student was in fact the same person, and the census residence rules would then be applied to designate the group quarters enumeration as correct and the home enumeration as erroneous. There would be no need to make the family choose the student's "usual residence."

With regard to communication of residence rules, cognitive research on the A.C.E. questionnaires and interviewer prompts could lead to interviewing procedures that better help respondents understand the Bureau's rules and reasons for them. The 2000 A.C.E. demonstrated that it is not enough to rely on respondents' willingness to follow the rules (e.g., when parents report a college student at home), which is a major reason for incorporating nationwide matching technology into the 2010 A.C.E. process. However, cognitive research could probably improve the interview process in ways that would improve the quality of residence information reported to the A.C.E. (Such research is also likely to improve the census enumeration.)

Furthermore, if plans to use mobile computing devices and global positioning system (GPS) technology for address listing and nonresponse follow-up in 2010 come to fruition, then there is the opportunity to reduce geocoding errors in the E-sample. Such technology could also be used to minimize geocoding errors in the listing operations conducted to build the independent P-sample address list.

With regard to understanding the effects of census records that are excluded from the A.C.E. matching (*II*s), such records in 2010 would probably be whole-person imputations. The large number of records that were reinstated from the special summer 2000 housing unit unduplication operation should not affect 2010, given that matching to detect duplicate addresses (and people) will probably be built into the enumeration process. However, there could still be large numbers of whole-person and whole-household imputations. In order to more fully understand the effects of an adjustment for population groups and geographic areas, it is important to analyze the distribution of such imputations and, in particular, how they may affect synthetic error.

For movers, with the nationwide matching capability that has been developed, it should be possible to use the PES-B procedure for a 2010 A.C.E., instead of the cumbersome PES-C procedure that was used in 2000. The speed of the 2000 P-sample interviewing reduced the number of movers and, consequently, their effects on the dual-systems estimates, but the quality of their data was markedly below that for nonmovers and inmovers. Finding where inmovers resided on Census Day would be greatly facilitated by nationwide matching, so that a PES-B procedure would be feasible and likely to provide improved data quality compared with its use in the 1990 PES.

Finally, with regard to poststratification, the Revision II effort to revise the E-sample poststrata on the basis of analyzing the A.C.E. data themselves was commendable. Further work on poststratification could be conducted with the 2000 data in planning for 2010, and plans for using the 2010 A.C.E. data to refine the poststrata could also be developed. Care should be taken to develop poststrata that do not produce the anomalies that were observed in Revision II from the use of E-sample poststrata for proxy enumerations for which no counterparts were developed on the P-sample side. The use of statistical modeling for developing poststrata from the A.C.E. data should also be part of the poststratification research for 2010.

We are confident that these and other improvements could be made in an A.C.E.-type coverage evaluation program for the 2010 census with sufficient attention to research, development, and testing in the next few years. The U.S. General Accounting Office (2003a:1) reported that the Census Bureau "obligated about

$207 million to [the 2000 census A.C.E. and the predecessor ICM program] from fiscal years 1996 through 2001, which was about 3 percent of the $6.5 billion total estimated cost of the 2000 Census" (see also U.S. General Accounting Office, 2002b). An equivalent expenditure for an A.C.E.-type program in 2010 would be money well spent to ensure that adequate data become available with which to evaluate not only net, but also gross coverage errors. Such errors could be more heavily weighted toward omissions, and not erroneous enumerations, in 2010 compared with the 2000 experience.

Recommendation 6.1: **The Census Bureau and administration should request, and Congress should provide, funding for the development and implementation of an improved Accuracy and Coverage Evaluation Program for the 2010 census. Such a program is essential to identify census omissions and erroneous enumerations and to provide the basis for adjusting the census counts for coverage errors should that be warranted.**

The A.C.E. survey in 2010 must be large enough to provide estimates of coverage errors that provide the level of precision targeted for the original (March 2001) A.C.E. estimates for population groups and geographic areas. Areas for improvement that should be pursued include:

1. **the estimation of components of gross census error (including types of erroneous enumerations and omissions), as well as net error;**
2. **the identification of duplicate enumerations in the E-sample and nonresidents in the P-sample by the use of new matching technology;**
3. **the inclusion of group quarters residents in the A.C.E. universe;**
4. **improved questionnaire content and interviewing procedures about place of residence;**
5. **methods to understand and evaluate the effects of census records that are excluded from the A.C.E. matching (*II*s);**

6. **a simpler procedure for treating people who moved between Census Day and the A.C.E. interview;**
7. **the development of poststrata for estimation of net coverage errors, by using census results and statistical modeling as appropriate; and**
8. **the investigation of possible correlation bias adjustments for additional population groups.**

6–E.2 Improved Demographic Analysis for 2010

We support the usefulness of demographic analysis for intercensal population estimation and for helping to identify areas of possible enumeration problems in the census and the A.C.E. For this reason, it is important for the Census Bureau to continue its efforts to obtain additional funding for research and development of demographic analysis methods, particularly for estimates of immigrants, and to develop methods for estimating uncertainty in demographic analysis population estimates. Such developmental work needs to be conducted with other federal statistical agencies that have relevant data and that make use of postcensal population estimates.

> *Recommendation 6.2:* **The Census Bureau should strengthen its program to improve demographic analysis estimates, in concert with other statistical agencies that use and provide data inputs to the postcensal population estimates. Work should focus especially on improving estimates of net immigration. Attention should also be paid to quantifying and reporting measures of uncertainty for the demographic estimates.**

6–E.3 Time for Evaluation and Possible Adjustment

The original A.C.E. data collection, matching, estimation, and initial evaluation were carried out according to carefully specified and controlled procedures with commendable timeliness (see Finding 6.1 in Section 6–A.11). However, the experience with the subsequent evaluations and A.C.E. Revision II demonstrates that the development of net coverage estimates that are judged to be sufficiently reliable to use in evaluation of the census counts—and possible adjustment—should not be rushed. Similarly, even if the process for evaluating census operations and data items is improved relative to 2000 (see Chapter 9), that process—which is important to verifying the quality of the census content—requires ample time. Consequently, the panel believes that adequate evaluation of the census block-level data for congressional redistricting is not possible by the current deadline of 12 months after Census Day. The Congress should consider changing this deadline to provide additional time for evaluation and delivery of redistricting data.

> *Recommendation 6.3:* **Congress should consider moving the deadline to provide block-level census data for legislative redistricting to allow more time for evaluation of the completeness of population coverage and quality of the basic demographic items before they are released.**

Chapter 7

Assessment of Basic and Long-Form-Sample Data

The content of the 2000 census, as in past censuses, included basic demographic items plus a wide range of social and economic characteristics. The basic items (or complete-count items) were asked of everyone, whether they received the short or the long form; the additional items (or sample items) were asked of people selected for the long-form sample of about one-sixth of the population (see Appendix B for the list of items). The demographic items have widespread use, particularly as they form the basis of small-area population estimates that the Census Bureau develops for years following each census (see Section 2–C). The additional long-form-sample items on such topics as income, employment, education, occupation and industry, transportation to work, disabilities, housing costs, and others are used extensively by federal, state, and local government agencies, the private sector, academic researchers, the media, and the public (see Section 2–D).

Users need to understand the quality of the basic and the sample data to interpret census results appropriately. The Census Bureau needs to understand data quality to determine how best to improve census processes to produce high-quality information and how to inform users about its strengths and weaknesses. Past censuses provided a rich array of basic and long-form-sample data quality measures from studies of nonresponse, exact matches with surveys and administrative records, content reinterviews with samples of respondents, and experiments to determine response effects of alternative questionnaire formats and wording (see, e.g., Bureau of the Census, 1964, 1970, 1975a,b, 1982a,b, 1983a,b, 1984). To date, data quality measures are somewhat sparser for the 2000 census.

The panel requested and received detailed tabulations of basic and long-form-sample item imputation rates for the 2000 and 1990 censuses and more limited information on item nonresponse in the Census 2000 Supplementary Survey (C2SS).[1] (An analysis commissioned by the Census Bureau used these tabulations; see Schneider, 2003.) The panel also compared the consistency of basic characteristics for people in the census-based E-sample who matched cases in the independent P-sample of the 2000 Accuracy and Coverage Evaluation (A.C.E.) Program.[2] A Content Reinterview Survey, conducted primarily by telephone in June–November 2000 of 20,000 long-form recipient households, provided indexes of inconsistency between census and survey responses for most questionnaire items for one randomly chosen member of each household. Unlike previous censuses, the 2000 Content Reinterview Survey did not try to measure systematic response biases by including probing questions to determine the most accurate response (see Singer and Ennis, 2003). A set of questionnaire experiments in 2000 examined forms design, listing of household members, and race and ethnicity questions (see Martin et al., 2003). At a later date, information on response variance and bias will be available from an exact match of long-form census records and the April 2000 Current Population Survey.

In this chapter, we briefly discuss the usefulness of three types of available 2000 census data quality measures: imputation rates, consistency measures, and variability and sample loss for the long-form sample (7–A). We then review available data quality measures

[1]The C2SS surveyed 700,000 households, or 1.8 million people, by mail with computer-assisted telephone and in-person follow-up; it is a precursor to the planned American Community Survey (see Appendix I.3).

[2]The 2000 P-sample surveyed 0.3 million households, or 0.7 million people, in about 11,000 block clusters, using computer-assisted telephone and in-person interviewing; the E-sample contained about the same number of households and people as the P-sample, drawn from the 2000 census records in the same block clusters (see Chapters 5 and 6).

for three population groups: all household members (7–B); household members in the long-form sample (7–C); and group quarters residents in the long-form sample (7–D). Each part concludes with a summary of findings and recommendations for 2010. Appendixes G and H describe the imputation and other data processing procedures that affect the basic and long-form-sample items, respectively. Appendix F reviews alternative item imputation procedures that may be more accurate than the current "hot-deck" procedures.

7–A AVAILABLE QUALITY MEASURES

7–A.1 Imputation Rates

The census enumeration will always have nonresponse: some households may not want to be found or are overlooked; respondents for participating households may not answer every question because they do not want to answer a particular item or do not know the answer; some responses are unintelligible and are voided in data processing. Since 1960, data processing for the complete-count census has used computer-based imputation for whole-household and item nonresponse (see Appendix C.5.d).[3] The long-form sample, like other household surveys, has used imputation for missing items; it accounts for household nonresponse by weighting respondent cases. Imputation makes census data more useful because analysts do not have to discard cases with missing values.[4] Imputation by the Census Bureau is also more efficient and facilitates consistency in uses of the data than if each analyst were to develop his or her own imputation procedure.

However, imputation is a source of error. Because imputation commonly uses reported values, the distribution of values after imputation will be inaccurate to the extent that cases requiring imputation differ from cases for which there are responses, in ways that are not or cannot be made part of the imputation procedure. Furthermore, the relationships of two or more variables may be distorted

[3]By "complete-count census," we mean the 100 percent enumeration, including short forms and the basic-item responses on long forms.

[4]One long-form-sample item first asked in 1980—namely, ancestry—is not imputed; instead, a "not reported" category is tabulated, as was common practice for census items prior to 1960.

if imputation levels are high and imputation techniques do not take account of these relationships. Consequently, except when cases with reported and missing values are similar in their characteristics or auxiliary information is available with which to improve the accuracy of the imputations, higher missing data rates will indicate poorer data quality.

For the 2000 census and the C2SS, codes on the data records distinguish imputations based on the use of another person's or household's information ("allocations") from "assignments" based on known information for the specific record. For example, first name was used to assign values to a large fraction of the records with missing sex;[5] answers to the race question were used to assign values to some cases with missing Hispanic origin; and answers to questions on housing costs were used to assign values to a large fraction of long-form cases missing housing tenure. Codes on the 1990 census records did not distinguish between imputations and assignments, so our tables for the 2000 census sometimes show both types of rates—the imputation/assignment rate is comparable to 1990 and indicates nonresponse;[6] the imputation rate per se indicates the fraction of cases that required a donor record to supply missing values.

7–A.2 Consistency Measures

Comparing the consistency of responses to the same question in two or more data sources can help identify possible reporting biases, although it is often not possible to say which source is more accurate. Consistency measures can also indicate response variability if responses tend to differ according to such factors as data collection mode, question format, and who answered for the household.

7–A.3 Sample Loss and Variability (Long Form)

Estimates from the long-form sample, like other surveys, are subject to variability from sampling and also from unit nonresponse

[5]This type of assignment was not possible in the 1990 census, which did not capture names except for records in the PES E-sample.

[6]The imputation/assignment rate is not exactly the same as a nonresponse rate because reported values that are inconsistent with other reported values may be blanked and another value imputed or assigned.

that further reduces the effective sample size. The long-form records for respondents are weighted to agree with complete-count totals (see Appendix H.2). This weighting effectively adjusts for sampling rates, instances of whole-household nonresponse, and additional sample loss due to some households having provided only minimal data. The 1990 and 2000 long-form sampling probabilities varied, by design, from 1 in 2 households to 1 in 8 households, depending on the population size and type of geographic area. To receive a nonzero weight, households had to include at least one member who reported at least two basic items and two long-form items. The measures of variability, or variance, that are constructed for the long-form sample take account of the weighting (see Appendix H.2), but not of the variability introduced by item imputation.

7–B QUALITY OF BASIC DEMOGRAPHIC CHARACTERISTICS

The basic demographic characteristics in 2000, asked on both the short and long forms, were age; month, day, and year of birth; sex; ethnicity (Hispanic origin); race; relationship to household reference person (first person listed on the questionnaire); and housing tenure (own or rent). The 1990 census included additional basic items (see Appendix B).

7–B.1 Imputation Rates for Complete-Count Basic Items

Table 7.1 provides imputation rates for the basic demographic items from the 2000 and 1990 census complete counts (separately for short and long forms) for total household members, people in households headed by blacks, and people in households headed by Hispanics.[7] These rates include cases of whole-household imputation in addition to individual item imputation.

In 2000 the combined whole-household and item imputation rates for all household members ranged from 2.3 percent for sex to 5.4 percent for ethnicity (Hispanic origin). The rates for long-form recipients were higher, ranging from 3 percent for sex to 9.3 percent for housing tenure. The rates for members of minority households

[7]Complete-count data were made available in the P.L. 94-171 file for redistricting and in Summary Files 1 and 2 (see Box 2.1).

Table 7.1 Basic Item Imputation Rates, 2000 and 1990 Complete-Count Census, by Type of Form and Race/Ethnicity, Household Population

Population Group and Form Type	Age[a]	Sex	Race	Ethnicity (Hispanic Origin)	Relation-ship to Head	Tenure
2000 Census						
Total Household Population	4.8	2.3	5.1	5.4	3.4	4.8
Short-Form Recipients	4.6	2.1	5.1	5.3	3.2	3.9
Long-Form Recipients	6.0	3.0	5.5	6.1	4.4	9.3
Black Non-Hispanic	7.6	3.9	4.9	10.5	5.6	7.3
Short-Form Recipients	7.1	3.5	4.6	10.3	5.2	6.0
Long-Form Recipients	10.2	5.8	7.0	11.6	7.9	14.5
Hispanic	7.5	4.2	17.2	6.5	6.3	5.9
Short-Form Recipients	7.4	4.1	17.4	6.4	6.0	5.2
Long-Form Recipients	8.4	4.9	15.9	7.6	7.5	10.3
1990 Census						
Total Household Population	3.1	1.9	2.6	10.5	3.3	3.1
Short-Form Recipients	3.0	1.9	2.8	11.7	3.4	3.1
Long-Form Recipients	3.5	1.7	2.1	4.1	2.9	2.7
Black Non-Hispanic	5.6	3.6	3.5	18.4	6.1	5.4
Short-Form Recipients	5.4	3.6	3.5	20.3	6.2	5.6
Long-Form Recipients	6.6	3.1	3.1	7.8	5.5	4.2
Hispanic	3.9	2.6	9.3	7.8	5.9	4.3
Short-Form Recipients	3.8	2.7	10.0	8.3	6.1	4.5
Long-Form Recipients	4.6	2.4	5.5	4.8	5.0	3.1

NOTES: Rates include whole-household imputations (types 2–5; see Box 4.2); wholly imputed persons (type 1); and item imputations. Race and ethnicity population groups are defined by the response of the household reference person.

Household population totals for 2000: 273.6 million total (83.4 percent on short forms, 16.6 percent on long forms); 32.4 million people in households headed by non-Hispanic blacks (84.8 percent on short forms, 15.2 percent on long forms); 33.4 million people in households headed by Hispanics (85.4 percent on short forms, 14.6 percent on long forms).

Household population totals for 1990: 242 million total (83.1 percent on short forms, 16.9 percent on long forms; 28.0 million people in households headed by non-Hispanic blacks (84.8 percent on short forms, 15.2 percent on long forms); 21.2 million people in households headed by Hispanics (85.4 percent on short forms, 14.6 percent on long forms).

[a] Excludes imputation of age from date of birth.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 and 1990 Household Census Edited Files (HCEF), provided to the panel spring 2003.

were also higher. Of particular note are the response patterns to the ethnicity and race questions: while 5 percent of total household members did not respond to these items, 11 percent of members of households headed by blacks did not respond to the ethnicity question, and 17 percent of members of households headed by Hispanics did not respond to the race question, perhaps believing it did not apply. For the 10 percent of census tract neighborhoods with the highest percentages of basic item imputations, the imputation rates were one-half to three times higher than the total population rates shown (data not shown).

In 1990 the combined whole-household and item imputation rates for household members were generally below the corresponding 2000 rates (Table 7.1). A reason for generally lower basic item imputation rates in 1990 compared with 2000 was that the 1990 census used telephone and field follow-up for missing or inconsistent data content, but the 2000 census did not. The exception was the ethnicity item on the short form, for which imputation rates in 1990 were twice as high as the 2000 rates (except for Hispanics). A reason for higher short-form imputation rates for ethnicity in 1990 than in 2000 was that the content review and follow-up procedures for mailed-back short forms in 1990 were trimmed back for budgetary reasons, so that only a one-tenth sample of short forms were reviewed and sent to follow-up if necessary (see Appendix C.3.c). Reordering the race and ethnicity items so that ethnicity came before race in 2000 (and not after as in 1990) also contributed to lower item imputation rates for ethnicity in 2000 compared with 1990.

In 2000 about 1.3 percentage points of the total household population imputation rates shown in Table 7.1 were due to whole-household imputations (types 2–5—see Box 4.2). For these cases data from neighboring households were used to supply complete records of basic information for members of households for which information on members' characteristics, and sometimes household size, was missing. Whole-household imputations contributed 1.2 percentage points to the imputation rates for short-form records and 1.8 percentage points to the imputation rates for long-form records. Whole-household imputation rates were highest for enumerator long forms (5.2 percent), followed by enumerator short forms (4.7 percent), with self-response forms by mail, telephone, Internet, or the Be Counted program including very few whole-household imputations (less than 0.1 percent).[8] In 1990 whole-household imputations contributed 0.7 percentage point to the total household population imputation rates shown.

In 2000 about 0.9 percentage point of the total household population imputation rates shown in Table 7.1 was due to wholly imputed persons (type 1—see Box 4.2 and Table 4.1), which occurred when there was not room to report basic characteristics for all household members on the questionnaire. The corresponding figure for 1990 was 0.2 percentage point. Imputations for missing members of enumerated households were made item by item, using information about the other household members to construct a reasonable household composition (e.g., imputing race and ethnicity to be consistent with other household members).

Imputation rates for some basic items would have been higher in 2000 than shown in Table 7.1 if the 2000 imputation procedures had not been able to take advantage of names and other information for assigning rather than imputing missing values (see Section 7–A.1). For example, first names were used to assign sex for about 1 percent of household members (data not shown); if this procedure had not been feasible, the imputation rate for sex for total household members in 2000 would have been 3.3 percent, not 2.3 percent as in Table 7.1.

Basic item imputation rates from the complete count for the nation as a whole and for large population groups were reasonably low for the most part, but some small geographic areas and population groups required much more imputation, which users should consider in their analyses. For example, imputation rates for race at the county level reached as high 17 percent, while imputation rates for ethnicity at the county level reached as high as 35 percent (see Section 8–C.2; see also Appendix H.3.b).

7–B.2 Missing Data Patterns for Basic Items

An analysis by Zajac (2003) provides information on patterns of missing responses in 2000 census records; that is, percentages of person records that are missing one, two, or three or more of the

[8]From tabulations by U.S. Census Bureau staff of the 2000 and 1990 Household Census Edited Files, provided to the panel in spring 2003.

five basic person items. We also computed these statistics for the 2000 census records in the A.C.E. E-sample and for the records in the independent P-sample. Table 7.2 provides these data from all three sources; the E-sample percentages exclude whole-household and whole-person imputations, as well as reinstated cases from the special Master Address File (MAF) unduplication operation, which could not be matched to P-sample cases.

More P-sample persons answered all five items—95 percent—than did census persons—87 percent (the corresponding figure for the E-sample—data not shown—is 89 percent). This result is not surprising because the interviewing for the P-sample was more carefully controlled than was the census enumeration. However, the census and the P-sample were much closer in the percentage of respondents who answered at least four items (97.6 percent P-sample, 96.1 percent 2000 census). The most commonly omitted basic items in the census were age and ethnicity (data not shown).

Rates of answering all five basic person items in the census varied by whether the household responded for itself or answered to an enumerator (Table 7.2, panel A). Members of self-responding households (mail, Internet, telephone, Be Counted) were more likely to answer all five items (90 percent) than were members of households visited by enumerators (79 percent). By the race/ethnicity and housing composition of the A.C.E. block cluster (Table 7.2, panel B), household members living in white and some other race owner and renter block clusters were most likely to answer all five questions (92 and 89 percent, respectively); household members living in Hispanic renter block clusters were least likely to answer all five questions (77 percent). These data are from the A.C.E. E-sample and underestimate the extent of nonresponse in the census; in contrast, the P-sample achieved a high level of reporting of all five basic person items for all neighborhood types—92 to 95 percent.

7–B.3 Consistency of Responses to Basic Items

Comparing census cases in the E-sample that matched P-sample cases revealed low rates of inconsistent reporting of basic items for the household population as a whole. Thus, 4.7 percent of matched cases (unweighted) had conflicting values for housing tenure; 5.1

Table 7.2 Percentage of Household Members Reporting Basic Items, 2000 Census, 2000 A.C.E. E-Sample and Independent P-Sample (weighted)

PANEL A	Number of Items Reported			
Sample	All Five	Four	Three	Two or Less
Census Total[a]	87.0	9.1	1.2	2.8
Self Enumerations	89.8	7.9	1.1	1.1
Interviewer Enumerations	79.4	11.9	1.6	7.1
P-Sample Total	94.9	2.7	0.8	1.6

PANEL B Household Members Reporting All Five Items by Neighborhood Type[b]	E-Sample[c]	P-Sample
American Indian and Alaska Native		
Owner	91.7	94.9
Renter	85.4	93.2
Hispanic		
Owner	80.6	94.5
Renter	77.2	94.1
Black		
Owner	85.0	94.5
Renter	81.7	93.8
Native Hawaiian and Other Pacific Islander		
Owner	84.7	94.0
Renter	85.4	92.2
Asian		
Owner	86.7	93.9
Renter	80.4	93.9
White and Other		
Owner	91.6	95.4
Renter	88.9	94.2

[a] Census percentages reporting two or fewer items include people in wholly imputed households (imputation types 2–5, see Box 4.2) and wholly imputed people (type 1); census percentages for interviewer enumerations include all people in households that were contacted in the coverage edit and follow-up operation of mail returns to obtain basic characteristics for missing household members.

[b] Neighborhood (A.C.E. block cluster) type determined by Census Bureau staff from 1990 characteristics (A.C.E. block clusters were defined as one or more contiguous blocks, intended to contain about 30 housing units on average—see Appendix E.1.a).

[c] The E-sample excludes whole-household and whole-person imputations and reinstated records due to the special summer 2000 unduplication operation (see Section 4–E).

SOURCE: For 2000 census, Zajac (2003:Table 34), adjusted to include people in wholly-imputed households; for E-sample and P-sample, tabulations by panel staff of P-Sample and E-Sample Dual-System Estimation Output Files (U.S. Census Bureau, 2001b), provided to the panel February 16, 2001, weighted using TESFINWT.

percent had conflicting values for age and sex group; and 3.9 percent had conflicting values for race/ethnicity domain (Farber, 2001a:Table 1). Reasons for conflicting values could include reporting error in one or both samples, differences in question format and mode and time of collection, different household respondents for the census enumeration and P-sample interview, different imputation methods, errors in imputation, and errors in matching.

Rates of inconsistency were higher for matched cases for which the characteristic in question was imputed in one or both samples than for nonimputed cases: thus, 12 percent of cases with imputed race or ethnicity, 22 percent of cases with imputed housing tenure, and 36 percent of cases with imputed age or sex were inconsistent between the E-sample and P-sample, compared with 3, 4, and 3 percent, respectively, of nonimputed cases (Farber, 2001a:Table 1). This result indicates that, at an individual level, imputations were often not accurate, which could be consequential for analyses of public use microdata samples and for small geographic areas. However, because overall imputation rates were low, the effects of imputation error were not large for the household population as a whole. Overall distributions by age and sex group, housing tenure, and race/ethnicity domain remained very similar in both the E-sample and P-sample (Farber, 2001a:Tables A-1, A-2, A-3).

When population groups were defined by multiple characteristics, instead of a single variable, then very high rates of inconsistency often occurred, particularly for imputed cases (see Farber, 2001a:Tables E-1 through E-64). As a best-case example, people who were classified in one or both samples as non-Hispanic white owners in medium-sized mailout/mailback areas with high response rates in the Midwest were classified inconsistently in another poststratum 10 percent of the time overall, 6 percent for nonimputed cases, and 41 percent for imputed cases, which accounted for only 11 percent of this group. As a worst-case example, people who were classified in one or both samples as American Indians and Alaska Natives off reservations were classified inconsistently in another poststratum 59 and 57 percent of the time overall for owners and renters, respectively; 55 and 52 percent for nonimputed owner and renter cases, respectively; and 74 and 77 percent for imputed owner and renter cases, respectively. Imputations accounted for 21 percent of each of these two groups (owners and renters).

7–B.4 Basic Item Imputation and Inconsistency Rates: Summary of Findings

In general the 2000 census complete count exhibited reasonably good data quality as measured by low imputation rates and low rates of inconsistency for individual items (less than 5 percent). However, there was considerable variability in the completeness of reporting (answering all or most items) by type of neighborhood (in contrast to the P-sample, which uniformly obtained a high level of complete reporting). There was also evidence of reporting problems for particular items, such as high rates of nonresponse to the race item by members of Hispanic-headed households and high rates of nonresponse to the ethnicity item by members of black non-Hispanic-headed households. In general, as we discuss in Chapter 8, responses to race and ethnicity items are sensitive to minor variations in question placement, format, and wording.

Comparisons of matched E-sample and P-sample cases indicated that rates of inconsistency were much higher for imputed cases than for nonimputed cases. Rates of inconsistency were also much higher for some groups than others. However, aggregate distributions were very much the same. Not answerable by any of the available evidence is whether there were systematic reporting errors for any of the basic items or whether imputation introduced bias, although further analysis of matched E-sample and P-sample cases could be helpful in this regard.

The census had higher missing data rates for basic items in 2000 than in 1990. Not known is whether the field follow-up to reduce missing data rates in 1990 obtained less accurate responses than the heavier reliance on computer-based imputation in 2000. We repeat that the Census Bureau's 2010 testing program should include tests of the trade-offs in costs and accuracy between computer imputation and additional field work for missing data (see Recommendation 4.2).

> **Finding 7.1:** Rates of missing data in 2000 were low at the national level for the basic demographic items asked of everyone (complete-count items)—age, sex, race, ethnicity, household relationship, and housing tenure. Missing data rates for these items ranged from 2 to 5 percent (including records for people with one or more missing

items and people who were wholly imputed). Rates of inconsistent reporting for the basic items (as measured by comparing responses for census enumerations and matching households in the independent Accuracy and Coverage Evaluation survey) were also low. However, some population groups and geographic areas exhibited high rates of missing data and inconsistent reporting for one or more of the basic items. No assessments have yet been made of reporting errors for such items as age, nor of the effects of imputation on the distributions of basic characteristics or the relationships among them.

7–C QUALITY OF BASIC AND ADDITIONAL LONG-FORM DATA

Data from the census long-form sample are eagerly awaited and widely used throughout government, academia, and the private sector. They are made available in Summary Files 3 and 4 and in public use microdata samples, which contain individual records for a sample of long-form cases, processed to protect individual confidentiality (see Box 2.1). For 2010 the current design calls for a short-form-only census, with long-form-type data collected in the new, continuously fielded American Community Survey (see Section 3–D).

7–C.1 Imputation Rates for Basic Items in the Long-Form Sample

Table 7.3 shows household member weighted imputation rates for the basic demographic items from the 2000 census long-form sample and three other surveys: (1) the Census 2000 Supplementary Survey; (2) the 2000 P-sample survey; and (3) the 1990 census long-form sample. Note that the 2000 and 1990 census long-form-sample rates are lower than the complete-count rates shown in Table 7.1. The reason is that whole-household imputations and members of nonsample-data-defined households were not given a weight in the long-form sample; instead, weighting was used to adjust for their nonresponse. (Sample-data-defined households in 2000 and 1990 were households with at least one member who reported at least 2 basic items and 2 sample items.) Consequently, a full understanding of the completeness of response in the long-form sample requires examining not only item imputation rates but also weighting adjustments (see Section 7–C.5).

Generally, imputation rates for basic items were lowest for the P-sample and for the 1990 census long-form sample, whether on self returns or enumerator returns. Imputation rates, even excluding assignments, were highest (although still below 5 percent, except for housing tenure) for the 2000 census long-form sample; again, there was not much difference between self and enumerator returns. Housing tenure had a much higher rate of nonresponse in the 2000 long-form sample (8 percent) than in the 1990 long-form sample (1 percent), which may have been due to the placement of the question (see Section 7–C.2). Imputation rates in the 2000 long-form sample for basic items by race and ethnicity and by type of geographic area (e.g., central city, rural) showed relatively little variation; imputation rates for census tracts with the highest percentage of basic item imputations were about twice the overall rates (see Appendix H).

7–C.2 Imputation Rates for Additional Long-Form Items

Table 7.4 compares nonresponse rates for the 2000 and 1990 long-form samples for 15 person items and 12 housing items that were asked solely of the long-form sample, by type of return (self, enumerator).[9] The items are listed in order of their appearance on the 2000 long-form questionnaire. With the exception of the items for English-speaking ability and year structure built, the 2000 long-form sample exhibited higher nonresponse rates than occurred in 1990. For some items, the differences were dramatic: for example, the 2000 and 1990 nonresponse rates were 16 and 9 percent, respectively, for occupation and 16 and 1 percent, respectively, for monthly rent (which was a complete-count item in 1990). For income, 30 percent of long-form cases in 2000 failed to report some or all income items, compared with only 13 percent of long-form cases in 1990.[10] The much higher nonresponse rates for most housing items in 2000

[9]See Appendix H (Tables H.6 and H.7) for comparable tabulations for all person and housing long-form items.

[10]Denominators of imputation rates for additional long-form items use the appropriate universe for the question.

Table 7.3 Basic Item Imputation Rates, 2000 and 1990 Census Long-Form Sample, Census 2000 Supplementary Survey, and 2000 P-Sample, by Type of Rate and Form, Household Population (weighted)

Population Group and Form Type	Age[a]	Sex	Race	Ethnicity (Hispanic Origin)	Relation-ship to Head	Housing Tenure
2000 Census Long-Form Sample						
Imputations/Assignments	2.6	1.6	3.2	4.0	2.7	8.0
Self Responses	1.9	1.5	3.7	4.7	2.6	7.6
Enumerator Responses	4.3	2.0	2.0	2.4	3.0	8.9
Imputations Only	2.6	0.9	3.2	3.6	2.3	4.3
Census 2000 Supplementary Survey						
Imputations Only	2.4	0.5	2.4	3.6	1.6	1.4
Self Responses	3.6	0.7	3.4	5.7	0.7	1.6
Enumerator Responses	0.8	0.2	1.1	0.7	1.2	1.0
2000 Accuracy and Coverage Evaluation P-Sample						
Imputations[b]	2.3	1.6	1.2	2.2	1.7[c]	1.9
1990 Census Long-Form Sample						
Imputations[b]	0.9	0.8	1.1	3.4	1.9	1.4
Self Responses	0.8	0.9	1.0	4.0	1.7	1.4
Enumerator Responses	1.2	0.7	1.6	1.7	2.4	1.4

NOTES: Imputation/assignment rates (percents) are included for comparability with 1990 imputation rates; imputation/assignment rates include item imputations plus assignments of values for missing or inconsistent responses based on known information for the specific record. —; not available. Census self responses include mail, telephone, Internet (2000 only), and Be Counted returns; enumerator responses include forms obtained in nonresponse follow-up, list/enumerate, and other field operations.

[a] Excludes imputation of age from date of birth.
[b] Includes "assignments."
[c] Relationship was not imputed in the P-sample; the rate reported is the missing data rate.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 and 1990 Sample Census Edited Files (SCEF), and the C2SS file, provided to the panel spring 2003. P-sample tabulations by panel staff from P-Sample Dual-System Estimation Output File (U.S. Census Bureau, 2001b), provided to the panel February 16, 2001, weighted using TESFINWT.

compared with 1990 could in part have resulted from the different placement of the housing items on the two long forms: in 2000, the housing items followed all of the basic and additional items for the first person; in 1990, the housing items followed all of the basic items for all household members and were followed in turn by the additional items for all household members.

By type of return, nonresponse rates for most long-form items were higher for enumerator returns than for self-returns.[11] For some items, the differences were minor, but for others they were large. For example, in 1990, 21 percent of people on enumerator returns failed to answer some or all of the income items, compared with 11 percent of people on self returns; in 2000, 40 percent of people on enumerator returns failed to answer some or all of the income items, compared with 26 percent of people on self returns—a disturbingly high percentage in itself. By race and ethnicity (defined for the household reference person; see Appendix H, Table H.2), 40 percent of non-Hispanic blacks and 34 percent of Hispanics failed to answer some or all of the income questions in 2000, compared with 25 percent of Asians and 28 percent of non-Hispanic whites.

Table 7.5 shows imputation rates (excluding assignments) for the same 15 person items and 12 housing items for the 2000 census long-form sample and the Census 2000 Supplementary Survey, by type of return. For every item shown except year structure built, imputation rates were higher for the 2000 long-form sample compared with the C2SS. The differences in favor of the C2SS ranged from 0.4 percentage points for marital status to 11.2 percentage points for property taxes. For the 2000 long-form sample, for most person items, imputation rates for enumerator returns exceeded those for self returns, sometimes by large margins. In contrast, for the C2SS, most person-item imputation rates were lower for enumerator returns than for self returns. The exception for the person items shown was income, for which enumerator returns had higher imputation rates than self returns for both the 2000 long-form sample and the C2SS, although

[11]Percentages of self returns in the 2000 and 1990 long-form samples (76 and 73 percent, respectively) are not the same as final mail return rates, which were 71 percent in both years for the long form. One reason is that proportionately more enumerator returns than mail returns were dropped from the long-form sample because they were not sample-data-defined, so their nonresponse was handled by weighting instead of imputation (see Section 7–C.5).

Table 7.4 Imputation/Assignment Rates for Selected Long-Form Items, 2000 and 1990 Census Long-Form Samples, by Type of Response, Household Population (weighted)

	2000 Long Form			1990 Long Form		
Item	Total	Self	Enum-erator	Total	Self	Enum-erator
Marital Status	3.4	2.3	6.2	0.9	0.8	0.9
Educational Attainment	7.2	5.2	12.0	4.5	3.8	6.1
English-Speaking Ability	7.6	7.3	7.9	8.5	8.3	9.0
Place of Birth	9.2	7.8	12.5	5.1	4.3	7.1
Residence 5 Years Ago	8.6	7.4	11.6	5.2	3.6	9.3
Mobility Disability	10.0	10.5	8.6	5.1	4.7	6.4
Work Disability[a]	11.4	12.2	9.3	7.4	7.4	7.6
Veteran Status	8.2	6.8	11.8	4.8	4.0	7.0
Employment Status Recode	11.1	10.2	13.4	3.8	3.0	6.2
Place of Work (State)	9.7	7.3	15.5	7.2	6.2	9.7
Transportation to Work	8.2	6.0	13.3	4.6	3.7	7.3
Occupation Last Year	16.1	14.3	20.4	9.1	7.9	12.5
Weeks Worked Last Year	20.2	19.7	21.3	14.7	14.6	15.0
Wage and Salary Income	20.0	15.0	32.6	10.0	7.7	16.3
Income, All Sources						
100 Percent Imputed	24.5	18.9	38.5	11.7	9.1	19.1
Some Imputed[b]	29.7	25.5	40.3	13.4	10.9	20.5
Units in Structure[c]	4.4	4.9	3.0	1.6	1.8	1.2
Year Structure Built	11.7	9.3	18.0	23.0	20.2	30.7
Number of Rooms[c]	6.2	6.2	6.4	0.4	0.4	0.5
Complete Plumbing	3.4	3.5	3.1	1.7	1.7	1.8
Complete Kitchen	3.4	3.5	3.1	1.8	1.8	1.8
Fuel Used for Heating	7.4	6.3	10.1	2.9	2.7	3.4
Annual Electric Cost	18.5	15.3	26.9	5.5	4.4	8.5
Monthly Rent[c]	15.6	13.2	19.2	1.3	1.1	1.6
Property Taxes	32.0	27.0	49.6	12.2	10.3	19.4
Value of Property[c]	13.3	12.3	16.6	3.3	3.3	3.4

NOTES: 2000 rates (percents) are imputation/assignment rates for comparability, as a measure of nonresponse, with 1990 imputation rates (see notes to Table 7.3).

Self responses (76 percent of 2000 long-form sample; 73 percent of 1990 long-form sample) included mail, telephone, Internet (2000 only), and Be Counted returns; enumerator responses included forms obtained in nonresponse follow-up, list/enumerate, and other field operations.

—; not available.

[a] In 1990, "work disability" refers to a disability that prevents working; in 2000, the term refers to a disability that makes it difficult to work.

[b] Includes 100 percent of income imputed.

[c] Basic (complete-count) item in 1990.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 and 1990 Sample Census Edited Files (SCEF), provided to the panel spring 2003.

the difference was smaller in the C2SS. For most housing items, imputation rates for enumerator returns were higher than the rates for self returns for both surveys.

At least three factors argue why the C2SS could be expected to achieve lower imputation rates than the long-form sample and, in particular, why the C2SS interviewers could be expected to outperform the census enumerators. First, the C2SS interviewers were more experienced and much better trained than the temporary census staff, which meant that they were better able to obtain responses from reluctant respondents. Second, the C2SS interviews used computer-assisted telephone and in-person interviewing. (By design the C2SS used sampling for households that did not mail back a return and did not respond during the initial telephone-follow-up phase; the personal interview sampling rate was 1 in 3 of nonresponding households.) Third, the goal for the C2SS was to collect all of the information. In contrast, the essential goal for the 2000 census was to obtain a complete count—if household respondents balked at answering the additional long-form-only questions, the enumerators were not as likely to press hard for a response as would the C2SS interviewers. Yet these factors were apparently not sufficient to overcome respondents' reluctance or inability to answer questions on income and housing costs in the C2SS as well as the census.

7–C.3 Missing Data Patterns for Additional Items

No analysis has been conducted to date of patterns of response and nonresponse to the additional long-form items; that is, whether people tended to omit single items or clusters of items or most items. The panel made a limited set of tabulations of nonresponse patterns of cases in the A.C.E. E-sample who fell into the 2000 long-form sample and whose records were augmented at the panel's request by the additional long-form items. These tables focused on the income and employment items, which exhibited high nonresponse rates.

A tabulation of people age 16 and over in the E-sample long-form records found that 71 percent answered all 9 income items, 11 percent answered 6–8 items, 2 percent answered only 3–5 items, 8 percent answered only 1 item, and 8 percent did not answer any of

Table 7.5 Imputation Rates for Selected Long-Form Items, 2000 Long-Form Sample and Census 2000 Supplemental Survey, by Type of Response, Household Population (weighted)

	2000 Long Form			Census 2000 Supplementary Survey		
Item	Total	Self	Enum-erator	Total	Self	Enum-erator
Marital Status	2.2	1.4	4.3	1.8	2.4	1.0
Educational Attainment	7.2	5.2	12.0	4.8	4.9	4.7
English-Speaking Ability	7.6	7.3	7.9	6.0	10.5	2.3
Place of Birth	9.2	7.8	12.5	6.4	8.1	4.1
Residence 5 Years Ago	5.8	4.3	9.6	4.0	5.6	1.8
Physical Disability	7.6	7.1	8.9	5.2	7.4	2.1
Work Disability	11.4	12.2	9.3	5.9	8.3	2.2
Veteran Status	7.5	6.1	11.0	4.7	6.1	2.5
Employment Status Recode	11.1	10.2	13.4	6.0	8.2	2.6
Place of Work - State	9.7	7.3	15.5	5.8	6.5	4.8
Transportation to Work	7.6	5.4	13.0	4.6	5.5	3.3
Occupation Last Year	14.9	13.2	19.2	9.5	11.1	7.1
Weeks Worked[a]	19.3	18.6	20.9	9.6	11.1	7.3
Wage and Salary Income	20.0	15.0	32.6	16.4	13.0	21.4
Income, All Sources[a]						
100 Percent Imputed	24.5	18.9	38.5	20.0	16.1	25.7
Some Imputed[b]	29.7	25.5	40.3	23.9	20.7	28.6
Units in Structure	4.4	4.9	3.0	1.4	1.6	1.0
Year Structure Built	11.7	9.3	18.0	13.4	7.4	22.8
Number of Rooms	6.2	6.2	6.4	2.6	3.4	1.4
Complete Plumbing	3.4	3.5	3.1	1.0	1.4	0.3
Complete Kitchen	3.4	3.5	3.1	0.9	1.3	0.3
Fuel Used for Heating	7.4	6.3	10.1	2.1	1.6	2.8
Electric Cost[c]	17.1	13.6	26.1	6.9	4.3	11.0
Monthly Rent	15.6	13.2	19.2	5.3	4.2	6.3
Property Taxes	32.0	27.0	49.6	20.8	13.7	35.4
Value of Property	13.3	12.3	16.6	9.7	6.0	17.4

NOTES: Rates (percents) exclude assignments. In 2000, self responses included mail, telephone, Internet, and Be Counted returns; enumerator responses included forms obtained in nonresponse follow-up, list/enumerate, and other field operations. In the C2SS, self responses included mail; enumerator responses included forms obtained in telephone and in-person follow-up.

[a] For 1999 in the 2000 census long-form sample; for last 12 months in the C2SS.
[b] Includes 100 percent of income imputed.
[c] Annual cost in the 2000 census long-form sample; last month's cost in the C2SS.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 Sample Census Edited File (SCEF) and the Census 2000 Supplementary Survey edited file, provided to the panel spring 2003.

the 9 items.[12] Of those cases that did not respond to the entire bank of income questions, 99 percent also failed to answer the questions on whether the person worked last week or last year; in contrast, 97 percent of complete responders to the income questions, and even 91 percent of those who omitted some of the income questions, did answer both the worked last week and worked last year questions. These few tabulations suggest that the distribution of respondent behavior could be bimodal, at least for the most intrusive and difficult questions on income and employment. A majority responded to all or most items, a smaller but still sizeable group responded to very few, and the smallest group was in between these two extremes.

7–C.4 Consistency of Responses to Long-Form Items

The 2000 Content Reinterview Survey provides an index of inconsistency for most of the long-form questionnaire items, comparing responses to the census with responses to the Content Reinterview Survey for 20,000 recipients of the 2000 long form. The index that was computed for each item is the ratio of simple response variance to total variance (times 100). Historically, index values below 20 have been considered to be low, values between 20 and 50 have been considered to be moderately problematic, and values above 50 have been considered to be very problematic (Singer and Ennis, 2003:7–9). The Content Reinterview Survey computations included only individuals for whom the items were reported in both the census and the survey, excluding imputed and edited values.

Table 7.6 provides values of the index of inconsistency from the 2000 Content Reinterview Survey and its 1990 counterpart for 18 person items and 11 housing items. Schneider (2003:Tables 1, 2) provides index values for additional 2000 long-form-sample items for which there were no corresponding 1990 values. The items exhibit a wide range of inconsistency, ranging from 3.2 percent in 2000 and 4.9 percent in 1990 for place of birth to 80.5 percent in 2000 for work disability and 73.6 percent in 1990 for self-care disability. Overall, of the long-form-sample items shown in Table 7.6, eight of the 2000 items and eight of the 1990 items have low values of

[12]The E-sample did not include the worst filled-out long-form cases, so the percentage of E-sample cases that did not answer any of the income questions was lower than the corresponding long-form-sample percentage.

the index of inconsistency (below 20), indicating that the data are measured reliably in the census (although whether the reporting is unbiased is not known). Thirteen of the 2000 items and fourteen of the 1990 items have index values between 20 and 50, which makes their reliability of moderate concern. The remaining eight of the 2000 items and seven of the 1990 items have high values of the index of inconsistency (50 or above), indicating that the data are not measured reliably but, rather, are subject to considerable reporting variance for what could be a number of reasons, such as different respondents, different data collection modes, the difficulty of answering the question, or poor question design. Income amounts in 2000 also tend to have high index values (data not shown). Some of the items with high values are measuring rare occurrences (e.g., few households lack plumbing facilities and few households have mobile home costs); for such items, discrepancies in reporting between the census and the Content Reinterview Survey tend to inflate the value of the index more than would be true for items that are more evenly distributed.

A number of items shown in Table 7.6 were less consistently reported in 2000 than in 1990, while others were more consistently reported. Users should be aware of these changes, as they may affect the reliability of trend analysis. Specifically, nine items had index values that were 5 percentage points or higher than the corresponding values in 1990. These items were Hispanic origin (5 percent increase), race (7 percent increase, perhaps due to the new option to indicate more than one race), mobility disability (17 percent increase), work disability (35 percent increase), veteran status (10 percent increase), complete plumbing facilities (31 percent increase), vehicles available (5 percent increase), business on property (16 percent increase), and agricultural sales (10 percent increase). Conversely, eight items had index values in 2000 that were lower than the corresponding 1990 values by 5 or more percentage points. These items were self-care disability (22 percent decrease), length of military service (17 percent decrease), work last year (22 percent decrease), usual hours worked (6 percent decrease), year structure built (11 percent decrease), lot size (7 percent decrease), monthly rent (12 percent decrease), and meals in rent (33 percent decrease).

Table 7.6 Index of Inconsistency for Selected Long-Form-Sample Items, 2000 and 1990 Content Reinterview Surveys (weighted)

	Content Reinterview Survey	
Item	2000 Census	1990 Census
Hispanic Origin[a]	17.2	12.2
Race[a]	23.1	16.3
School Enrollment	13.5	17.3
Educational Attainment	36.5	32.3
Ancestry	30.7	26.5
Speaks Non-English Language	22.7	26.9
English-Speaking Ability	59.5	60.3
Place of Birth	3.2	4.9
Citizenship	9.8	10.9
Year of Entry	18.9	23.0
Self-Care Disability	51.7	73.6
Mobility Disability	64.5	47.1
Work Disability	80.5	45.7
Veteran Status	18.7	8.5
Length of Service	41.6	58.8
Work Last Year	24.3	45.9
Weeks Worked	57.5	56.8
Usual Hours Worked	34.3	40.1
Housing Tenure[a]	19.4	13.3
Units in Structure[b]	20.8	21.9
Year Structure Built	29.3	40.6
Complete Plumbing Facilities	85.2	53.8
Heating Fuel	17.7	14.0
Vehicles Available	37.1	32.1
Business on Property[b]	65.8	50.0
Lot Size	20.9	27.8
Agricultural Sales	52.0	41.7
Monthly Rent[b]	23.2	34.7
Meals in Rent	38.2	71.6

NOTE: The index of inconsistency is the ratio of the simple response variance to the total variance for an item times 100; for items with three or more categories, it is the aggregate index for the whole item; it ranges from 0 to 100 (see Singer and Ennis, 2003:7–9).

[a] Basic (complete-count) item in 2000 and 1990.
[b] Basic (complete-count) item in 1990.

SOURCE: Schneider (2004:Tables 1, 2).

7–C.5 Weighting and Variance Estimation

Because the long form is a sample survey, there is another option for treating missing data in addition to imputation—namely, adjustments can be made to the survey weights of respondents so that they represent nonrespondents as well. Such weighting is a replacement for whole-household imputation, which represented a considerably higher percentage of long-form-sample households than of the complete count in 2000 and 1990. The reason is that only households in which at least one member reported at least two basic items and two additional items were given a nonzero weight in the final edited long-form sample.

Table 7.7 provides a measure of household sample loss in the 2000 and 1990 long-form surveys. The 2000 long-form sample experienced somewhat less sample loss than the 1990 long-form sample: 99 percent of expected long forms were received, and 93 percent of households received were retained in the sample (i.e., given nonzero weights). The corresponding figures for the 1990 long-form sample were 98 percent of expected long forms received and 91 percent of received household long forms retained in the sample. In the 10 percent of census tracts with the highest rates of basic item imputations, only 84 and 79 percent of received household long forms were retained in the 2000 long-form sample and 1990 long-form sample, respectively.

Of the 1.2 million long-form-sample household returns in 2000 that were not sample defined and were dropped from the 2000 Sample Census Edited File, 54 percent were proxy returns for which a neighbor or landlord provided information for the household. Proxy returns in 2000 accounted for 6.2 percent of long forms received (they were 19.4 percent of enumerator long forms received). Because thee-fifths of long-form proxy returns were not sample-data-defined, the percentage of proxy returns in the 2000 Sample Census Edited File was only 2.7 percent.[13]

A reason for the somewhat higher rates of sample retention in 2000 than in 1990 could have been the layout of the questionnaire. As we described above, the 2000 long form led off with all of the

[13]From tabulations by U.S. Census Bureau staff from the 2000 Sample Census Edited File provided to the panel in spring 2003. Proxy returns were 4.8 percent of 2000 short forms (20.9 percent of enumerator short forms).

Table 7.7 Whole-Household Nonresponse in the 2000 and 1990 Census Long-Form Samples

Measure	2000 Long-Form Sample: Total Households	2000 Long-Form Sample: Households in Worst 10% of Tracts	1990 Long-Form Sample: Total Households	1990 Long-Form Sample: Households in Worst 10% of Tracts
Percent, Long Forms Received of Number Expected	98.5	96.6	97.8	94.0
Percent, Households Retained in Edited Long-Form Sample of Number of Forms Received	93.2	84.3	91.2	78.8
No. Long Forms Expected from Households (millions)	17.9	1.3	15.9	1.1

NOTES: Households not retained in the edited long-form sample include wholly imputed households from the complete-count processing (types 2–5; see Box 4.2) and households in which no person had at least two basic and two long-form items reported (i.e., they were not sample data-defined). Worst 10 percent census tracts were defined as those with the highest rates of basic item imputations.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 and 1990 Sample Census Edited Files (SCEF), provided to the panel spring 2003.

person items for the first person, whereas the 1990 long form asked all of the basic items for each household member, followed by the housing items, followed by the sample person items. It was easier in 2000 to meet the criterion for being "sample-data-defined," so long as the first person answered the basic items and, say, marital status and education level, which came first among the additional person items.

Table 7.8 provides another measure of loss for the 2000 long-form sample; it shows the percentages of non-sample-data-defined persons by race and ethnicity of the household reference person and type of return. For the total long-form population, the rates of non-sample-data-defined persons range from 1.6 percent of white self returns to 25.6 percent of black enumerator returns. In the 10

Table 7.8 Whole-Person Nonresponse in the 2000 Long-Form Sample, by Race of Reference Person

			Non-Hispanic	
Measure	Total Persons	Hispanic	Black	White
Percent Non-Sample-Data-Defined Persons of:				
Total Persons on Long Forms	8.4	11.0	14.2	7.1
Self Returns	2.1	4.7	3.7	1.6
Enumerator Returns	20.9	18.5	25.6	20.7
Persons on Long Forms in Worst 10% Census Tracts	18.5	15.4	23.6	17.3
Self Returns	4.8	6.8	4.8	2.5
Enumerator Returns	31.5	24.1	39.3	33.9
No. Persons on Long Forms Received (millions)				
Total	45.4	4.9	4.9	33.1
Worst 10% Census Tracts	3.7	1.3	1.0	1.1

NOTES: Non-sample-data-defined persons include persons in wholly imputed households and other non-sample-data-defined households (see text), plus wholly imputed persons (type 1) in enumerated households. Wholly imputed persons did receive a sample weight. Worst 10 percent census tracts were defined as those with the highest rates of basic item imputations.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 and 1990 Sample Census Edited Files (SCEF), provided to the panel spring 2003.

percent of census tracts with the highest percentage of basic item imputations, the rates range from 2.5 percent of white self-returns to 39.3 percent of black enumerator returns.

For some groups and geographic areas, the levels of sample loss are large, adding to the variability, and, hence, uncertainty, of long-form-sample estimates. Moreover, because the Census Bureau provides variability estimates that do not account for item imputation and, separately, provides item imputation rates based on the people who were retained in the sample, it is easy to overlook the fact that weighting and item imputation are both methods of dealing with sample loss. Ideally, the two kinds of sample loss should be considered together. For example, if income imputation rates are combined with sample loss due to households with no or minimal response, then it could result that the effective sample size for a poverty estimate for the total household population in 2000 would be only 60 percent of the original sample size (8 percent sample loss of persons from Table 7.8 plus 30 percent nonresponse to one or more income items from Table 7.5).

7–C.6 Long-Form-Sample Data Quality: Summary of Findings and Recommendations

The additional long-form-sample items exhibited higher item imputation rates in 2000 than the basic items. More serious, the long-form-sample data quality, as measured by nonresponse, deteriorated in 2000 compared with 1990. While sample loss was not quite as high as in 1990, imputation rates for many items were considerably higher, reaching levels as great as 32 percent for property taxes and 30 percent for some or all income items (compared with 12 and 13 percent, respectively, for these two items in 1990). A major reason for the disparities in rates was the effort devoted to telephone and field follow-up for missing data items in 1990; such effort was almost nonexistent in 2000.

When sample loss is considered together with item imputation, the variability in the 2000 long-form-sample estimates could be much more than expected from the original sample selection probabilities. However, the Census Bureau's variance estimates for the long-form sample took account of sample selection rates and whole-household sample loss, but not item imputation. What we do not know is the extent of bias in the 2000 long-form-sample estimates that might be attributable to the high rates of imputation.

With regard to measures of response variance or consistency of reporting for the long-form items, the 2000 Content Reinterview Survey showed a wide range of values for an index of inconsistency, as did a similar survey in 1990. A total of 8 of the 2000 items and 7 of the 1990 items had index values greater than 50, indicating that the data were not reliably measured; another 13 items in 2000 and 14 items in 1990 had index values between 20 and 50, indicating that the data were only moderately reliable.

Little information is available with regard to possible response bias for the long-form-sample items (e.g., systematic overreporting or underreporting of income). Because the 2000 Content Reinterview Survey did not ask probing questions to try to obtain an accurate response, it did not provide measures of response bias. Comparisons of long-form-sample estimates with other sources have been performed for only a few variables thus far. Aggregate comparisons with the 2000 April Current Population Survey (CPS) and the C2SS found relatively consistent estimates of the percent poor in 1999: the estimated poverty rates were 12.4 percent for the long-form sample, 11.9 percent for the CPS, and 12.2 percent for the C2SS (Schneider, 2004:16). However, aggregate comparisons with the 2000 April CPS found sizeable discrepancies in estimates of employed and unemployed people (Clark et al., 2003). Thus, the census estimate of the number of employed people was 5 percent lower than the 2000 April CPS estimate, while the census estimate of the unemployment rate was 2.1 percentage points higher than the CPS rate, representing a 50 percent larger number of unemployed people (7.9 million in the census compared with 5.2 million in the CPS). The differences were much more pronounced for blacks and Hispanics than for other groups and much larger than those found in similar comparisons for 1990. They appear to be due in part to changes in question wording and imputation procedures from those used in 1990.

> **Finding 7.2:** For the household population, missing data rates were at least moderately high (10 percent or more) for over one-half of the 2000 census long-form-sample items and very high (20 percent or more) for one-sixth of the long-form-sample items. Missing data rates also varied widely among population groups and geographic areas. By comparison with 1990, missing data rates were higher in 2000 for most long-form-sample items asked in both years and substantially higher—by 5 or more percentage points—for one-half of the items asked in both years. In addition, close to 10 percent of long-form-sample households in 2000 (similar to 1990) provided too little information for inclusion in the sample data file. When dropped households and individually missing data are considered together, the effective sample size that is available for analysis for some characteristics is 60 percent or less of the original long-form-sample size.
>
> Many long-form-sample items had moderate to high rates of inconsistent reporting, as measured in a content reinterview survey. Few assessments have yet been

made of systematic reporting errors for the long-form-sample items, although aggregate comparisons of employment data between the 2000 census and the Current Population Survey (CPS) found sizeable discrepancies in estimates of employed and unemployed people—much larger than the discrepancies found in similar comparisons for 1990. No analysis of the effects of item imputation and weighting on the distributions of characteristics or the relationships among them has yet been undertaken, although analysis determined that changes in imputation procedures contributed to the 50 percent higher unemployment rate estimate in the 2000 census compared with the April 2000 CPS.

Recommendation 7.1: **Given the high rates of imputation for many 2000 long-form-sample items, the Census Bureau should develop procedures to quantify and report the variability of the 2000 long-form estimates due to imputation, in addition to the variability due to sampling and weighting adjustments for whole-household weight adjustments. The Bureau should also study the effects of imputation on the distributions of characteristics and the relationships among them and conduct research on improved imputation methods for use in the American Community Survey (or the 2010 census if it includes a long-form sample).**

Recommendation 7.2: **The Census Bureau should make users aware of the high missing data rates and measures of inconsistent reporting for many long-form sample items, and inform users of the 2000 census long-form-sample data products (Summary Files 3 and 4 and the Public Use Microdata Samples) about the need for caution in analyzing and interpreting those data.**

In particular, users should review Census Bureau documentation of imputation and weighting procedures, examine imputation rates and estimates of standard errors provided by the Bureau, be alert for User Notes from the Bureau about data errors and other reports on data quality, and inform Census Bureau staff of data anomalies for investigation.

7–D QUALITY OF GROUP QUARTERS DATA

Residents of group quarters accounted for 7.8 million people in the 2000 census, up from 6.7 million in 1990. The census is the only source at present of detailed information for residents of all types of group quarters, including prisons, juvenile institutions, nursing homes, hospitals and schools for the handicapped, military quarters, shelters, group homes, and other group quarters.

The quality of the data for group quarters residents in the 2000 census long-form sample was poor in comparison with the data for household residents and also in comparison with the group quarters data in 1990. Table 7.9 shows 2000 imputation/assignment rates and comparable 1990 imputation rates for selected person data items for the total group quarters population, prison inmates, and students in college dormitories.[14] In 2000 missing data rates for the items shown reached as high as 50 percent for all group quarters residents and as high as 75 percent for prison inmates. Generally, missing data rates were highest for inmates of institutions (prisons, juvenile institutions, nursing homes, hospitals and schools for the handicapped) and lowest for college students and the military. The particularly high rates for institutional residents were probably due in part to the high rates of use of administrative records to provide information instead of enumeration of residents. The Census Bureau was not prepared for the widespread resort to administrative records by institutions (see Section 4–F.2), and, very often, the available records did not contain long-form-type information, or institutions were unwilling to provide such information.

The great extent of missing data among group quarters residents in 2000 raises a question as to whether the Census Bureau should have published long-form-sample estimates for some or all types of group quarters. Given the decision to publish, it was unfortunate that many tabulations in census data products (e.g., age, employment, income) combined group quarters and household residents.

[14]See Appendix H (Table H.8) for rates for all of the basic and additional items for group quarters residents in nine categories of type of facility.

Table 7.9 Imputation/Assignment Rates for Selected Person Items, 2000 and 1990 Census Long-Form Samples, by Type of Residence, Group Quarters Population (weighted)

	2000			1990		
	Total Group Quarters	Inmates of Prisons	Students in College Dormitories	Total Group Quarters	Inmates of Prisons	Students in College Dormitories
Age[a]	3.8	5.5	3.4	1.5	2.1	1.3
Sex	3.0	2.7	1.9	0.6	1.1	0.2
Race	4.5	5.4	5.4	1.8	2.7	1.4
Ethnicity	8.0	11.8	7.1	7.6	16.8	3.4
Marital Status	18.0	30.9	8.1	4.2	11.1	1.2
Educational Attainment	39.3	53.8	19.2	17.9	24.6	2.8
English-Speaking Ability	33.9	56.8	16.3	22.1	29.8	11.0
Place of Birth	40.2	54.0	22.2	19.2	31.7	6.7
Citizenship	36.5	53.0	19.9	14.0	24.7	3.9
Residence 5 Years Ago	44.9	70.6	23.7	18.1	33.5	4.7
Mobility Disability	46.9	66.2	22.3	16.7	31.5	6.3
Work Disability[b]	47.7	66.7	22.7	18.1	34.2	6.0
Grandchildren at Home	30.0	36.5	0.5	—	—	—
Veteran Status	39.6	57.5	21.6	18.0	29.2	5.7
Occupation Last Year	46.9	75.4	30.7	21.3	44.2	11.1
Weeks Worked Last Year	42.8	72.5	29.1	21.4	40.2	13.2
Wages Last Year	50.1	74.3	34.7	27.4	49.7	15.0
Population (millions)	7.78	1.98	2.06	6.66	—	—

NOTES: —; not available.

[a] Excludes imputation of age from date of birth.

[b] In 1990, "work disability" refers to a disability that prevents working; in 2000, the term refers to a disability that makes it difficult to work.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 and 1990 Sample Census Edited Files (SCEF), provided to the panel spring 2003.

Combining the data makes it harder for users to compare census results for such statistics as the poverty rate and unemployment rate with other household surveys, which typically do not include the military or institutional residents. Combining the data also obscures the differences in data quality between group quarters residents and household members. The difference is further obscured because published item imputation (allocation) rates that accompany the long-form-sample data products combined group quarters and household member rates.

No systematic investigation has yet been undertaken of the effects on distributions of characteristics of the high rates of missing data and the imputation procedures used. However, the discovery by users of very high unemployment rates in some communities, such as college towns, led to a determination by the Census Bureau that a particular combination of missing responses to some questions and reports of availability for work by residents of group quarters resulted in an inappropriate imputation of unemployed status to many such residents. The problem affected residents of noninstitutional group quarters, such as students in college dormitories, people living in group homes, and others. Residents of institutions showed similar reporting patterns, but the imputation did not allow an unemployed status for them. The problem is described in User Note 4 for Summary Tape File 3 (U.S. Census Bureau, 2003d:Data Note 4),[15] and the magnitude is such that the Census Bureau reissued employment status tabulations for states, counties, and places in late 2003 to exclude group quarters residents, limiting the tabulations to household residents only.[16]

Finding 7.3: For group quarters residents, missing data rates for most long-form-sample items were very high in 2000 (20 percent or more for four-fifths of the items and 40 percent or more for one-half of the items). The 2000 rates were much higher than missing data rates for household members and considerably higher than missing data rates for group quarters residents in 1990. The

[15] The note is also available at http://www.census.gov/prod/cen2000/doc/sf3.pdf [2/25/04].

[16] Available at http://www.census.gov/population/www/census2000/phc-28.html [2/25/04].

2000 missing data rates were particularly high for prisoners, residents of nursing homes, and residents of long-term-care hospitals perhaps because of heavy reliance on administrative records for enumerating them. Few assessments have yet been made of systematic reporting errors for group quarters residents for long-form-sample items, nor of the effects of imputations on the distributions of characteristics or the relationships among them. However, a systematic error was found in the imputation of employment status for people living in noninstitutional group quarters because of a particular pattern of missing data. The result was a substantial overestimate of unemployment rates for these people, so much so that the Census Bureau reissued employment status tabulations for household members only, excluding group quarters residents.

Earlier we called for a complete redesign of the enumeration procedures for group quarters residents in the 2010 census (see Recommendation 4.4 in Section 4–F.3). That redesign should include consideration of changes to questionnaire content as well. If the American Community Survey is fully funded as a replacement for the census long-form sample, then it is likely to and should provide detailed data for group quarters residents. (At present, the C2SS and other precursors to the ACS include only household residents.) The designers for the ACS should consider how best to obtain long-form-type data for different types of group quarters. For institutions, the use of administrative records may make most sense provided the cooperation of the facility staff can be obtained. It may also be that accurate responses to some of the long-form-sample questions (e.g., income and employment last year) are too difficult to obtain in institutional settings, either from records or the residents themselves, at least without special training of interviewers and other measures to elicit responses. For other types of group quarters, a household-type questionnaire may work best. This small, but important, population merits dedication of sufficient resources for research and testing on questionnaire design and content and enumeration procedures that can produce useful, high-quality information for policy making, program planning, and other purposes.

Recommendation 7.3: **The Census Bureau should publish distributions of characteristics and item imputation rates, for the 2010 census and the American Community Survey (when it includes group quarters residents), that distinguish household residents from the group quarters population (at least the institutionalized component). Such separation would make it easier for data users to compare census and ACS estimates with household surveys and would facilitate comparative assessments of data quality for these two populations by the Census Bureau and others.**

CHAPTER 8

Race and Ethnicity Measurement

MEASUREMENT OF RACE HAS BEEN A CRITICAL PART of census-taking from the first census in 1790; measurement of ethnicity and national origin in one or another form goes back to the 1850 census. Today, census data on race and ethnicity (Hispanic origin) are used to assess the compliance of legislative redistricting plans with the Voting Rights Act and to enforce equal opportunity employment laws, among other applications. These data also facilitate research on such diverse topics as income, disability, and migration flows.

Classification by race has become an increasingly complex enterprise as the composition of the U.S. population has changed and the notions of race and ethnicity have shifted in meaning. Racial classification has been used differently throughout U.S. history, at times to discriminate and at other times to measure and strive to end discrimination. Hence, even basic definitions and approaches to measurement are controversial. Restricting racial identity to a single racial category has become increasingly difficult as the population with parents from different racial groups increases in size; changing immigration patterns have also posed challenges for classifying—and even self-identifying—by race.

In the past, federal standards for collection of racial and ethnic data limited the flexibility of federal statistical and program agencies to collect data reflecting the changing racial and ethnic composition of the U.S. population. The 2000 census marked a radical departure from past practices in the collection of racial data. Specifically, the 2000 census is the first major implementation of revised guidelines for the collection of racial and ethnic data issued by the U.S. Office of Management and Budget (OMB). These guidelines instruct federal agencies collecting data on race and ethnicity to allow respondents to "mark one or more races to indicate what this person considers himself/herself to be," thus relaxing the past practice of limiting racial classification of individuals in the United States to one and only one of a few mutually exclusive racial categories.[1] The full implications of this change are yet to be understood.

This chapter provides additional background on the manner in which race and ethnicity data were collected in the 2000 census. After presenting a brief history of the collection of race and ethnicity data (8–A), we discuss the standardization of the federal collection of race and ethnicity data and changes therein, as embodied in the OMB guidelines (8–B). We discuss differences between the 1990 and 2000 census questions on race and Hispanic origin, the implications of those differences, and the results (8–C). After examining the quality of the race and ethnicity data collected in 2000, comparisons are made, when possible, with the quality of those data from the 1990 census. Finally, we discuss the implications of the findings for the collection of racial and ethnic data in future censuses (8–D).

8–A HISTORICAL OVERVIEW

Since its inception in 1790, every United States census has included questions on race. The federal collection of data on race and ethnicity has its origins in the balance of power between the North and the South when the United States was first established. The original text of Article 1, Section 2, of the Constitution tied the allocation of representatives in Congress to population counts; the

[1]The few instances prior to 2000 of including multiracial categories in census responses all appeared before the shift from enumerator-identification to self-identification of race in 1970. "Mulatto" was included in 1850–1890, 1910, and 1920, "quadroon" and "octoroon" in 1890, and "part Hawaiian" in 1960.

weight applied to each person differed by civil status. Specifically, for the purpose of reapportionment, each free person was to count as a whole person, each slave was to count as three-fifths of a person, and Indians not taxed were to be excluded. The first census in 1790 required enumerators to classify people as white (identifying separately white males above and below the age of 16 and white females), free blacks, and slaves. In context, because there were no enslaved whites, this classification by civil status effectively replicated what would have been the classification by race of whites and blacks (Indians not taxed were not counted). The decennial census has continued to collect racial data—even after the differential civil status articulated in Article 1, Section 2, of the Constitution was stricken by the Fourteenth Amendment—with the specific question and categories shifting every 10 years.

Over the past 200 years there has been tremendous variation in the collection of race and ethnicity data in the census; Table 8.1 lists the racial categories in the census by year from 1850 to 2000.[2] In fact, the race and ethnicity items have rarely remained the same from one census to the next. As new populations have entered the United States, the number of census items related to race, ethnicity, and ancestry has fluctuated. Items on ethnicity and national origin were first included in the census of 1850, as part of an effort to distinguish recent European immigrants from the native-born white population. These items related to nativity, parental birthplace, immigration, citizenship status, language status, and time of arrival in the United States. Although ethnicity questions initially reflected a desire to distinguish among different groups of immigrants with European origins, the saliency of those groups shifted as they were assimilated into an increasingly heterogeneous white racial group.

In order to account for non-European immigrants, beginning with the 1910 census, the race question has included an "other" category allowing the enumerator (later a household member) to write in a race not otherwise on the list. Subsequent to passage of the 1965 Immigration and Nationality Act (which abolished the national origins system of the National Origins Act of 1924, as well as the Asiatic barred zone of the Immigration Act of 1917), additional

[2]Prior to the 1850 census, de facto racial categories were indicated by civil status: free whites, all other free persons, and slaves.

Table 8.1 Census Race Categories, 1850–2000

Year	White	Black	Native Peoples	Asian	Hawaiian and Other Pacific Islanders	Other
1850	—	Black, mulatto	—	—	—	—
1860	—	Black, mulatto	Indian	—	—	—
1870	White	Black, mulatto	Indian	Chinese	—	—
1880	White	Black, mulatto	Indian	Chinese	—	—
1890	White	Black, mulatto, quadroon, octoroon	Indian	Chinese, Japanese	—	—
1900	White	Black	Indian	Chinese, Japanese	—	—
1910	White	Black, mulatto	Indian	Chinese, Japanese	—	Other (+ write in)
1920	White	Black, mulatto	Indian	Chinese, Japanese, Filipino, Hindu, Korean	—	Other (+ write in)
1930	White	Negro	Indian	Chinese, Japanese, Filipino, Hindu, Korean	—	(Other races, spell out in full)
1940	White	Negro	Indian	Chinese, Japanese, Filipino, Hindu, Korean	—	(Other races, spell out in full)
1950	White	Negro	American Indian	Chinese, Japanese, Filipino	—	(Other race—spell out)
1960	White	Negro	American Indian	Chinese, Japanese, Filipino	Hawaiian, Part Hawaiian	—
1970	White	Negro or black	Indian (American)	Chinese, Japanese, Filipino, Korean	Hawaiian	Other (print race)
1980	White	Black or negro	Indian (American), Eskimo, Aleut	Chinese, Japanese, Filipino, Korean, Vietnamese, Asian Indian	Hawaiian, Guamanian, Samoan	Other (specify)
1990	White	Black or negro	Indian (American), Eskimo, Aleut	Chinese, Japanese, Filipino, Korean, Vietnamese, Asian Indian	Hawaiian, Guamanian, Samoan, Other Asian or Pacific Islander	Other race (Print race)
2000	White	Black, African American, or Negro	American Indian or Alaska Native	Chinese, Japanese, Filipino, Korean, Vietnamese, Asian Indian, Other Asian	Native Hawaiian, Guamanian or Chamorro, Samoan, Other Pacific Islander	Some other race (Print race)

SOURCE: National Research Council (1995b:Table 7.1).

Asian categories were included to account for changing immigration flows from that region. Racial categories, once introduced, have not remained sacrosanct: Mexican was used only once, in the 1930 census; none of the mixed black-white designations (i.e., mulatto, quadroon, octoroon) have reappeared since the 1920 census; and part Hawaiian made its sole appearance in the 1960 census.

As the ethnic groups from Europe were assimilated as whites, other ethnic differentiation became increasingly salient. The Hispanic origin question was first introduced in the 1970 census at the behest of Daniel Patrick Moynihan, then a domestic policy adviser in the Nixon White House (Farley, 2001). Because production of the census questionnaires had already begun, the question was included on only 5 percent of the 1970 census forms. Later, in 1974, Congress passed Public Law 94-311, which formally required the collection and publication of Hispanic-origin data.

The introduction of the Hispanic-origin question roughly coincided with another fundamental change in the collection of race and ethnicity data in the U.S. census. Prior to the 1960 census, the enumerator determined the racial classification of each person. The transition in the 1960 and 1970 censuses to census designs centered on mailed questionnaires necessitated a corresponding move from enumerator-reporting of race by observation to mainly self-reporting. In 1960, census short forms were delivered by the U.S. Postal Service to residences in areas of the country covering about 82 percent of the population. Recipients were asked to fill out the forms and be prepared to return them to the enumerators. The best estimate is that 60 percent of households had the forms filled out and waiting before the enumerator arrived. If the form was not filled out, or if a household had not received a form, or if a household resided in the remaining areas of the country covering about 18 percent of the population, then race was identified by the enumerator. The 1970 census was the first in which racial identity was meant to be entirely self-reported in mailback areas (to the extent that a member of the household is responsible for reporting his or her own race and the race of others in the household) instead of observer-reported. In practice, the extent of household reporting has varied with mailback response; in some cases of follow-up by census enumerators, nonrespondents may have their race reported by a neighbor or landlord or by enumerator observation. Also, the extent of mailback areas has varied in censuses from 1970 to 2000 (see Section 4–A.1 and Box 4.1).

More recent censuses have had opportunities for individuals to report variations in racial and ethnic identity not otherwise captured by the standard categories: the census allows and codes write-ins for the race, ethnicity, and ancestry items. On the short form in the 2000 census, an individual could write in a tribal affiliation if identifying as an American Indian or Alaska Native, an "other" Spanish/Hispanic/Latino group, a specific "other" Asian group, a specific "other" Pacific Islander group, or some other race (not listed elsewhere).

Although the categories and questions about race have shifted across the years, until 2000 there was one constant: each person was classified as a member of one, and only one, racial category. In the more than 200 years that the United States has collected data on race, the 2000 census is the first that has allowed people to identify themselves as belonging to more than a single racial category.

8–B STANDARDIZING FEDERAL COLLECTION: THE OMB GUIDELINES

Despite this long history of collecting data on race and ethnicity in the census, as well as in household surveys and program agency administrative records, there was no coordination among federal agencies collecting such data until the OMB released Statistical Policy Directive 15 in 1977.[3] This directive was meant to standardize the federal collection of data on race and ethnicity by dictating minimum standards for collection and reporting. As originally formulated, the directive required that data be collected and reported for a minimum of four major racial groups (white, black, Asian and Pacific Islander, and American Indian or Alaskan Native) and two ethnicities (Hispanic origin and non-Hispanic origin). More detailed categories were permitted, insofar as the categories could be aggregated into the specified minimum set of categories. Directive 15 formally codified what had been implicit in questions

[3]Technically, the Statistical Policy Directive—under that name—was released in 1978, following a reorganization of OMB functions. The material commonly known as Statistical Directive 15 was released on May 12, 1977, as Exhibit F in OMB Circular A-46.

on race throughout the history of the census: the underlying conception of race, as implemented by the U.S. government, was inherently categorical with mutually exclusive categories.

The distinction between race and ethnicity, however, was not immediately apparent from the directive's guidelines. Although a preference was expressed for a two-question format that would allow each racial category to be cross-classified by Hispanic origin, a single combined race and ethnicity question was permitted. This flexibility allowed for the collection of irreconcilable race and ethnicity data. Whereas the two-question format allowed each individual to select both race and ethnicity, the one-question format allowed an individual to indicate only one of five categories—each of the four races and the one ethnicity (Hispanic). In the first formulation, race and ethnicity are two distinct dimensions: Hispanics could be of any race. The second formulation required mutually exclusive identification with *either* a race *or* Hispanic ethnicity. In the second formulation, Hispanics did not belong to any racial group—Hispanics were treated as a distinct group, not allowing for overlapping ethnic identity with classification into a racial category. The racial and ethnic populations classified under these two schemes differ significantly. This blurring of the concepts of race and ethnicity is indicative of the lack of a coherent theoretical structure underlying U.S. attempts to measure race and ethnicity.

Beyond the inconsistencies present in Directive 15 from the outset, the directive also failed to anticipate the changing racial and ethnic landscape of the United States. The changes in race and ethnic data collected in each census prior to the implementation of Directive 15 are indicative of changes across time in both the population and the context of measurement. Hence, it is not surprising that in recent decades the federal standards for the measurement of race and ethnicity and the reality of racial and ethnic identities in the U.S. population became increasingly mismatched.

Directive 15 guided the collection of race and ethnicity data in both the 1980 and the 1990 censuses, as well as numerous other federal data collection efforts. By 1993 enough questions had been raised about the relevance of the 1977 standards that Representative Thomas Sawyer (D-Ohio), who headed the House subcommittee with jurisdiction over the census, arranged to discuss these issues at hearings about the 2000 census. One of the key issues of discussion at the hearings was the measurement and classification of multiracial individuals. How were the children of interracial marriages to identify themselves if the standards remained unchanged? Shortly thereafter, the OMB announced that the racial categories in Directive 15 were of decreasing value, and therefore it would undertake a complete review of the 1977 standards.

OMB established the Interagency Committee for the Review of the Racial and Ethnic Standards in March 1994 to review and evaluate the standards, consider changes, and make recommendations to OMB for revisions.[4] The decision to review the standards spurred the Bureau of Labor Statistics and the Census Bureau to field three surveys in order to examine alternative formulations of the race and ethnicity questions. In addition to including alternative race and ethnicity questions on the Census Bureau's 1996 National Content Survey, additional options were explored in the Bureau of Labor Statistics' 1995 Supplement on Race and Ethnicity to the Current Population Survey (CPS) and the Census Bureau's 1996 Race and Ethnicity Targeted Test (RAETT).

In the 1996 National Content Survey (NCS), the Census Bureau tested three changes to the race and ethnicity questions as presented in the 1990 census: (1) adding a multiracial response category, (2) placing the Hispanic-origin question immediately before the race question, and (3) combining both changes. A "mark one or more" option was not tested in the NCS. The CPS supplement also examined the effects of the inclusion of a single "multiracial" category in the list of racial categories. In addition, it tested the inclusion of Hispanic origin as a "race" category and usage of alternate names for some of the racial and ethnic categories. In their analysis of the CPS supplement, Tucker and Kojetin (1996:5) report that, when specifically asked about their preferences, "a substantial majority of Hispanics ... preferred to identify themselves as Hispanic in the race question, rather than in the separate ethnicity question." However, the combined question format led to significantly lower reports of Hispanic origin than the two-question format. Like the NCS, the CPS supplement did not examine the effect on reporting Hispanic-origin of combining the race and Hispanic origin question

[4]See National Research Council (1996) for the summary of a workshop convened by the Committee on National Statistics in conjunction with the OMB-initiated review of Directive 15.

when respondents are allowed to mark one or more categories.

Of the three tests of alternative race and ethnicity questions prior to the revision of Directive 15, the Census Bureau's Race and Ethnic Targeted Test was the only one to test the "mark one or more" format. The RAETT also tested a combined race and ethnicity question, allowing respondents to mark one or more categories, and found that there was no decline in reporting of Hispanic origin in comparison to the two-question format. Moreover, in comparison to the nonresponse rates to both the race and ethnicity questions, nonresponse was significantly reduced by use of the combined format (Hirschman et al., 2000). Along with public discussion, the results of these experiments with alternative formats for the race and ethnicity questions informed the deliberations and conclusions of the interagency committee.

Responding to recommendations announced by the interagency committee in July 1997, the OMB issued "Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity" on October 30, 1997. (The revised guidelines no longer carry the designation "Statistical Policy Directive 15," as they were not released as part of an OMB directive, circular, or bulletin.) The minimum number of racial categories for reporting was changed from four to five; the Asian or Pacific Islander category was divided into an Asian category and a Native Hawaiian or Other Pacific Islander category, while the other three categories (white, black, and American Indian) remained unchanged. No other ethnic groups were added to the minimum standards, although several labels were modified and the definition of American Indians expanded to include Central and South American Indians. OMB also reiterated its preference for race to be self-identified and for the retention of distinct race and Hispanic-origin questions, although a single, combined question would still be permitted. Finally, respondents were to be allowed to self-identify with more than one racial group. Tabulations are required for the total number of individuals identifying with more than one group, as well as any combinations permitted given data quality and confidentiality concerns. The 2000 census adopted the revised standards, with the exception that OMB allowed the census to continue to include an "other" category for those respondents who did not identify with any of the racial groups otherwise listed in the question. The revised standards were to be implemented in all federal data collection efforts by 2003.

8–C RACE AND ETHNICITY DATA IN 1990 AND 2000

8–C.1 Questions and Results

The 2000 census is the first major federal data collection effort to adopt and implement the revised standards. To reiterate, in accord with the revised OMB guidelines, there are three major differences between the 1990 and 2000 race and ethnicity questions. First, respondents were instructed to "mark one or more" instead of "fill one circle" for the race question. Second, the Asian and Pacific Islander category was split into two distinct categories. Third, the Hispanic origin question was moved in 2000 to immediately precede the race question, whereas in 1990 it followed the race question with other questions on age, year of birth, and marital status placed in between. This last change was meant to encourage higher response rates to the race questions among Hispanics. Figures 8.1 and 8.2 illustrate the relevant sections of the 1990 and 2000 census forms, respectively. Altogether, these changes render the 1990 and 2000 data on race and ethnicity less than strictly comparable.

In 1990, in a total population of 248,709,873 people, the census indicated that 80.3 percent of the U.S. population was white, 12.1 percent black, 0.8 percent American Indian, Eskimo or Aleut, 2.9 percent Asian or Pacific Islander, and 3.9 percent other race. Whereas in the 1990 census the U.S. population was classified entirely into five mutually exclusive racial categories, cross-classified by one ethnicity (Hispanic, non-Hispanic), the 2000 census does not allow us to replicate that classification scheme. Because individuals were allowed to mark more than one race, the racial categories are no longer mutually exclusive. In 2000, in a total population of 281,421,906 people, published estimates indicated that 2.4 percent of the population marked two or more races, while 97.6 percent marked a single racial category: 75.1 percent of the U.S. population reported white alone, 12.3 percent black or African American alone, 0.9 percent American Indian and Alaska Native alone, 3.6 percent Asian alone, 0.1 percent Native Hawaiian or Other Pacific Islander alone, and 5.5 percent some other race alone.

Correction of a recently discovered error in the processing of race on 2000 census forms obtained from enumerators resulted in a 14.7 percent decrease in the population reporting two or more races, from

2.4 percent to 2.1 percent of the total population.[5] Changes for single race groups were minimal: from 75.1 percent to 75.4 percent white alone, from 12.3 percent to 12.4 percent black alone, no change at 0.9 percent for American Indian and Alaska Native alone, no change at 0.1 percent for Native Hawaiian or Other Pacific Islander alone, and from 5.5 percent to 5.4 percent for some other race alone.

Based on the published estimates, two or more race respondents in 2000 accounted, as one would expect from the increase in multiracial marriages, for higher percentages of children than older people—over 8 percent of children ages 0–4 were reported as multiracial in 2000 compared with only 1 percent of people age 70 and older. The West had the highest percentage of multiracial people (4 percent), and the Midwest the lowest percentage (1.5 percent). Nearly one-third (32.6 percent) of the multiracial population identified themselves as Hispanics (2,224,082 of 6,826,228 individuals marking two or more races also indicated that they were of Hispanic origin).

The proportion of individuals reporting that they were of Hispanic origin increased by approximately one-third between 1990 and 2000. For individuals of all races, 9.0 percent reported Hispanic origin in 1990 and 12.5 percent reported Hispanic origin in 2000. In both 1990 and 2000, almost 97 percent of individuals reporting "other race" or "some other race" were Hispanics. In addition, 6.3 percent of Hispanics made up nearly one-third of those people checking more than one race in 2000 (most of whom checked "white" and "Some Other Race"). These results appear to be consistent with those from the 1996 RAETT (discussed above), which indicate that many Hispanics (almost 50 percent) think of their ethnicity as a race.

8–C.2 Quality of Race and Ethnicity Data in the 2000 Census

Self-identification of race, Hispanic origin, and ancestry questions means that responses are based on self-perception and therefore are subjective, but at the same time, by definition, whatever response is recorded is an accurate response. This means that the concept of accuracy is meaningless for self-reported race and ethnicity data. However, we do have two metrics by which we can

[5]The Census Bureau plans to issue a User Note about the error (Cresce, 2003).

4. Race
Fill ONE circle for the race that the person considers himself/herself to be.
If **Indian (Amer.)**, print the name of the enrolled or principal tribe. ⟶

- ☐ White
- ☐ Black or Negro
- ☐ Indian (Amer.) (Print the name of the enrolled or principal tribe.) [_ _ _ _ _ _ _ _]
- ☐ Eskimo
- ☐ Aleut

Asian or Pacific Islander (API)

☐ Chinese	☐ Japanese
☐ Filipino	☐ Asian Indian
☐ Hawaiian	☐ Samoan
☐ Korean	☐ Guamanian
☐ Vietnamese	☐ Other API

If **Other Asian or Pacific Islander (API)**, print one group, for example: Hmong, Fijian, Laotian, Thai, Tongan, Pakistani, Cambodian, and so on. ⟶ [_ _ _ _ _ _ _ _]

If **Other race**, print race. ⟶ ☐ Other race (Print race)

Questions 5 (Age and year of birth) and 6 (Marital Status)

7. Is this person of Spanish/Hispanic origin?
Fill ONE circle for each person.

- ☐ No (not Spanish/Hispanic)
- ☐ Yes, Mexican, Mexican-Am., Chicano
- ☐ Yes, Puerto Rican
- ☐ Yes, Cuban
- ☐ Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)

If **Yes, other Spanish/Hispanic**, print one group. ⟶ [_ _ _ _ _ _ _ _]

Figure 8.1 Race and Hispanic Origin Questions, 1990 Census

➜ **NOTE: Please answer BOTH Questions 5 and 6.**

5. Is this person Spanish/Hispanic/Latino? *Mark ☒ the* ***"No"*** *box if* ***not*** *Spanish/Hispanic/Latino.*

- ☐ **No,** not Spanish/Hispanic/Latino
- ☐ Yes, Puerto Rican
- ☐ Yes, Mexican, Mexican Am., Chicano
- ☐ Yes, Cuban
- ☐ Yes, other Spanish/Hispanic/Latino—*Print group*

6. What is this person's race? ***Mark ☒ one or more races*** *to indicate what this person considers himself/herself to be.*

- ☐ White
- ☐ Black, African Am., or Negro
- ☐ American Indian or Alaska Native—*Print name of enrolled or principal tribe*
- ☐ Asian Indian
- ☐ Japanese
- ☐ Native Hawaiian
- ☐ Chinese
- ☐ Korean
- ☐ Guamanian or Chamorro
- ☐ Filipino
- ☐ Vietnamese
- ☐ Samoan
- ☐ Other Asian—*Print race.*
- ☐ Other Pacific Islander—*Print race.*
- ☐ Some other race—*Print race.*

Figure 8.2 Race and Hispanic Origin Questions, 2000 Census assess the quality of these data. The first is by examining item imputation rates (i.e., the percentage of the population for whom data were missing and subsequently had values assigned or imputed). The second is by examining consistency in self-reports of race and ethnicity (i.e., the extent to which the responses to race and ethnicity questions were the same or different across multiple responses for the same individuals). Even when the respondent is the same, however, consistency of reporting of race and ethnicity may be affected by the context in which the reporting occurs (see below). When measured even over short time intervals, consistency of reporting may also be affected by changes in what is salient to the respondent. (See del Pinal, 2003, which summarizes several studies of race and ethnicity reporting in the 2000 census.)

Item Imputation Rates

To what extent are race and ethnicity data missing and therefore require imputation? If a respondent fails to provide answers to a question in the census, special procedures are used to impute (or allocate) the response. The Hispanic-origin item continues to have the highest imputation rate of any of the basic items collected from everyone. However, the imputation rate for the Hispanic-origin item was substantially lower for the 2000 census (5.4 percent for household members) than the 1990 census (10.5 percent; see Section 7–B.1). This significant improvement is undoubtedly due in no small part to placing the question to precede the race question in 2000 rather than to follow both the race question and several other questions (age, year of birth, and marital status) in 1990.

In contrast, the imputation rate for the race item approximately doubled between 1990 and 2000: it was 5.1 percent in 2000 compared with 2.6 percent in 1990. Although it is not yet clear what accounts for the increase in the imputation rate for the race item, some portion of this change is undoubtedly due to the changing ethnic composition of the population. Hispanics were more likely than non-Hispanics to be nonrespondents to race, and the share of the population that is Hispanic increased significantly between 1990 and 2000, from 9.0 to 12.5 percent.

Imputation Rates Across Geographic Areas

Imputation varied greatly across smaller geographic areas. Although some counties required no imputation for one or the other of the race and ethnicity items, there are others with imputation rates for these items far above the national average. For example, among counties in the 48 contiguous states, the imputation rate for race was as high as 16.7 percent in Costilla County, Colorado, and, similarly, the imputation rate for Hispanic origin reached 35 percent in Concho County, Texas. Figures 8.3 and 8.4 illustrate the variation in imputation rates for the Hispanic-origin and race items both across and within county population categories. County-level variation in imputation rates is not strictly related to population size.

To emphasize the extent to which these imputation rates vary across and within various geographical boundaries, we examined the tract-level imputation rates for three of the largest counties in the nation: Los Angeles County, California; Cook County, Illinois; and Harris County, Texas. These counties make up part of the Los Angeles, Chicago, and Houston Primary Metropolitan Statistical Areas, respectively. For each of these areas, plots of the mean imputation rates (tract-level) for Hispanic origin and race are shown in Figures 8.5 and 8.6. Inspection of the figures shows that there is significant variation of imputation rates among tracts (within counties) in each of the selected areas. This variation at lower levels of geographic aggregation indicates that small-area estimates for race and ethnic groups may be quite problematic.

Measures of Inconsistency

Consistency in the reporting of race is a complicated issue. Research demonstrates that context affects both self-reports and classification by others of an individual's race. Analysis of data from the National Longitudinal Study of Adolescent Health indicates that there is differential reporting of race by students when they are in school and when they are at home (Harris, 2003). In the same paper, Harris shows that there is significant variation by race and gender in observer identification of other people's race. As suggested by Waters (1990), self-identification for many people is based on context, and people have multiple identities that may differ by context. The

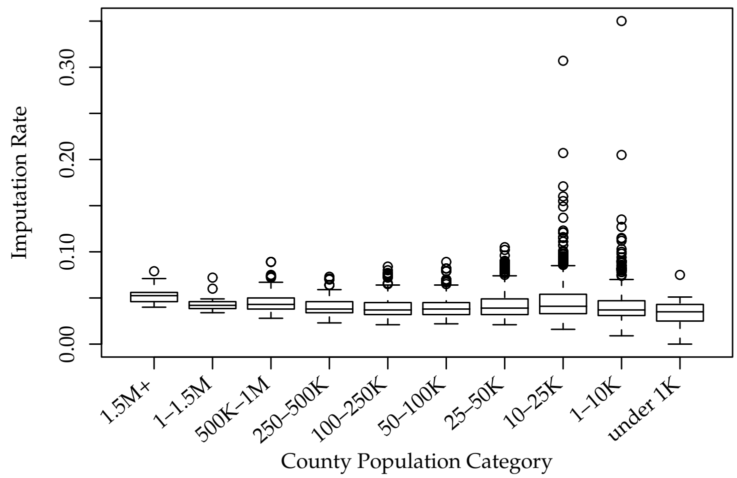

Figure 8.3 Imputation Rates for Hispanic Origin by Population (County Level)

NOTES: In this boxplot, each box is bounded by the first and third quartiles (25th and 75th percentiles) of the distribution of data values for the category. The center line of the box is the median value; the "whisker" lines extend from the ends of the box to the most extreme data points within 1.5 times the interquartile range. Points outside the whiskers may be considered outliers for the distribution. The data are for all counties in the 48 contiguous states, excluding zero population counties. Imputation rates are based on total population, including household members and group quarters residents. Population categories are in millions (M) or thousands (K).

SOURCE: Summary File 1, accessed from the Census Bureau's Web site, http://factfinder.census.gov.

same may also be true for observer-identification. The political context for the current census categories—whether there are positive, neutral, or negative effects to identifying with a certain group—is only one context for identification. The auspices and purposes of collecting race and ethnicity data provide another context for self-identification. For example, a person may respond differently to the same wording of a question on race when asked in the census and when asked in a job application form.

We can assess consistency by comparing reports from the census enumeration with those from the 2000 Accuracy and Coverage

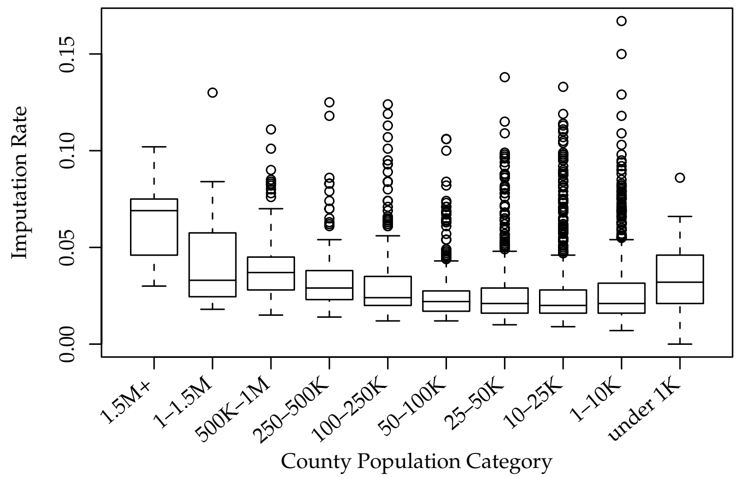

Figure 8.4 Imputation Rates for Race by Population (County Level)

NOTE: See Notes to Figure 8.3.

SOURCE: Summary File 1, accessed from the Census Bureau's Web site, http://factfinder.census.gov.

Evaluation (A.C.E.) survey (see Chapter 6). A sample survey was conducted in selected blocks (generating the P-sample), the data from which were then matched to the set of census enumerations from the sample blocks (the census E-sample). The A.C.E. collected race and ethnicity information using different data collection procedures and likely different respondents for people who matched to the census E-sample. There are therefore many factors that may have contributed to the inconsistency of reports.

Reports of ethnicity were substantially more consistent in the two samples than reports of race. Comparison of matched cases when ethnicity was not imputed in either sample found a high rate of consistent reporting overall—98.4 percent of cases were in agreement on Hispanic or not Hispanic origin between the two samples. Consistency of race reporting was not as high—only 91.5 percent of cases were in agreement between the two samples. Consistency

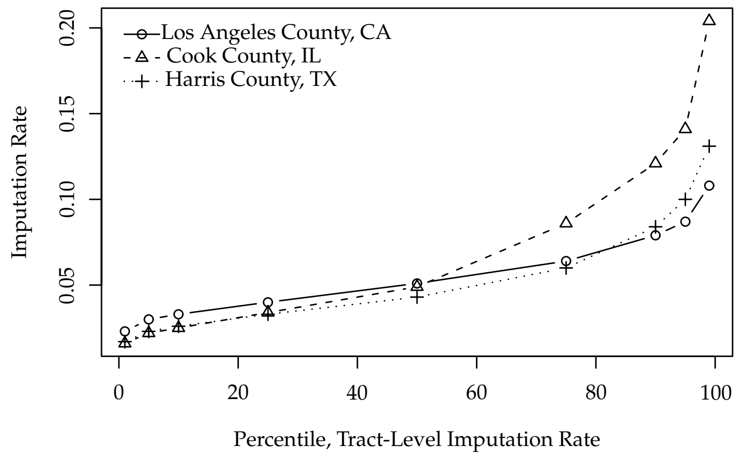

Figure 8.5 Imputation Rates for Hispanic Origin in Census Tracts of Selected Counties

NOTE: Points on each line are the cumulative percentage imputation rates for census tracts at the 1st, 5th, 10th, 25th, 50th, 75th, 90th, 95th, and 99th percentiles of the distribution of tract-level imputation rates within a county.

SOURCE: Summary File 1, accessed from the Census Bureau's Web site, http://factfinder.census.gov.

rates as percentages of the reported census (E-sample) race category were quite high for whites (94 percent), blacks (93 percent), and Asians (89 percent); considerably lower for Some Other Race (72 percent), Native Hawaiian and Other Pacific Islander (60 percent), and American Indian and Alaska Native (56 percent); and very low for Two or More Races, for which only 35 percent of the census E-sample responses agreed with the P-sample response. This significant inconsistency in multirace reporting indicates that membership in the multiracial population is highly unstable and therefore that the multiracial population identified on Census Day is only one of many multiracial populations that might have been enumerated (Harris, 2003).

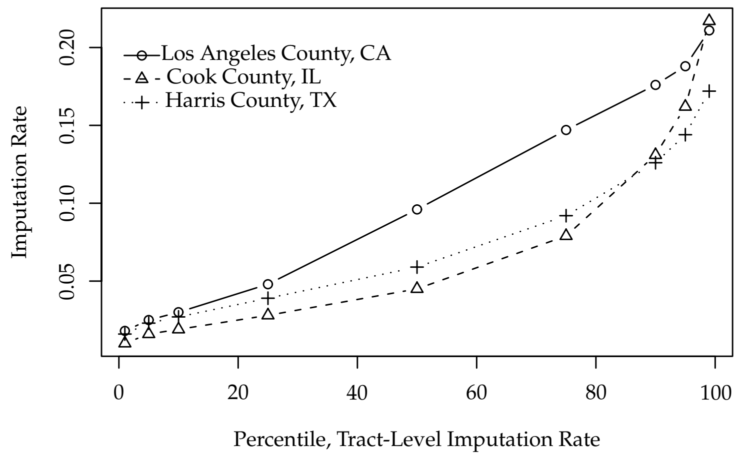

Figure 8.6 Imputation Rates for Race in Census Tracts of Selected Counties

NOTE: See note to Figure 8.5.

SOURCE: Summary File 1, accessed from the Census Bureau's Web site, http://factfinder.census.gov.

Examination of the characteristics of cases with different race categories reported compared with those with the same category reported reveal some interesting differences. Non-Hispanics were much more likely than Hispanics to have the same race reported in the two samples. Fully 45 percent of Hispanics have different races reported. Again, this indicates there may be a mismatch between the race and ethnicity classification scheme and the manner in which Hispanics understand themselves to fit into the scheme. Age also played a role; among Hispanics and non-Hispanics, younger people were less likely to have the same race reported than older people. Renters were more likely to have different races reported than owners. By region, Westerners were the most likely and Midwesterners the least likely to have different races reported in the two samples.

8–D FUTURE REQUIREMENTS

Race is consequential—racial classification can create a social reality that has real and enduring consequences. It can affect access to resources, the distribution of wealth and employment, interpersonal relationships, and place and type of residence. OMB provides federal statistical and program administration agencies with classification standards for maintaining, collecting, and reporting data on race and ethnicity. These data, in turn, are used by multiple entities for diverse purposes—statistical purposes, program administration, business planning, research, and civil rights enforcement, to name just a few. The way the federal government chooses to measure race and ethnicity has profound implications.

The research program surrounding the revision of Directive 15 and accompanying the 2000 census is only a beginning. We still do not have a deep understanding of the way that individuals classify themselves and how others classify them. The changes to Directive 15 did not produce a coherent organizing concept on which to base the federal system of racial and ethnic classification (see Harris, 2003, for discussion and argument for a particular organizing concept). We recognize that it could be difficult to gain acceptance of one concept, given the different uses of federal data on race and ethnicity, such as identifying members of groups that are victims of historical and current racial discrimination as defined in specific statutes in contrast to studying social behavior of self-defined ethnic groups. Nevertheless, continued lack of a coherent foundation is likely to lead to further instability in both the response to and content of questions on race and ethnicity.

Because race and ethnicity are social constructs (see Harris, 2003), definitions of racial and ethnic categories change over time and differ over cultures. Consequently, there can be considerable ambiguity in their definition at any one moment. Such ambiguity was evident in the 2000 census results, as well as in comparison data sources. Many factors are known to affect race and ethnicity reporting, including who responds, the setting for response, question wording and placement, and whether and what kinds of examples are given for race categories to prompt reporting. Imputation is also a factor—5 percent of household members did not report race in 2000, and the same percentage did not report ethnicity (fewer than 1 percent failed to report both).

A continued distinction between race and Hispanic origin may already reflect an outdated mode of thought about race and ethnicity. After reallocation to other categories based on written responses to the "some other race" category, 97 percent of those remaining in the category are Hispanic. In turn, they make up about half of all individuals indicating that they are of Hispanic origin, suggesting that a large portion of that population considers Hispanicity to be conceptually coherent with groups classified as races. Analysis of the 1996 RAETT indicated that the combined race/ethnicity question, which allowed multiple categories to be checked, did not result in a reduction of the proportion of the population indicating Hispanic origin. This result is in contrast to findings from the 1996 NCS and the 1995 CPS supplement on race and ethnicity, both of which indicated that a combined race and ethnicity question would reduce the proportion of the population reporting Hispanic origin. However, unlike the RAETT, neither the NCS nor the CPS supplement tested a combined question with a "check one or more" option. Given that OMB has adopted the "check one or more" format, it may therefore be appropriate to continue to explore the implications of making a combined race and Hispanic-origin question the standard format for future censuses.

Many of the available analyses suggest that race and ethnicity (particularly race) are fluid concepts, for which reporting may be affected by what appear to be minor changes in context, format, and procedures. Users need to be aware of the high levels of inconsistent reporting for some race groups in their analyses. The Census Bureau needs to continually conduct testing and research on race and ethnicity to evaluate data quality and plan refinements in question design for future censuses.

The Bureau tested minor changes in the race and ethnicity questions in a national mailout survey in 2003 (Martin et al., 2003). The major change in the test was to eliminate the "Some Other Race" category for one of the experimental samples. This change would bring the census race question into line with the OMB guidelines and make the census race data comparable with data from other surveys and other agencies.

Because it is likely that concepts of race and ethnicity will continue to change, the Census Bureau should consider tests of more striking modifications of the standard race and ethnicity questions. For example, tests could be conducted of asking people to write in their race(s) rather than check a box, of including Hispanic as a racial category and dropping a separate question on ethnicity, or of asking people to report not only their self-perception, but also their perception of how others would report them. Such tests, while not necessarily representing feasible options for the 2010 census, could shed considerable light on the meaning of race and ethnicity in the United States today, which would be useful for census planning and for public enlightenment.

> **Finding 8.1:** People who marked more than one race category in the 2000 census (the first to allow this reporting option) accounted for over 2 percent of the total population and as much as 8 percent of children ages 0 to 4, suggesting that the multirace population will grow in numbers. Nearly one-third of multirace respondents were of Hispanic origin, as were 97 percent of people checking only "some other race." Together, multirace and some other race Hispanic respondents accounted for about one-half of all Hispanics, indicating the ambiguities confronting measurement of race for the Hispanic group. Consistency of reporting of Hispanic origin (as measured by responses of E-sample households compared with matching P-sample households) was very high (98 percent); consistency of race reporting was also high for non-Hispanic whites, blacks, and Asians, but quite low for multirace respondents, and only moderately high for other groups. Both missing data rates and distributions for ethnicity and race are sensitive to differences in question format, order, and wording.

> *Recommendation 8.1:* **The Census Bureau should support—both internally and externally, in cooperation with other statistical agencies—ongoing, intensive, and innovative research and testing on race and ethnicity reporting. Particular attention should be given to testing formats that increase consistency of reporting and to methods for establishing comparability between old and new definitions and measures.**

CHAPTER 9

Management and Research

IN THIS CHAPTER, WE ADDRESS two elements that are critical to a successful census. They are: the organizational and management structure within the Census Bureau for taking the census together with the mechanisms for heeding internal and external advice (9–A) and the research, testing, and experimentation program that is required for evaluating one census and planning the next (9–B).

9–A ORGANIZATION AND MANAGEMENT STRUCTURE

9–A.1 2000 Census Organization

The 2000 census organizational structure was similar in broad outline to the 1990 structure (see Thompson, 2000). Of the Census Bureau's eight directorates in 2000, each headed by an associate director, seven were involved in the census: decennial census, field operations, communications, information technology, demographic programs, methodology and standards, and finance and administration. Five positions were key: the director, the deputy director, the associate and assistant directors for decennial census, and the associate director for field operations.

The director, a presidential appointee, principally handled relations with Congress, the Commerce Department, and other key stakeholders. The deputy director helped make planning and operational decisions when there was disagreement among other managers; from January 2001 to June 2002, the deputy director also served as acting director during the change of presidential administrations. The associate director for decennial census oversaw all decisions for planning, budget, and operations. Under the associate director, the assistant director for decennial census served as the 2000 census chief operating officer and had direct line responsibility for the Decennial Management Division, the Geography Division, the Decennial Statistical Studies Division (which implemented and evaluated the 2000 Accuracy and Coverage Evaluation Program, or A.C.E.), and the Decennial Systems and Contracts Management Office, which housed the programmers responsible for data processing and supervised contractors for three of the four data capture centers and other contractor operations.

The associate director for field operations oversaw data collection by the Field Division, which at the time of the 2000 census encompassed 12 permanent regional offices, 12 regional census centers, the Puerto Rico area office, 520 temporary local census offices (and 6 offices in Puerto Rico), the National Processing Center in Jeffersonville, Indiana, and the Technologies Management Office (which develops computer-assisted data collection questionnaires and procedures). The Field Division also employed and directed the massive temporary corps of enumerators.

Other Census Bureau divisions involved in the 2000 census included the Population Division and the Housing and Household Economic Statistics Division in the demographic programs directorate and the Planning, Research, and Evaluation Division (PRED) in the methodology and standards directorate.

The Bureau developed a variety of mechanisms for internal coordination. An executive-level steering committee for the census, meeting every 2 weeks, was formed in 1997. It included the director, the deputy director, the two principal associate directors, the associate and assistant directors for decennial census and field operations, the associate director for communications, and the congressional affairs office chief. In 2000 the Executive Steering Committee for A.C.E. Policy (ESCAP) was formed to review the plans for the A.C.E., monitor A.C.E. implementation, review A.C.E. evaluation studies, and make recommendations to the director or acting director about whether or not to adjust the census counts for an expected measured net undercount (see Section 5–D.2). Staff teams cutting across directorates were formed as needed for specific tasks.

For external oversight, the Bureau answered to an unprecedented array of groups (see Section 3–C). They included the Congressional Monitoring Board (presidential and congressional appointees), the U.S. House Subcommittee on the Census of the Committee on Government Reform (and successor subcommittees charged with census oversight), the Department of Commerce Inspector General, and the U.S. General Accounting Office. External input was sought from our Panel to Review the 2000 Census, the Panel on Research on Future Census Methods, the Secretary of Commerce's 2000 Census Advisory Committee, the Census Bureau's Professional Associations Advisory Committee, and advisory committees for minority groups.

9–A.2 Assessment

The basic structure of the Census Bureau—directorates including divisions, each largely consisting of staff with particular expertise—is typical of large data collection organizations. It is also not surprising that almost every directorate in the Census Bureau would contribute staff and expertise to the decennial census, the centerpiece of the Bureau's activity. However, the panel's observations, discussions with key Census Bureau staff, and individual administrative experiences suggest that the census organization was not as effective as it could have been to ensure smooth-running, high-quality operations and data.

Three major problems stand out in the panel's view. First, there was no project director for the census with authority over the census-related work of all staff involved in census operations. It is common in data collection organizations (e.g., Statistics Canada, private survey firms) to use a matrix approach in which staff from expertise-based divisions are assembled as a team to work on a particular project under the direct authority of a project leader (or similar title). The associate director for decennial census came closest to such a person for the 2000 census, but he did not have direct authority over some staffs, such as those in the field and demographic programs directorates.

Second, across and within directorates, divisions tended to operate as individual fiefdoms. For example, the Population Division developed the address list for group quarters with little coordination with the Geography Division staff that had responsibility for compiling the Master Address File (MAF). The Geography Division staff tended to focus on assigning structure addresses to the correct geographic areas; they were less sensitive to the varieties of housing unit types across the nation (e.g., single-family houses carved up into apartments with a common mailing address) that required careful handling to avoid duplication or omission. As another example, the PRED Division, which was charged to evaluate census operations, was deliberately kept independent from the decennial directorate. However, this separation was carried to such an extreme that PRED staff and staff of other evaluation units (including the A.C.E. staff and subject-matter specialists in the demographic programs directorate) were often not adequately apprised of or able to inform each other's efforts. Moreover, operational staff were often not cognizant of or sensitive to the needs of evaluation staff for accessible data files and other information. (See also Morganstein et al., 2003, who make a similar point with regard to the lack of priority attached to quality assurance during data collection.)

For many tasks, project groups were formed from staffs of multiple divisions; however, such groups tended to operate not as project teams but as committees, in which each member's allegiance and lines of communication were primarily with the home division. Consequently, responsibilities were often not clearly defined, and interactions were often highly bureaucratic. For other tasks, needed coordination and feedback across divisions was not obtained on a timely basis. For example, the PRED staff apparently had no input into the design of the MAF database structure; the coding schemes used in the MAF made it difficult to identify the sources that contributed to each address and led to delays in completing crucial assessments. As another example, contractors for the data capture centers had no input into questionnaire design and printing, which put the data capture operation at risk (see Titan Corporation, 2003).

Symptomatic of and contributing to problems in coordination was that standards for documentation of data files and operations and for specification of key variables (e.g., flags for types of imputations) were not uniform across divisions or, in some cases, across branches within a division. Consequently, the quality of available documentation varies widely, as does the usability of data files that are critical for quality control and evaluation. Generally, the A.C.E Program and other operations specified by the Decennial Statistical Studies Division (e.g., long-form sample weighting) followed good practices for documentation, and the A.C.E. files are easy to use for many applications. However, other data files were not designed to facilitate evaluation; these include the MAF, management operations files, and data processing files. Also, important operations such as the system for processing complete-count data and imputing for item nonresponse in the long-form sample were not well documented, which hampers both internal and external evaluation.

Third, communication with and involvement of outside resource people and stakeholders was far from optimal. In particular, communication channels among local governments, geographers located in the Census Bureau's regional offices, and headquarters Geography Division staff were often muddied. There was too little feedback from headquarters to the regions and localities about schedule changes and other matters that made it difficult for localities to participate effectively in local review. There was also too little involvement of localities in such operations as determining sensible "blue lines" (the map lines demarcating areas for mailout/mailback procedures from update/leave procedures). Case studies by participants in the Local Update of Census Addresses (LUCA) Program document many instances of inappropriate blue-line designations, which not only complicated local review but also created logistical problems for delivery of census questionnaires (see Working Group on LUCA, 2001:Chap. 4).

We could not and have not undertaken a comprehensive management review of the 2000 census, and so our critique of the census organizational structure is deliberately limited in nature. However, we believe that the problems we have identified affected important aspects of census operations and impeded the ability to conduct imaginative, timely, and useful evaluations of those operations (see also IBM Business Consulting Services, 2003; Morganstein et al., 2003).

Finding 9.1: From the panel's observations and discussion with key Census Bureau staff, it appears that the decentralized and diffuse organization structure for the 2000 census impeded some aspects of census planning, execution, and evaluation. There was no single operational officer (below the level of director or deputy director of the Bureau) clearly in charge of all aspects of the census; the structure for decision-making and coordination across units was largely hierarchical; and important perspectives inside the Bureau and from regional offices, local partners, and contractors were not always taken into account. These aspects of the 2000 management structure affected two areas in particular: (1) development of the Master Address File (MAF), which experienced numerous problems, and (2) the program to evaluate census processes and data quality, from which results were slow to appear and are often of limited use for understanding the quality of the 2000 census or for planning the 2010 census.

Finding 9.2: The quality of documentation and usability varies among internal 2000 census data files and specifications that are important for evaluation. Generally, the A.C.E. Program followed good practices for documentation, and the A.C.E. files are easy to use for many applications. However, the lack of well-documented and usable data files and specifications hampered timely evaluation of other important aspects of the census, such as the sources contributing to the Master Address File and the implementation of imputation routines.

9–B EVALUATION PROGRAM

9–B.1 Completing 2000 Census Evaluations

The Census Bureau developed an ambitious evaluation program for the 2000 census. The evaluations that were conducted to assess coverage error in the census and to evaluate the two major coverage measurement programs—A.C.E. and demographic analysis—were generally of high quality, informative, and completed on a timely schedule (see discussion in Chapter 6).

A large number of evaluations of the census itself were planned to cover virtually every aspect of the data and operations (see listing in Appendix I). Some projects were subsequently cancelled or combined with other projects because of budget limitations or because they were superseded by an A.C.E.-related evaluation. The list of completed evaluations, however, is still sizable. These evaluations began to appear only in summer 2003—late for informing users and 2010 data planners (see Appendix I).[1] We conjecture that many evaluations were hampered by the need to devote additional resources to the various A.C.E. evaluations and by the fact that many key data files were not well designed for evaluation purposes. Consequently, the evaluation staff had to devote substantial time to obtaining usable data, leaving relatively little time for analysis.

We applaud the effort the Census Bureau devoted to evaluation of the 2000 census. Yet we must note the serious deficiencies of many (but by no means all) of the evaluation studies released to date. Too often, the evaluations do not clearly answer the needs of relevant audiences, which include 2000 census data users who are interested in data quality issues that bear on their analyses and Census Bureau staff and others who are concerned with the lessons from the 2000 experience that can inform 2010 planning. No single evaluation will necessarily speak to both audiences, but every evaluation should clearly speak to at least one of them.

Yet many of the completed evaluations are accounting-type documents rather than full-fledged evaluations. They provide authoritative information on such aspects as number of mail returns by day, complete-count item nonresponse and imputation rates by type of form and data collection mode, and enumerations completed in various types of special operations (e.g., urban update/leave, list/enumerate). This information is valuable but limited. Many reports have no analysis as such, other than simple one-way and two-way tabulations. Reports sometimes use different definitions of variables, such as type of form (self or enumerator), and obtain data from files at different stages of processing, with little attempt to reconcile such differences. Almost no reports provide tables or other

[1]We thank the Census Bureau for providing advance copies of evaluation reports to us (and also to the U.S. General Accounting Office, the Department of Commerce inspector general, and relevant congressional offices).

analyses that look at operations and data quality for geographic areas.

Such limited operational evaluations are of little benefit to users, who need more context than is provided in many of the evaluation reports and who need additional meaningful detail, such as variations in item imputation rates among population groups and geographic areas and patterns of nonresponse for more than one variable considered at a time. (Census Bureau summary files provide item imputation rates for geographic areas, but these rates are for the total applicable universe for a single variable. Many of the rates combine the household and group quarters populations.)

Similarly, 2010 planners need analysis that is explicitly designed to answer important questions for research and testing to improve the 2010 census. One useful analysis, for example, might focus on factors that explain variations in mail response rates, cross-sectionally and in comparison with 1990 rates, given how critical mail response is to the timeliness of census operations and the quality of census data.

Imaginative data analysis techniques—using multiple regression, exploratory data analysis, and graphical analysis—could yield important findings as well as facilitate effective presentation of results. For example, a graphical analysis showing geographic areas with particularly high or low whole-household imputation rates compared with 1990 could vividly summarize a wealth of data in a manner that would be helpful to users and suggest hypotheses for future research and testing as part of the 2010 planning process.

Topic reports from the evaluation program that summarize individual evaluations (see Appendix I) also vary in usefulness. Some of them provide a well-rounded picture of broad aspects of census operations, such as data capture and outreach programs, but others do not add much beyond the individual evaluations. Both the individual evaluations and topic reports tend to focus on specific slices of census operations (e.g., data capture, data processing). There are almost no studies that focus on an outcome of interest, such as the higher rates of whole-household and whole-person imputations in 2000 compared with 1990, or seek to describe patterns among geographic areas and population groups that would be important for users to understand, or seek to analyze explanatory factors and their importance for 2010 research and testing.

It is not too late for more imaginative and focused analysis to be conducted to benefit data users and the 2010 census planning process. In fact, several additional data sources exist that could support rich analyses of many important topics. At the urging of several panels of the Committee on National Statistics (see, e.g., National Research Council, 1988, 2000a), the Census Bureau for the first time is developing a Master Trace Sample of information from all of the various 2000 census operations for a sample of addresses (processes related specifically to long-form items are not included—see Hill and Machowski, 2003). Once fully completed, this file will support evaluations of specific 2000 operations on enumerations, and it should also enable 2010 planners to query the sample as a means of simulating the likely effects of various procedures being considered for 2010. For example, the Master Trace Sample might be used to evaluate the relationship of the number of enumerator visits to the final status of an enumeration (e.g, household response, proxy response, occupancy status not known), which, in turn, could inform decisions about the optimal number of callbacks to require in nonresponse follow-up in 2010. Developing this database—and working with the research community in order to best use it—should remain a top priority.

Other useful 2000 databases for analysis include the A.C.E. P-sample and E-sample files; the Person Duplication Study files of duplicate enumerations of the census and the E-sample; extracts from the Master Address File; an exact match of census records with the March 2000 Current Population Survey (CPS); and the 2000 Census Public Use Microdata Samples (PUMS) files. Each of these data sets could support important analyses. For example, the Person Duplication Study files could be analyzed to answer such questions as: What were the different types of duplicate enumerations in 2000? Among which population groups and geographic areas were they most likely to occur? What do the results suggest for 2010 research and testing and, specifically, for plans to eliminate duplicates during the actual enumeration? The 2000 PUMS files could be analyzed to answer such questions as: What were the patterns of missing long-form data among population groups and across geographic areas? What would be the effects on distributions of using alternative imputation methods? What do the results suggest for improving response in 2010 or the American Community Survey?

The census-March CPS exact match file could be used not only to study imputations, but also to investigate possible reporting errors for census (or CPS) items. MAF extracts, perhaps linked to the Master Trace Sample, could be studied to learn more about the factors contributing to duplication.

We understand that resources are limited, yet careful consideration of priority areas for investigation with already-available data from the 2000 census could lead to important research findings for 2010 planning and could further inform users about 2000 census data quality.

> *Recommendation 9.1:* **The Census Bureau should mine data sources created during the 2000 census process, such as the A.C.E. data, Person Duplication Studies, extracts from the Master Address File, a match of census records and the March 2000 Current Population Survey, and the Master Trace Sample. Such data can illuminate important outstanding questions about patterns of data quality and factors that may explain them in the 2000 census and suggest areas for research and testing to improve data quality in the 2010 census and the American Community Survey.**

9–B.2 Strengthening the Evaluation Component for 2010

It is critically important for the Census Bureau to strengthen the evaluation component of the 2010 census, beginning with the tests and other analyses that are planned to lead up to the 2008 dress rehearsal. To do this, Bureau staff must first step back and ask what are the priority questions that evaluations should address. Accounting-type reports are important, but they are not a substitute for evaluations that seek to inform users about important aspects of data quality or to develop and test hypotheses regarding factors that contribute to key census outcomes that are important for future census planning.

Second, census operations must be fully specified, and all operational data systems must be fully documented and designed to facilitate evaluation. A recurring theme in the 2000 census evaluation reports is the difficulties encountered by the evaluation staff because important processes and data files were not properly documented. The Bureau is currently reengineering its software development procedures for all of the data processing, management information, and administrative computer systems that will be required for 2010 (see National Research Council, 2003a). This reengineering should incorporate procedures for each system to provide data that are readily accessible for quality control and evaluation purposes, thereby creating a Master Trace Sample (or samples) on a real-time basis. The availability of evaluation information as part of census operations will be critical if the Bureau is to implement computerized procedures for reducing duplicates as part of the census itself. It will also be critical to timely evaluation and corrective actions for other census procedures and for timely evaluation for planning the census in 2020. Evaluation-friendly software systems should be designed and tested as soon as possible as part of the 2010 census testing program.

The 2010 evaluation program should also make use of statistical tools for data analysis and presentation that can help make sense of large amounts of information. For example, graphical analysis by geographic areas could not only inform postcensus evaluation but also be a useful management tool during the census to identify potential problem areas. Generally, the Census Bureau should seek ways to include real-time evaluations as part of the census, in order to help identify unexpected problems and facilitate appropriate corrective action. An ongoing evaluation program will also be critical to the ability of the planned American Community Survey to change course as needed in order to provide useful data to the public. The preparation by Population Division staff of local-area housing unit estimates from administrative records in the first half of 2000 is an example of a real-time evaluation that was critical to improving the enumeration. These evaluations strongly suggested a sizable net overcount of housing units in the MAF, which led to the special unduplication operation in summer 2000 (see Section 4–E).

A critical element of a successful evaluation program is sufficient well-trained and experienced staff to identify evaluation needs, specify evaluation requirements for census data systems and documentation, and carry out the evaluations. The Census Bureau, like other federal government agencies, is facing loss of experienced staff due to retirement and other reasons. In addition, the need for a large A.C.E. Program in 2010 will place heavy demands on research and evaluation staff. The Census Bureau will need to give priority to development of technical staff resources and seek ways to augment those resources through such means as involving outside researchers in some studies (as was done for some of the 2000 evaluations).

Finally, the 2010 research, testing, and evaluation program must be open to sharing preliminary results with outside researchers for critical review and feedback. The 2010 evaluation program should also be open to sharing data sets (including microdata with appropriate confidentiality safeguards) that can make it possible for outside researchers to conduct useful studies of the population and contribute to understanding of census data quality.

With regard to the sharing of preliminary results, it is too time-consuming, lengthy, and cumbersome a process to wait until evaluation results have been completely vetted inside the Census Bureau to share them with researchers. Admittedly, the highly critical environment surrounding the 2000 census would caution any agency to be careful about releasing preliminary results, but constrained resources in the Bureau for research place a premium on effective ways of obtaining outside input from the research community. Sharing of preliminary results (clearly labeled as such) should become part of the Census Bureau culture.

> *Recommendation 9.2:* **In addition to pursuing improvements for coverage evaluation in 2010 (see recommendation 6.1), the Census Bureau must materially strengthen the evaluation component for census operations and data quality in 2010 (and in the current testing program) in the following ways:**
>
> 1. **Identify important areas for evaluations to meet the needs of users and census planners and set evaluation priorities accordingly;**
> 2. **Design and document data collection and processing systems so that information can be readily extracted to support timely, useful evaluation studies;**
> 3. **Use graphical and other exploratory data analysis tools to identify patterns (e.g., mail return rates, imputation rates) for geographic areas and popu-**

lation groups that may suggest reasons for variations in data quality and ways to improve quality (such tools could also be useful in managing census operations);

4. **Explore ways to incorporate real-time evaluation during the course of the census;**
5. **Give priority to development of technical staff resources for research, testing, and evaluation; and**
6. **Share preliminary analyses with outside researchers for critical assessment and feedback.**

Recommendation 9.3: **The Census Bureau should seek ways to expand researcher access to microdata from and about the 2000 census in order to further understanding of census data quality and social science knowledge. Such data files as the 2000 A.C.E. E-sample and P-sample output files, for example, should be deposited with the Bureau's Research Data Centers. To help the Bureau evaluate population coverage and data quality in the 2010 census, the Bureau should seek ways—using the experience with the Panel to Review the 2000 Census as a model—to furnish preliminary data, including microdata, to qualified researchers under arrangements that protect confidentiality.**

CHAPTER 10

Detailed Findings and Recommendations

FOR THE READER'S CONVENIENCE, we list below the 19 specific findings for 2000 and the 16 recommendations for 2010 that appear in Chapters 3–4 and 6–9 of this report. See Section 1–D for the panel's general, overall findings.

OVERALL CENSUS DESIGN

Finding 3.1: The lack of agreement until 1999 on the basic census design among the Census Bureau, the administration, and Congress hampered planning for the 2000 census, increased the costs of the census, and increased the risk that the census could have been seriously flawed in one or more respects.

Recommendation 3.1: **The Census Bureau, the administration, and Congress should agree on the basic census design for 2010 no later than 2006 in order to permit an appropriate, well-planned dress rehearsal in 2008.**

Recommendation 3.2: **The Census Bureau, the administration, and Congress should agree on the overall scheme for the 2010 census and the new American Community Survey (ACS) by 2006 and**

preferably earlier. Further delay will undercut the ability of the ACS to provide, by 2010, small-area data of the type traditionally collected on the census long form and will jeopardize 2010 planning, which currently assumes a short-form-only census.

ASSESSMENT OF 2000 CENSUS OPERATIONS

Finding 4.1: The use of a redesigned questionnaire and mailing strategy and, to a more limited extent, of expanded advertising and outreach—major innovations in the 2000 census—contributed to the success achieved by the Census Bureau in stemming the decline in mail response rates observed in the two previous censuses. This success helped reduce the costs and time of follow-up activities.

Recommendation 4.1: **The Census Bureau must proceed quickly to work with vendors to determine cost-effective, timely ways to mail a second questionnaire to nonresponding households in the 2010 census, in order to improve mail response rates, in a manner that minimizes duplicate enumerations.**

Finding 4.2: Contracting for selected data operations, using improved technology for capturing the data on the questionnaires, and aggressively recruiting enumerators and implementing nonresponse follow-up were significant innovations in the 2000 census that contributed to the timely execution of the census.

Finding 4.3: The greater reliance on imputation routines to supply values for missing and inconsistent responses in 2000, in contrast to the greater reliance on telephone and field follow-up of nonrespondents in 1990, contributed to the timely completion of the 2000 census and to containing the costs of follow-up. It is not known whether the distributions of characteristics and the relationships among characteristics that resulted from imputation (particularly of long-form content) were less accurate than the distributions and relationships that would have resulted from additional follow-up.

Recommendation 4.2: **Because the 2000 census experienced high rates of whole-household nonresponse and missing responses for individual long-form items, the Census Bureau's planning for the 2010 census and the American Community Survey should include research on the trade-offs in costs and accuracy between imputation and additional field work for missing data. Such research should examine the usefulness of following up a sample of households with missing data to obtain information with which to improve the accuracy of imputation routines.**

Finding 4.4: The use of multiple sources to build a Master Address File (MAF)—a major innovation in 2000—was appropriate in concept but not well executed. Problems included changes in schedules and operations, variability in the efforts to update the MAF among local areas, poor integration of the address list for households and group quarters, and difficulties in determining housing unit addresses in multiunit structures. Changes were made to the MAF development: a determination late in the decade that a costly complete block canvass was needed to complete the MAF in mailout/mailback areas and a determination as late as summer 2000 that an ad hoc operation was needed to weed out duplicate addresses. Problems with the MAF contributed to census enumeration errors, including a large number of duplicates.

Finding 4.5: The problems in developing the 2000 Master Address File underscore the need for a thorough evaluation of the contribution of various sources, such as the U.S. Postal Service Delivery Sequence File and the Local Update of Census Addresses Program, to accuracy of MAF addresses. However, overlapping operations and unplanned changes in operations were not well reflected in the coding of address sources on the MAF, making it difficult to evaluate the contribution of each source to the completeness and accuracy of the MAF.

Recommendation 4.3: **Because a complete, accurate Master Address File is not only critical for the 2010 census, but also important for the 2008 dress rehearsal, the new American Community Survey, and other Census Bureau surveys, the Bureau must develop more effective procedures for updating and correcting the MAF**

than were used in 2000. Improvements in at least three areas are essential:

1. **The Census Bureau must develop procedures for obtaining accurate information to identify housing units within multiunit structures. It is not enough to have an accurate structure address.**
2. **To increase the benefit to the Census Bureau from its Local Update of Census Addresses (LUCA) and other partnership programs for development of the MAF and the TIGER geocoding system, the Bureau must redesign the program to benefit state and local governments that participate. In particular, the Bureau should devise ways to provide updated MAF files to participating governments for statistical uses and should consider funding a MAF/TIGER/LUCA coordinator position in each state government.**
3. **To support adequate assessment of the MAF for the 2010 census, the Census Bureau must plan evaluations well in advance so that the MAF records can be assigned appropriate address source codes and other useful variables for evaluation.**

Finding 4.6: The enumeration of people in the 2000 census who resided in group quarters, such as prisons, nursing homes, college dormitories, group homes, and others, resulted in poor data quality for this growing population. In particular, missing data rates, especially for long-form items, were much higher for group quarters residents than for household members in 2000 and considerably higher than the missing data rates for group quarters residents in 1990 (see Finding 7.3). Problems and deficiencies in the enumeration that undoubtedly contributed to poor data quality included: the lack of well-defined concepts of types of living arrangements to count as group quarters; failure to integrate the development of the group quarters address list with the development of the Master Address File; failure to plan effectively for the use of administrative records in enumerating group quarters residents; errors in assigning group quarters to the correct geographic areas; and poorly controlled tracking and case management for group quarters. In addition, there was no program to evaluate the completeness of population coverage in group quarters.

Recommendation 4.4: **The Census Bureau must thoroughly evaluate and completely redesign the processes related to group quarters populations for the 2010 census, adapting the design as needed for different types of group quarters. This effort should include consideration of clearer definitions for group quarters, redesign of questionnaires and data content as appropriate, and improvement of the address listing, enumeration, and coverage evaluation processes for group quarters.**

ASSESSMENT OF COVERAGE IN 2000

Finding 6.1: The 2000 Accuracy and Coverage Evaluation (A.C.E.) Program operations were conducted according to clearly specified and carefully controlled procedures and directed by a very able and experienced staff. In many respects, the A.C.E. was an improvement over the 1990 Post-Enumeration Survey, achieving such successes as high response rates to the P-sample survey, low missing data rates, improved quality of matching, low percentage of movers due to more timely interviewing, and substantial reductions in the sampling variance of coverage correction factors for the total population and important population groups. However, inaccurate reporting of household residence in the A.C.E. (which also occurred in the census itself) led to substantial underestimation of duplicate enumerations in 2000 in the original (March 2001) A.C.E. estimates.

Finding 6.2: The Census Bureau commendably dedicated resources to the A.C.E. Revision II effort, which completely reestimated net undercount (and overcount) rates for several hundred population groups (poststrata) by using data from the original A.C.E. and several evaluations. The work exhibited high levels of creativity and effort devoted to a complex problem. From innovative use of matching technology and other evaluations, it provided substantial additional information about the numbers and sources of erroneous census enumerations and, similarly, information with which to correct the residency status of the independent A.C.E. sample. It provided little additional information, however, about the numbers and sources of census omissions.

Documentation for the original A.C.E. estimates (March 2001), the preliminary revised estimates (October 2001), and the A.C.E. Revision II estimates (March 2003) was timely, comprehensive, and thorough.

Finding 6.3: We support the Census Bureau's decision not to use the March 2003 Revision II A.C.E. coverage measurement results to adjust the 2000 census base counts for the Bureau's postcensal population estimates program. The Revision II results are too uncertain to be used with sufficient confidence about their reliability for adjustment of census counts for subnational geographic areas and population groups. Sources of uncertainty stem from the small samples of the A.C.E. data that were available to correct components of the original A.C.E. estimates of erroneous enumerations and non-A.C.E. residents and to correct the original estimate of nonmatches and the consequent inability to make these corrections for other than very large population groups; the inability to determine which of each pair of duplicates detected in the A.C.E. evaluations was correct and which should not have been counted in the census or included as an A.C.E. resident; the possible errors in subnational estimates from the choice of one of several alternative correlation bias adjustments to compensate for higher proportions of missing men relative to women; the inability to make correlation bias adjustments for population groups other than blacks and nonblacks; and the possible errors for some small areas from the use of different population groups (poststrata) for estimating erroneous census enumerations and census omissions. In addition, there is a large discrepancy in coverage estimates for children ages 0–9 when comparing demographic analysis estimates with Revision II A.C.E. estimates (2.6 percent undercount and 0.4 percent net overcount, respectively).

Finding 6.4: Demographic analysis helped identify possible coverage problems in the 2000 census and in the A.C.E. at the national level for a limited set of population groups. However, there are sufficient uncertainties in the revised estimates of net immigration (particularly the illegal component) and the revised assumption of completeness of birth registration after 1984, compounded by the difficulties of classifying people by race, so that the revised demographic analysis estimates cannot and should not serve as the definitive standard of evaluation for the 2000 census or the A.C.E.

Finding 6.5: Because of significant differences in methodology for estimating net undercount in the 1990 Post-Enumeration Survey Program and the 2000 Accuracy and Coverage Evaluation Program (Revision II), it is difficult to compare net undercount estimates for the two censuses. Nevertheless, there is sufficient evidence (from comparing the 1990 PES and the original A.C.E.) to conclude that the national net undercount of the household population and net undercount rates for population groups were reduced in 2000 from 1990 and, more important, that differences in net undercount rates between historically less-well-counted groups (minorities, children, renters) and others were reduced as well. From smaller differences in net undercount rates among groups and from analysis of available information for states and large counties and places, it is reasonable to infer that differences in net undercount rates among geographic areas were also probably smaller in 2000 compared with 1990. Despite reduced differences in net undercount rates, some groups (e.g., black men and renters) continued to be undercounted in 2000.

Finding 6.6: Two factors that contributed to the estimated reductions in net undercount rates in 2000 from 1990 were the large numbers of whole-person imputations and duplicate census enumerations, many of which were not identified in the original (March 2001) A.C.E. estimates. Contributing to duplication were problems in developing the Master Address File and respondent confusion about or misinterpretation of census "usual residence" rules, which resulted in duplication of household members with two homes and people who were enumerated at home and in group quarters.

Recommendation 6.1: **The Census Bureau and administration should request, and Congress should provide, funding for the development and implementation of an improved Accuracy and Coverage Evaluation Program for the 2010 census. Such a program is essential to identify census omissions and erroneous enumerations and to provide the basis for adjusting the census counts for coverage errors should that be warranted.**

The A.C.E. survey in 2010 must be large enough to provide estimates of coverage errors that provide the level of precision targeted for the original (March 2001) A.C.E. estimates for population groups and geographic areas. Areas for improvement that should be pursued include:

1. **the estimation of components of gross census error (including types of erroneous enumerations and omissions), as well as net error;**
2. **the identification of duplicate enumerations in the E-sample and nonresidents in the P-sample by the use of new matching technology;**
3. **the inclusion of group quarters residents in the A.C.E. universe;**
4. **improved questionnaire content and interviewing procedures about place of residence;**
5. **methods to understand and evaluate the effects of census records that are excluded from the A.C.E. matching (*II*s);**
6. **a simpler procedure for treating people who moved between Census Day and the A.C.E. interview;**
7. **the development of poststrata for estimation of net coverage errors, by using census results and statistical modeling as appropriate; and**
8. **the investigation of possible correlation bias adjustments for additional population groups.**

Recommendation 6.2: **The Census Bureau should strengthen its program to improve demographic analysis estimates, in concert with other statistical agencies that use and provide data inputs to the postcensal population estimates. Work should focus especially on improving estimates of net immigration. Attention should also be paid to quantifying and reporting measures of uncertainty for the demographic estimates.**

Recommendation 6.3: **Congress should consider moving the deadline to provide block-level census data for legislative redistricting to allow more time for evaluation of the completeness of popula-**

tion coverage and quality of the basic demographic items before they are released.

ASSESSMENT OF BASIC AND LONG-FORM-SAMPLE DATA

Finding 7.1: Rates of missing data in 2000 were low at the national level for the basic demographic items asked of everyone (complete-count items)—age, sex, race, ethnicity, household relationship, and housing tenure. Missing data rates for these items ranged from 2 to 5 percent (including records for people with one or more missing items and people who were wholly imputed). Rates of inconsistent reporting for the basic items (as measured by comparing responses for census enumerations and matching households in the independent Accuracy and Coverage Evaluation survey) were also low. However, some population groups and geographic areas exhibited high rates of missing data and inconsistent reporting for one or more of the basic items. No assessments have yet been made of reporting errors for such items as age, nor of the effects of imputation on the distributions of basic characteristics or the relationships among them.

Finding 7.2: For the household population, missing data rates were at least moderately high (10 percent or more) for over one-half of the 2000 census long-form-sample items and very high (20 percent or more) for one-sixth of the long-form-sample items. Missing data rates also varied widely among population groups and geographic areas. By comparison with 1990, missing data rates were higher in 2000 for most long-form-sample items asked in both years and substantially higher—by 5 or more percentage points—for one-half of the items asked in both years. In addition, close to 10 percent of long-form-sample households in 2000 (similar to 1990) provided too little information for inclusion in the sample data file. When dropped households and individually missing data are considered together, the effective sample size that is available for analysis for some characteristics is 60 percent or less of the original long-form-sample size.

Many long-form-sample items had moderate to high rates of inconsistent reporting, as measured in a content reinterview survey. Few assessments have yet been made of systematic reporting errors for the long-form-sample items, although aggregate comparisons of employment data between the 2000 census and the Current Population Survey (CPS) found sizeable discrepancies in estimates of employed and unemployed people—much larger than the discrepancies found in similar comparisons for 1990. No analysis of the effects of item imputation and weighting on the distributions of characteristics or the relationships among them has yet been undertaken, although analysis determined that changes in imputation procedures contributed to the 50 percent higher unemployment rate estimate in the 2000 census compared with the April 2000 CPS.

Recommendation 7.1: **Given the high rates of imputation for many 2000 long-form-sample items, the Census Bureau should develop procedures to quantify and report the variability of the 2000 long-form estimates due to imputation, in addition to the variability due to sampling and weighting adjustments for whole-household weight adjustments. The Bureau should also study the effects of imputation on the distributions of characteristics and the relationships among them and conduct research on improved imputation methods for use in the American Community Survey (or the 2010 census if it includes a long-form sample).**

Recommendation 7.2: **The Census Bureau should make users aware of the high missing data rates and measures of inconsistent reporting for many long-form-sample items, and inform users of the 2000 census long-form-sample data products (Summary Files 3 and 4 and the Public Use Microdata Samples) about the need for caution in analyzing and interpreting those data.**

Finding 7.3: For group quarters residents, missing data rates for most long-form-sample items were very high in 2000 (20 percent or more for four-fifths of the items and 40 percent or more for one-half of the items). The 2000 rates were much higher than missing data rates for household members and considerably higher than missing data rates for group quarters residents in 1990. The 2000 missing data rates were particularly high for prisoners, residents of nursing homes, and residents of long-term-care hospitals perhaps because of heavy reliance on administrative records for enumerating them.

Few assessments have yet been made of systematic reporting errors for group quarters residents for long-form-sample items, nor of the effects of imputations on the distributions of characteristics or the relationships among them. However, a systematic error was found in the imputation of employment status for people living in noninstitutional group quarters because of a particular pattern of missing data. The result was a substantial overestimate of unemployment rates for these people, so much so that the Census Bureau reissued employment status tabulations for household members only, excluding group quarters residents.

Recommendation 7.3: **The Census Bureau should publish distributions of characteristics and item imputation rates, for the 2010 census and the American Community Survey (when it includes group quarters residents), that distinguish household residents from the group quarters population (at least the institutionalized component). Such separation would make it easier for data users to compare census and ACS estimates with household surveys and would facilitate comparative assessments of data quality for these two populations by the Census Bureau and others.**

RACE AND ETHNICITY MEASUREMENT

Finding 8.1: People who marked more than one race category in the 2000 census (the first to allow this reporting option) accounted for over 2 percent of the total population and as much as 8 percent of children ages 0 to 4, suggesting that the multirace population will grow in numbers. Nearly one-third of multirace respondents were of Hispanic origin, as were 97 percent of people checking only "some other race." Together, multirace and some other race Hispanic respondents accounted for about one-half of all Hispanics, indicating the ambiguities confronting measurement of race for the Hispanic group. Consistency of reporting of Hispanic origin (as measured by responses of E-sample households compared with matching P-sample households) was very high (98 percent); consistency of race reporting was also high for non-Hispanic whites, blacks, and Asians, but quite low for multirace respondents, and only moderately high for other groups. Both missing data rates and distributions for ethnicity and race are sensitive to differences in question format, order, and wording.

Recommendation 8.1: **The Census Bureau should support—both internally and externally, in cooperation with other statistical agencies—ongoing, intensive, and innovative research and testing on race and ethnicity reporting. Particular attention should be given to testing formats that increase consistency of reporting and to methods for establishing comparability between old and new definitions and measures.**

MANAGEMENT AND RESEARCH

Finding 9.1: From the panel's observations and discussion with key Census Bureau staff, it appears that the decentralized and diffuse organization structure for the 2000 census impeded some aspects of census planning, execution, and evaluation. There was no single operational officer (below the level of director or deputy director of the Bureau) clearly in charge of all aspects of the census; the structure for decision-making and coordination across units was largely hierarchical; and important perspectives inside the Bureau and from regional offices, local partners, and contractors were not always taken into account. These aspects of the 2000 management structure affected two areas in particular: (1) development of the Master Address File (MAF), which experienced numerous problems, and (2) the program to evaluate census processes and data quality, from which results were slow to appear and are often of limited use for understanding the quality of the 2000 census or for planning the 2010 census.

Finding 9.2: The quality of documentation and usability varies among internal 2000 census data files and specifications that are important for evaluation. Generally, the A.C.E. Program followed good practices for documentation, and the A.C.E. files are easy to use for many applications. However, the lack of well-documented and usable data files and specifications hampered timely evaluation of other important aspects of the census, such as the sources contributing to the Master Address File and the implementation of imputation routines.

Recommendation 9.1: **The Census Bureau should mine data sources created during the 2000 census process, such as the A.C.E. data, Person Duplication Studies, extracts from the Master Address File, a match of census records and the March 2000 Current Population Survey, and the Master Trace Sample. Such data can illuminate important outstanding questions about patterns of data quality and factors that may explain them in the 2000 census and suggest areas for research and testing to improve data quality in the 2010 census and the American Community Survey.**

Recommendation 9.2: **In addition to pursuing improvements for coverage evaluation in 2010 (see recommendation 6.1), the Census Bureau must materially strengthen the evaluation component for census operations and data quality in 2010 (and in the current testing program) in the following ways:**

1. **Identify important areas for evaluations to meet the needs of users and census planners and set evaluation priorities accordingly;**
2. **Design and document data collection and processing systems so that information can be readily extracted to support timely, useful evaluation studies;**
3. **Use graphical and other exploratory data analysis tools to identify patterns (e.g., mail return rates, imputation rates) for geographic areas and population groups that may suggest reasons for variations in data quality and ways to improve quality (such tools could also be useful in managing census operations);**
4. **Explore ways to incorporate real-time evaluation during the course of the census;**
5. **Give priority to development of technical staff resources for research, testing, and evaluation; and**
6. **Share preliminary analyses with outside researchers for critical assessment and feedback.**

Recommendation 9.3: **The Census Bureau should seek ways to expand researcher access to microdata from and about the 2000 census in order to further understanding of census data quality and social science knowledge. Such data files as the 2000 A.C.E. E-sample and P-sample output files, for example, should be deposited with the Bureau's Research Data Centers. To help the Bureau evaluate population coverage and data quality in the 2010 census, the Bureau should seek ways—using the experience with the Panel to Review the 2000 Census as a model—to furnish preliminary data, including microdata, to qualified researchers under arrangements that protect confidentiality.**

Appendixes

Appendix A

Panel Activities and Prior Reports

Formed in response to a 1998 request by the U.S. Census Bureau, the Panel to Review the 2000 Census conducted a variety of activities to carry out its broad charge to review the statistical methods of the 2000 census. In this appendix, we provide additional detail on the meetings and previous publications of the panel.

A.1 LIST OF PANEL MEETINGS, WORKSHOPS, AND TRIPS BY PANEL MEMBERS

The Panel to Review the 2000 Census met 20 times as a whole between its inception in 1998 and the end of 2003. Additional detail on the panel's schedule of meetings is given in Table A.1.

In addition, members and staff of this panel were joined by members of the Panel on Research on Future Census Methods in visiting local and regional offices during various phases of census operations. Visits conducted during census data collection are listed in Table A.2, and visits conducted to monitor the Accuracy and Coverage Evaluation (A.C.E.) Program are listed in Table A.3.

Table A.1 Meetings of the Panel to Review the 2000 Census

Meeting	Date	Format
1	November 9, 1998	Open and closed sessions.
2	January 15, 1999	Open and closed sessions.
3		Panel chair and four members toured the National Processing Center in Jeffersonville, Indiana, with local census personnel.
4	March 19, 1999	Open and closed sessions.
5	June 28, 1999	Open and closed sessions.
6	October 6–7, 1999	Workshop, and open and closed sessions.
7	February 2–3, 2000	Workshop, and open and closed sessions.
8	May 8–9, 2000	Open and closed sessions.
9	October 2–3, 2000	Workshop, and closed session.
10	March 22, 2000	Closed session.
11	June 20–21, 2000	Closed session.
12	October 22–23, 2001	Closed session.
13	July 17–18, 2002	Closed session.
14	October 30–31, 2002	Closed session.
15	March 12, 2003	Open and closed sessions. Open session held jointly with Panel on Research on Future Census Methods.
16	April 25, 2003	Closed session.
17	May 19, 2003	Closed session.
18	July 17, 2003	Closed session.
19	September 15, 2003	Closed session.
20	November 18, 2003	Closed session.

A.2 PUBLICATIONS

A.2.a *The 2000 Census: Interim Assessment*

On October 9, 2001, the panel released its interim report in prepublication format. Titled *The 2000 Census: Interim Assessment*, the interim report assessed the Census Bureau's March 2001 recommendation regarding statistical adjustment of census data for redistricting and reviewed census operations. By design, the interim report did not address the Census Bureau's decision on adjustment for nonredistricting purposes, which was anticipated to occur on or about October 15 (the decision was actually announced on October 17). Subsequently, on November 26, the panel sent a letter report to William Barron, acting director of the Census Bureau. In the letter report, the panel reviewed the new set of evaluations prepared by the Census Bureau in support of its October decision.

In late 2001, these two reports—the letter report and the interim

Table A.2 Site Visits to Regional and Local Census Offices, 2000

Location	Date	Participants
Update/Leave		
Rapid City, SD	March 6	Heather Koball
Special Places		
Madison, WV	March 17	Heather Koball
List/Enumerate		
Remote Maine	April 6–7	Heather Koball
Nonresponse Follow-up		
Rochester, NY	April 20	Constance Citro
Boston/Cape Cod, MA[a]	April 28–29	David Binder[b], Heather Koball
Sampson County, NC	May 15	Allen Schirm[b], Heather Koball
Philadelphia, PA[a]	May 19	Lawrence Brown, Heather Koball
New Orleans, LA	May 26	L. Bruce Petrie, Heather Koball
New York City, NY[a]	May 26	Joseph Salvo[b], Andrew White
Los Angeles, CA[a]	June 2	Robert Bell, Andrew White
Chicago, IL[a]	June 9	Robert Hauser, Heather Koball
Miami, FL	June 12	Benjamin King[b], Heather Koball
Dallas, TX	June 21	L. Bruce Petrie, Heather Koball
Atlanta, GA	July 28	Robert Hauser, Heather Koball
Data Capture Center		
Baltimore, MD	March 30	David Binder[b], William Eddy, Sallie Keller-McNulty[b], Janet Norwood, Joseph Salvo[b], Allen Schirm[b], Michael Cohen, Heather Koball, Andrew White

NOTES: All visits (except to data capture centers) included panel members and staff accompanying census enumerators during their work; the Rochester, NY, visit was to the local census office before follow-up commenced.

[a] Visits included both a local and a regional A.C.E. office.
[b] Member, Panel on Research on Future Census Methods.

report—were issued as a combined volume that retained the title *The 2000 Census: Interim Assessment* (National Research Council, 2001a).

A.2.b LUCA Working Group Report

Jointly commissioned by the Panel to Review the 2000 Census and the Panel on Research on Future Census Methods, the Working Group on LUCA brought together representatives of state and local governments to assess participation in the Census Bureau's Local Update of Census Addresses (LUCA) Program. The Working Group was chaired by Joseph Salvo, director of the New York City Department of City Planning's Population Division and a member of the Panel on Research on Future Census Methods. Members of

Table A.3 Additional Site Visits to Accuracy and Coverage Evaluation Offices, 2000

Location	Date	Participants
Seattle, WA[a]	June 19	Michael Meyer[b]; Andrew White
McAllen, TX	June 27	Heather Koball
San Francisco, CA	June 30	Donald Ylvisaker[b]; Heather Koball
Greenville, MS	July 7	L. Bruce Petrie; Heather Koball
Los Angeles, CA[a]	July 10	Robert Bell; Heather Koball
Detroit, MI[a]	July 14	William Eddy; Andrew White
Newark, NJ, and New York City, NY	July 21	Joseph Salvo[b]; Allen Schirm[b]; Michael Cohen; Heather Koball
Chicago, IL[a]	August 4	Keith Rust[b]; Heather Koball

NOTES: All visits included panel members and staff accompanying A.C.E. interviewers during their work.

[a] Visits included both a local and a regional A.C.E. office.
[b] Member, Panel on Research on Future Census Methods.

the working group (and their local government affiliations) were: Shoreh Elhami (Delaware County, Ohio); Abby Hughes (formerly with Arkadelphia, Arkansas); Terry Jackson (Georgia Department of Community Affairs); Tim Koss (Snohomish County, Washington); and Harry Wolfe (Maricopa Association of Governments, Arizona). Patricia Becker of APB Associates served as consultant to the working group.

The working group found that the interaction between local governments and the Census Bureau in LUCA helped identify important problems in the Master Address File (MAF) and revealed limitations in the Bureau's methods to create and update the MAF. The report—including a detailed chapter on case studies of LUCA implementation in various communities—described strategies that local governments employed to locate and rectify address list gaps in their areas. Overall, the working group concluded that the LUCA program was beneficial for those communities that participated. The group suggested that the Census Bureau continue to build partnerships with local governments and to encourage broader participation by local authorities in any planned continuous address updating effort.

The working group's final report (Working Group on LUCA, 2001) was published and distributed by the panels and the Committee on National Statistics (CNSTAT). It is not sold by The National Academies Press, but it is available on the Internet through the CNSTAT Web site (http://www7.nationalacademies.org/cnstat/1LUCAReport.pdf [9/1/03]).

A.2.c Workshop Proceedings

As part of its work, the panel held three open workshops on topics related to the A.C.E. and possible adjustment of the census counts for population coverage errors. In 2001, the panel issued a proceedings volume for each of these workshop meetings; each volume is an edited transcript of the presentations and discussions.

The first workshop (National Research Council, 2001e) was held October 6, 1999. It considered issues of the A.C.E. design that had not yet been completely worked out by the Census Bureau staff. Topics discussed included methods and issues for determining poststrata for estimation, obtaining the final sample of block clusters from a larger initial sample, and imputing values for missing responses on characteristics needed to define poststrata.

The second workshop (National Research Council, 2001f) was held February 2–3, 2000. It covered the dual-systems estimation process, as planned for the 2000 census, from beginning to end.

The third workshop (National Research Council, 2001g) was held October 2, 2000. It laid out the process the Census Bureau planned to follow in order to reach a decision by March 1, 2001, on whether to adjust the census counts for purposes of congressional redistricting.

A.3 COMMISSIONED PAPER

To supplement its work, the panel commissioned a paper by David Harris, Department of Sociology and Institute for Social Research at the University of Michigan, to review the measurement of race and ethnicity in the 2000 census.

In the paper, Harris (2003) considers the process and method by which the Office of Management and Budget (OMB) revised its Directive Number 15 to create the racial and ethnic classifications for the 2000 census, and examines how the 2000 data can be used by government agencies and researchers. As part of this discussion, Harris looks at the social perspective on race that informed OMB's revisions, and offers a critique of the new classification system and its uses.

Harris concludes that OMB's classification and tabulation guidelines are often somewhat arbitrary and fail to acknowledge that racial classifications emerge from social interactions. Harris recommends that future attempts at data collection, and current users of the 2000 census data, employ a social theory of race. He also suggests that data users look to other data sources to help develop models for classification patterns for multiracial respondents.

A.4 LETTER REPORTS

Although the letter reports issued by the panel are readable via the National Academies Press Web site (http://www.nap.edu), we reprint here two of the panel's three letters that have not previously appeared in any of the bound panel reports. References and acronyms in the letter reports have been minimally edited for consistency with the rest of the volume.

A.4.a May 1999 Letter Report

Dr. Kenneth Prewitt
Director
U.S. Bureau of the Census
Room 2049, Building 3
Washington, DC 20233

Dear Dr. Prewitt:

As part of its charge, the new Panel to Review the 2000 Census offers this letter report on the Census Bureau's plans for the design of the Accuracy and Coverage Evaluation (A.C.E.) survey, a new post-enumeration survey. This survey is needed in light of the recent U.S. Supreme Court ruling regarding the use of the census for reapportionment.

In general, the panel concludes that the A.C.E. design work to date is well considered. It represents good, current practice in both sample design and poststratification design, as well as in the interrelationships between the two. In this letter the panel offers observations and suggestions for the Census Bureau's consideration as the work proceeds to complete the A.C.E. design.

Background

Because it is not possible to count everyone in a census, a post-enumeration survey is an important element of census planning. The survey results are combined with census data to yield an alternative set of estimated counts that are used to evaluate the basic census enumeration and that can be used for other purposes. For 2000, an Integrated Coverage Measurement (ICM) survey had been planned for evaluation and to produce adjusted counts for all uses of the census.[1] The recent U.S. Supreme Court ruling against the use of sampling for reapportionment among the states eliminates the need for a post-enumeration survey that supports direct state estimates, as was originally planned for the ICM survey. (The state allocations of the ICM sample design deviated markedly from a proportional-to-size allocation in order to support direct state estimation. Specifically, the ICM design required a minimum of 300 block clusters in each state.) Alternative approaches are now possible for both sample and poststratification designs for the 2000 A.C.E. survey. As a result, the planned A.C.E. post-enumeration survey will differ in several important respects from the previously planned ICM survey.

Plans for A.C.E. Sample and Poststratification Design

Our understanding of the current plans for the A.C.E. survey is based on information from Census Bureau staff.[2] Building on its work for the previously planned ICM, the Census Bureau will first identify a sample of block clusters containing approximately 2 million housing units and then will independently develop a new list of addresses for those blocks.[3] In a second stage, a sample of block clusters will be drawn from the initial sample to obtain approximately 750,000 housing units, which was the number originally planned for the ICM. (Larger block clusters will not be drawn

[1]See National Research Council (1999b).

[2]See Kostanich et al. (1999)

[3]The use of the term *block cluster* refers to the adjoining of one or more very small blocks to an adjacent block for the purpose of the A.C.E. sample design. Large blocks often form their own block clusters.

in their entirety; they will first be subsampled to obtain sampling units of 30–50 housing units. Because the costs of interviewing are so much greater than the costs of listing addresses, this subsampling approach allows the interviewed housing units to be allocated in a more effective manner.) Finally, in a third stage, a sample of block clusters will be drawn from the second-stage sample to obtain the approximately 300,000 housing units required for the A.C.E. sample. The target of 300,000 housing units for the A.C.E., which may be modified somewhat, will be based on a new set of criteria that are not yet final.

The Census Bureau is considering three strategies for selection of the 300,000 A.C.E. subsample from the 750,000 sample: (1) reducing the sample proportionately in terms of state and other block characteristics from 750,000 to 300,000; (2) reducing the sample by using varying proportions by state; or (3) differentially reducing the sample by retaining a higher proportion of blocks in areas with higher percentages of minorities (based on the 1990 census).[4] These options for selection of the 300,000 A.C.E. housing units from the 750,000 units first selected will be carefully evaluated. The plans include three evaluation criteria for assessing the options: (a) to reduce the estimated coefficients of variation for 51 poststratum groups (related to the 357-cell poststratification design discussed below); (b) to reduce the differences in coefficients of variation for race/ethnicity and tenure groups; and (c) to reduce the coefficients of variation for estimated state totals. (Option (3) above is motivated by criterion (b).) Without going into detail, it is also useful to mention that the Census Bureau has instituted a number of design changes from the 1990 Post-Enumeration Survey for the A.C.E. that will reduce the variation in sampling weights for blocks, which will reduce the sensitivity of the final estimates to results for individual blocks. This represents a key improvement in comparison with the 1990 design.

The current plan to produce poststrata involves modification of the 357-cell poststratification design suggested for use in 1990-based intercensal estimation. Current modifications under consideration by the Census Bureau include expansion of the geographic stratification for non-Hispanic whites from four regions to nine census

[4]The Census Bureau is aware that mixtures of strategies (2) and (3) are also possible, although such mixtures are not currently being considered.

divisions, adding a race/ethnicity group, changing the definition of the urbanicity variable, and adding new poststratification factors, such as mail return rate at the block level. Logistic regression, modeling inclusion in the 1990 census, is being used to help identify new variables that might be useful, as well as to provide a hierarchy of the current poststratification factors that will be used to guide collapsing of cells if that is needed. (In comparison, the analysis that generated the 357-cell poststratification was based on indirect measures of census undercoverage, such as the census substitution rate.)

The Census Bureau plan demonstrates awareness of the interaction of its modification of the 750,000 housing unit sample design with its modification of the 357 poststrata design. (On the most basic level, the sample size allocated to each poststratum determines the variance of its estimate.) The plan also makes clear that even though much of the information used to support this modification process must be based on the 1990 census, it is important that the ultimate design for the A.C.E. survey (and any associated estimation) allows for plausible departures from the 1990 findings. For example, significant differences between the 1990 and 2000 censuses could stem from the change in the surrounding block search for matches, the planned change in the treatment of A.C.E. movers, or changes in patterns and overall levels of household response.

Observations and Comments

Sample Design to Select the 300,000 Housing Units

Because of the need to keep the A.C.E. on schedule by initiating resource allocations that support the independent listing of the 2 million addresses relatively soon, as well as the need to avoid development and testing of new computer software, the Census Bureau has decided to subsample the 300,000 A.C.E. housing units from the 750,000 housing units of the previously planned ICM design. The panel agrees that operational considerations support this decision.

The cost of the constraint of selecting the 300,000 A.C.E. housing units from the 750,000 ICM housing units, in comparison with an unconstrained selection of 300,000 housing units, is modest. While the constrained selection will likely result in estimates with somewhat higher variances, the panel believes that careful selection of the subsample can limit the increase in variance so that it will not be consequential. (By careful selection, the panel means use of the suggested approaches of the Census Bureau, or new or hybrid techniques, to identify a method that best satisfies the criteria listed above.) This judgment by the panel, although not based on a specific analysis by itself or the Census Bureau, takes into account the fact that a large fraction of the 750,000 housing units of the ICM design are selected according to criteria very similar to those proposed for the A.C.E. design.

In addition, the panel notes that the removal of the requirement for direct state estimates permits a substantial reduction in sample size from the 750,000 ICM design in sparsely populated states, for which A.C.E. estimates can now pool information across states. As a result, the A.C.E. design could result in estimates with comparable reliability to that of the previously planned, much larger ICM design.

Given the freedom to use estimates that borrow strength across states, the final A.C.E. sample should reduce the amount of sampling within less populous states from that for the preliminary sample of 750,000 housing units. However, there is a statistical basis either for retaining a minimum A.C.E. sample in each state, or what is nearly equivalent, for retaining a sample to support an A.C.E. estimate with a minimum coefficient of variation. The estimation now planned for the A.C.E. survey assumes that there will be no important state effects on poststratum undercoverage factors. In evaluating the quality of A.C.E. estimates, it will be important to validate this assumption, which can only be done for each state if the direct state estimates are of sufficient quality to support the comparison, acknowledging that for some of these analyses one might pool data for similar, neighboring states. (Identification of significant state effects would not necessarily invalidate use of the A.C.E. estimates for various purposes but would be used as part of an overall assessment of their quality.)

This validation could take many forms, and it is, therefore, difficult to specify the precise sample size or coefficient of variation needed. We offer one approach the Census Bureau should examine for assessing the adequacy of either type of standard. Using the criteria for evaluating alternative subsample designs (i.e., the estimated coefficients of variation for 51 poststratum groups, the differences in coefficients of variation for race/ethnicity and tenure groups, and the coefficients of variation for state totals), the Census Bureau should try out various state minima sample sizes to determine their effects on the outputs. It is possible that a moderately sized state minimum sample can be obtained without affecting the above coefficients of variation to any important extent. There are a variety of ways in which the assumption of the lack of residual state effects after accounting for poststratum differences could be assessed, including regression methods. We encourage the Census Bureau to consider this important analytic issue early and provide plans for addressing it before the survey design is final.

The panel makes one additional point on state minima. The state minima will support direct state estimates that will be fairly reliable for many states. The Census Bureau should consider using the direct state estimates not only for validation, but also in estimation—in case of a failure of the assumption that there will be no important state effects on undercoverage factors. Specifically, the Census Bureau should examine the feasibility of combining the currently planned A.C.E. estimates at the state level with the direct state estimates, using estimated mean-squared error to evaluate the performance of such a combined estimate in comparison with the currently planned estimates. We understand that the necessity of prespecification of census procedures requires that the Census Bureau formulate an estimation strategy prior to the census, which adds urgency to this issue.

Finally, the panel has two suggestions with respect to the criteria used for assessing the A.C.E. sample design. First, there should be an assessment of the quality of the estimates for geographic areas at some level of aggregation below that of states, as deemed appropriate by the Census Bureau. (This criterion is also important for evaluating the A.C.E. poststratification design, discussed below.) Second, the importance of equalizing the coefficients of variation for different poststrata depends on how estimates for specific poststrata with higher coefficients of variation adversely affect the variance of estimated counts for certain areas. Coefficients of variation for poststrata that do not have much effect have less need to be controlled, assuming that the estimates for these poststrata do not have other uses.

Poststratification Plans

The 1990 census adjusted counts used 1,392 poststrata, but post-production analysis for calculating adjusted counts for intercensal purposes resulted in the use of 357 poststrata. The panel believes that the use of these 357 poststrata (and the hierarchy for collapsing poststratification cells) was a reasonable design for 1990, and that, in turn, the 1990 design is a good starting point in determining the poststrata to be used in the 2000 A.C.E. The Census Bureau is considering four types of modifications to the 357 poststrata design, although it has not yet set the criteria for evaluating various poststratification designs. Logistic regression will be used to identify new variables and interactions of existing variables that might be added to the poststratification. Finer poststrata have the advantage of greater within-cell homogeneity, potentially producing better estimates when carried down to lower levels of geographic aggregation. Some gains with respect to the important problem of correlation bias might also occur. However, stratifying on factors that are not related to the undercount will generally decrease the precision of undercount adjustments. The tradeoff between within-cell homogeneity and precision needs to be assessed to determine whether certain cells should be collapsed and whether additional variables should be used.

It is also important to examine the effects of various attempts at poststratification on the quality of substate estimates, especially since certain demographic groups are more subject to undercoverage, and so substate areas with a high percentage of these groups will have estimates with higher variances. (This argument is based on the fact that, as in the binomial situation, the mean and the variance of estimated undercounts are typically positively related.) We believe it is extremely important that analyses at substate levels of aggregation be conducted to inform both the sample design and the poststratification scheme. Furthermore, this issue needs to be studied simultaneously with that of the effect of the design and poststratification on the poststratum estimates. The fact that analysis of substate areas appears in both sample design and poststratification design is an indication of the important interaction between these two design elements and justifies the need for studies of them to be carried out simultaneously. The panel encourages the Census Bureau to work on them at the same time.

The panel notes that the decision to use a modification of the 357-strata system from 1990 for the A.C.E. poststratification design will probably not permit many checks against estimates from demographic analysis that use direct estimates from A.C.E. This limitation may increase the difficulty of identifying the precise source of large discrepancies in these comparisons. However, the panel does not view this as a reason not to proceed, since the precision of direct estimates at the finest level of detail of poststratification (using 1,392 strata in this context) could make such comparisons more difficult to interpret, and the estimates from demographic analysis are not extremely useful for this purpose (except for blacks, and then only nationally).

As work on both the sample design and poststratification design progresses, the Census Bureau should not rely entirely on information from the 1990 census: substantial differences might occur between the 1990 and the 2000 censuses that would lead to either a sample design or a poststratification design that was optimized for 1990 but that might not perform as well in 2000. Instead, the Census Bureau should use a sample design that moves toward a more equal probability design than 1990 information would suggest. Similarly, the Census Bureau, using whatever information is available since 1990 on factors related to census undercoverage, should develop a poststratification design that will perform well for modest departures from 1990.

Finally, when considering criteria for both sample design and poststrata, it is important to keep in mind that the goal of the census is to provide estimated counts for geographic areas as well as for demographic groups. Since the use of equal coefficients of variation for poststrata will not adequately balance these competing demands, the Census Bureau will need to give further attention to this difficult issue. The balancing of competing goals is not only a poststratification issue, but also a sample design issue. For example, if block clusters that contain large proportions of a specific demographic group are substantially underrepresented in the A.C.E. sample, the performance of the estimates for some areas could be affected.

Documentation

Given the importance of key decisions and input values for the A.C.E. design, it is important that they be documented. In particular, the Census Bureau should produce an accessible document in print or in electronic form that (1) gives the planning values for state-level, substate level, and poststratum level variances resulting from the decisions for the sample and poststratification designs and (2) provides the sampling weights used in the A.C.E. selection of block clusters.

Summary

From its review of the Census Bureau's current plans for design of the A.C.E. survey, the panel offers three general comments:

- The panel concludes that the general nature of the Census Bureau's work on the A.C.E. design represents good, current practice in sample design and poststratification design and their interactions.
- The panel recognizes that operational constraints make it necessary for the Census Bureau to subsample the A.C.E. from the previously planned ICM sample. The subsampling, if done properly, should not affect the quality of the resulting design if compared with one that sampled 300,000 housing units that were not a subset of the 750,000 housing units previously planned for the ICM.
- The panel believes that removal of the constraint to produce direct state estimates justifies the substantial reduction in the A.C.E. sample size from the ICM sample size. The planned A.C.E. could result in estimates with comparable reliability to that of the larger ICM design.

The panel offers three suggestions for the Census Bureau as it works to finalize the A.C.E. design, some of which the Census Bureau is already considering: (1) a method for examining how large a state minimum sample to retain; (2) some modifications in the criteria used to evaluate the A.C.E. sample design and poststratification, namely, lower priority for coefficients of variation for excessively detailed poststrata and more attention to coefficients of variation for substate areas; and (3) a possible change in the A.C.E. estimation procedure, involving use of direct state estimates in combination with the currently planned estimates. In addition, the Census Bureau should fully document key decisions for the A.C.E. design.

The panel looks forward to continuing to review the A.C.E. design and estimation as the Census Bureau's plans are further developed. The panel is especially interested in the evolving plans for poststratification design, including the use of logistic regression to identify additional poststratification factors; plans for the treatment of movers in A.C.E.; and the treatment of nonresponse as it relates to unresolved matches in A.C.E. estimation. In addition, after data have been collected, the panel is interested in the assessment of the effect of nonsampling error on A.C.E. estimation and the overall evaluation criteria used to assess the quality of A.C.E. estimates.

We conclude by commending you and your staff for the openness you have shown and your willingness to discuss the A.C.E. survey and other aspects of the planning for the 2000 census.

A.4.b November 2000 Letter Report

Dr. Kenneth Prewitt
Director
U.S. Census Bureau
Room 2049, Building 3
Washington, DC 20233

Dear Dr. Prewitt:

This letter comments on the Census Bureau's plans for deciding whether to release adjusted estimates of the population of states and substate areas from the 2000 census. That decision will be made in March 2001 so that the Bureau can meet its April 1 deadline under Public Law 94-171 for providing data to the states for redrawing congressional districts.

This is the second letter report of our Panel to Review the 2000 Census. The first report, issued in May 1999, commented on the Bureau's plans for the design of the Accuracy and Coverage Evaluation (A.C.E.) Survey. The results of that survey will be used along with other information to evaluate the census, and if deemed beneficial, will be combined with census data, using dual-systems estimation methods, to yield adjusted estimates of the population. In this letter report we comment on the Bureau's plans for evaluating the census data and the A.C.E. data and for deciding whether to release both the census population estimates, unadjusted for coverage errors, and the adjusted estimates.

The Census Bureau presented its plans for the evaluation and decision process that will lead up to the adjustment decision at an open panel workshop on October 2 in Washington, D.C. Bureau staff provided 16 papers that are now in draft form. The papers contain table shells that will be filled in, as data become available over the next few months, with information that is important for the evaluation and decision. The papers cover a variety of topics, including: overall census and A.C.E. quality indicators, quality of census processes, demographic analysis results, person interviewing, person matching and follow-up, missing data, variance estimates by size of geographic area, correlation bias, synthetic assumptions, and other topics. Bureau staff discussed the draft papers and responded to questions and comments from panel members and other workshop participants.

The panel commends the Census Bureau for the openness and thoroughness with which it has informed the professional community of the kinds of evaluations that it plans to conduct of the census and A.C.E. data prior to March 2001. The papers presented at the panel workshop provide evidence of the hard work and professional competence of Census Bureau staff in specifying a series of evaluations that can inform the adjustment decision.

The panel recognizes the difficult task faced by the Census Bureau in evaluating the census and A.C.E. data by the time that it must provide congressional redistricting data to the states. Since it will not be possible for the Bureau to complete all possible analyses by March 2001, it will have to act on the basis of analyses that can be conducted before that time. In view of that constraint, the panel concludes that the set of papers presented to the workshop in draft form reflect competent, professional work to develop an informative set of evaluations for the short term. The planned analyses appear to cover all of the evaluations that can reasonably be expected to be completed within the time available. Furthermore, they appear to be sufficiently comprehensive that they will likely provide support for a reasonably confident decision on adjustment in March.

However, since the numbers themselves, which are, of course, critical to the evaluation process, are not yet available, it is not possible at this time to comment on what the adjustment decision should be nor to conclude definitively that the planned short-term evaluations will be adequate to support the decision. Such commentary will be possible only after the Bureau has completed its work and has provided the supporting data to the professional community.

In addition to the evaluations that are planned specifically to inform the adjustment decision in March 2001, the Census Bureau has a longer term evaluation agenda that includes projects to assess all major systems used in the 2000 census and many aspects of census data quality. That agenda, which will take several years to complete, also includes evaluations that are related to the A.C.E., such as a study of error in the process by which census enumerations and A.C.E. enumerations are matched in A.C.E. blocks. The panel urges the Census Bureau to identify those longer term studies that are likely to provide useful information with which to evaluate the adjustment decision, to give priority to these evaluations, and to provide detailed plans for these evaluations and a schedule that allows for completing them as soon after March 2001 as possible. Users of census data need to know when and what kinds of evaluations will be available in the longer term, just as they have been made aware of the Bureau's plans for the evaluations to be completed by March.

The short-term evaluations that are planned to inform the adjustment decision in March will provide voluminous, complex data and analyses on a range of aspects of the census and the A.C.E. Review and assessment of this necessarily complex set of information will present a challenge for the Census Bureau's Executive Steering Committee for A.C.E. Policy (ESCAP), which is charged to recommend an adjustment decision to the Bureau director, as well as for the professional community and stakeholders. The panel believes it would be useful for all concerned parties for the Census Bureau to develop a summary tabular presentation of the factors affecting its decision.

One of the Census Bureau's 16 papers, "Data and Analysis to Inform the ESCAP Recommendation," is intended to summarize the analyses and the approach; it should usefully serve this purpose.

However, that paper is itself lengthy, and we encourage the Bureau to provide a summary table in addition that focuses on key pieces of evidence in the decision-making process. The summary table would indicate for each piece of evidence such information as the nature of the empirical findings, the type(s) of analysis that were the basis for the findings, the strengths and weaknesses of the analysis, and the implications of the findings for the adjustment decision. The summary should also note relevant types of evidence that are not yet available.

We understand that no simple formula will lead from the available data to a recommendation whether or not to issue adjusted census counts by March 2001 for use in congressional redistricting. A sound recommendation will ultimately rest on professional judgment informed by the available scientific evidence. However, we believe that preparation of a summary presentation of key evidence will assist ESCAP to integrate what will necessarily be a large volume of complex information, some of which may be conflicting, in reaching an adjustment decision. Such a presentation will also assist the professional community and stakeholders to understand the basis for the decision.

In summary, the panel appreciates the openness and professionalism with which you and your staff have undertaken to provide an extensive set of data for use in determining the quality of the enumerated census and, alternatively, of any adjustments that might be made to improve the census data. In furthering your efforts, we make the following two suggestions that were discussed above:

(1) The Census Bureau should prioritize its plans for previously planned long-range evaluation studies with a view toward completing those evaluations most directly relevant to the adjustment decision as soon as possible. While we understand that some studies, especially those involving new data collection, cannot be concluded before the statutory requirements for release of the data, we believe it would be useful for you to produce a schedule that allows for completing them soon after March 2001. The Bureau should make public its plans for these evaluations and their release.

(2) Although the Census Bureau has developed plans for comprehensive review of large bodies of data for use in its decision on

whether to release adjusted census counts, the panel believes that it would also be useful to have a summary presentation of key evidence. In particular, a summary table listing each piece of evidence and how it relates to the adjustment decision would be helpful to all parties concerned.

We conclude by thanking you and your staff for your cooperation in providing information for our workshop on the forthcoming decision process.

A.4.c November 2001 Letter Report

The panel issued a third letter report in November 2001, regarding the Census Bureau's October 2001 decision not to adjust 2000 census data for such purposes as fund allocation. This letter is not reproduced here since it was printed in full in our interim report (National Research Council, 2001a).

Appendix B

Questionnaire Items on the 2000 and 1990 Censuses and Census 2000 Supplementary Survey

2000 Census Item	Asked in 1990 Census?	Asked in Census 2000 Supplementary Survey?
Short Form		
Age	Yes	Yes
Date of birth (month, day, year)	Yes (year of birth only)	Yes
Sex	Yes	Yes
Hispanic origin (question preceded race)	Yes (question followed race)	Yes (question preceded race)
Race (option for multiple races)	Yes (no option for multiple races)	Yes (option for multiple races)
Relationship to household reference person	Yes	Yes (less detail)
Housing tenure (own or rent)	Yes	Yes

2000 Census Item	Asked in 1990 Census?	Asked in Census 2000 Supplementary Survey?
Long Form		
Marital status	Yes (short form)	Yes
Whether and how long responsible for grandchildren in home	No	Yes
School attendance/education attainment	Yes	Yes
Public or private school (for those enrolled)	Yes	Yes
Place of birth	Yes	Yes
Citizenship	Yes	Yes
Year of immigration	Yes	Yes
Language spoken at home	Yes	Yes
How well English spoken	Yes	Yes
Ancestry	Yes	Yes
Veteran status/period of service	Yes	Yes
Years of military service	Yes	Yes
Work disability	Yes	Yes
Disabled for going outside the home alone	Yes	Yes
Disabled for taking care of personal needs	Yes	Yes
Other disabilities (involving eyes, ears, cognition, physical activity)	No	Yes
Place of residence 5 years ago	Yes	Yes (1 year ago)
Employment status (current)	Yes	Yes
Occupation	Yes	Yes
Industry	Yes	Yes
Class of worker	Yes	Yes
Place of work	Yes	Yes
Means of transportation for work	Yes	Yes
Commuting time	Yes	Yes
Carpooling	Yes	Yes
Year last worked	Yes	Yes
Weeks worked in preceding year	Yes	Yes
Hours worked per week in preceding year	Yes	Yes
Income from earnings (last year)	Yes	Yes (last 12 months)
Income from self-employment (net) (last year)	Yes (farm and nonfarm separately)	Yes (last 12 months)

2000 Census Item	Asked in 1990 Census?	Asked in Census 2000 Supplementary Survey?
Long Form		
Income from Social Security (last year)	Yes	Yes (last 12 months)
Income from Supplemental Security Income (last year)	No (included in public assistance)	Yes (last 12 months)
Income from public assistance (last year)	Yes	Yes (last 12 months)
Income from interest, dividends, rent (last year)	Yes	Yes (last 12 months)
Income from pensions (last year)	Yes	Yes (last 12 months)
All other income (last year)	Yes	Yes (last 12 months)
Total income (last year)	Yes	Yes (last 12 months)
Type of property (e.g., whether includes a business; for single-family homes only)	Yes	Yes
Value	Yes (short form)	Yes
Rent (monthly contract rent)	Yes (short form)	Yes
Does rent include any meals?	Yes (short form)	Yes
Condominium status	Yes	Yes
Farm residence/sales (single-family homes only)	Yes (all housing units)	Yes (single-family homes)
Number of units in structure	Yes (short form)	Yes
Year structure built	Yes	Yes
Year reference person moved into unit	Yes	Yes
Number of rooms	Yes (short form)	Yes
Number of bedrooms	Yes	Yes
Complete plumbing facilities	Yes	Yes
Complete kitchen facilities	Yes	Yes
Telephone available	Yes	Yes
Type of heating fuel	Yes	Yes
Electricity costs (yearly)	Yes	Yes (last month)
Gas costs (yearly)	Yes	Yes (last month)
Water costs (yearly)	Yes	Yes
Oil, coal, etc. costs (yearly)	Yes	Yes
Mortgage payment (and whether includes taxes, insurance)	Yes	Yes
Homeowners insurance	Yes	Yes
Property taxes (owners only)	Yes	Yes
Payment for second mortgage/home equity loans	Yes	Yes

2000 Census Item	Asked in 1990 Census?	Asked in Census 2000 Supplementary Survey?
Long Form		
Condominium/mobile home fee	Yes	Yes
Number automobiles/vans/trucks	Yes	Yes

NOTES: The 2000 census, the 1990 census, and the Census 2000 Supplementary Survey provided room on the mailback questionnaire for characteristics of up to six, seven, and five household members, respectively. Section I.3 provides additional detail on the Census 2000 Supplementary Survey. Questionnaires should be consulted for precise question wording, which may differ between censuses.

SOURCE: Citro (2000d:Table 1) for 2000 and 1990; http://www.census.gov/c2ss for the Census 2000 Supplementary Survey.

Appendix C

Census Operations

This appendix provides additional detail on the operations of the 2000 census, noting differences from 1990 census procedures. It covers five topics:

- the Master Address File (MAF) (including local review and internal checks for duplicate addresses);
- questionnaire delivery and mail return (including redesign of mailings and materials and multiple response modes);
- field follow-up (including nonresponse follow-up, NRFU, and coverage improvement follow-up, CIFU);
- outreach efforts; and
- data processing (including data capture, coverage edit and telephone follow-up, unduplication of households and people, and other processing).

Two important parts of census data processing—editing and imputation—are described in greater detail in separate appendices for the basic (complete-count) data (Appendix G) and the long-form-sample data (Appendix H). General theory and approaches to item imputation are discussed in Appendix F.

C.1 MASTER ADDRESS FILE

The 2000 census was conducted primarily by mailing or delivering questionnaires to addresses on a computerized mailing list—the MAF—and asking residents to fill out the questionnaires and mail them back.[1] The Census Bureau first used mailout/mailback techniques with an address list in the 1970 census,[2] but the procedures to develop the 2000 MAF differed in several important respects from those used in past censuses (see Working Group on LUCA, 2001; Owens, 2000). The major difference from 1990 was that the 2000 MAF was constructed using more sources.

C.1.a Initial Development

The Census Bureau used somewhat different procedures to develop the MAF for areas believed to have predominantly city-style mailing addresses (house number and street) than for areas believed to have predominantly rural route and post office box mailing addresses (see Box C.1). City-style areas were those inside the "blue line," and non-city-style areas were those outside the "blue line."[3] For areas inside the blue line, the Bureau expected to have U.S. Postal Service carriers deliver questionnaires to most addresses on the list; for areas outside the blue line, the Bureau expected to use its own field workers to deliver questionnaires.

For remote rural areas, which have less than 1 percent of the population, Census Bureau enumerators developed the address list concurrently with enumerating households in person. For special places in which people live in nonresidential settings, such as college dormitories, prisons, nursing homes and other group quarters, the Bureau used a variety of sources to develop an address list.

[1]The Census Bureau refers to the version of the MAF that was used in the census as the Decennial Master Address File or DMAF. It is an extract of the full MAF, which includes business as well as residential addresses. Use of the term *MAF* in this report refers to the DMAF.

[2]Unaddressed short-form questionnaires were delivered by the U.S. Postal Service in the 1960 census to 80 percent of households, but residents were to hold the questionnaires for enumerators to pick up. At every fourth household, enumerators left a long-form questionnaire, which respondents were to fill out and mail back (National Research Council, 1995b:189).

[3]The blue line was a late-1997 Census Bureau demarcation.

Inside the "Blue Line"

As the starting point for the MAF for city-style areas inside the blue line, the Census Bureau took the 1990 census address list for these areas and updated it from the Delivery Sequence File (DSF) of the Postal Service. The DSF contains a listing of addresses to which mail is delivered, ordered by carrier routes. It is updated regularly. The Census Address List Improvement Act of 1994 (P.L. 103-430) allowed the Postal Service to share the DSF with the Bureau.

Although not part of its original plan, the Bureau determined that a complete field check of the city-style list should be conducted, which was done in a block canvass operation for all mailout/mailback areas conducted in January–May 1999. The reason for the complete block canvass was the determination that the DSF was not as accurate or as up to date in all areas as needed for the MAF. The Bureau also provided an opportunity for local review in 1998 and 1999 (see Section C.1.b). Approximately 101 million addresses were included in the MAF for areas inside the blue line at the time when questionnaires were labeled and prepared for mailing in July 1999. The Postal Service conducted an intensive check of the DSF in early 2000, and updates were made to the MAF based on that check prior to questionnaire delivery.

Outside the "Blue Line"

To develop the MAF for non-city-style areas, the Bureau first conducted a complete address listing operation in July 1998–February 1999. The 1990 list was not used. There was also a local review program for areas outside the blue line in 1999. Approximately 21 million addresses were included on the MAF for areas outside the blue line at the time when questionnaires were labeled and prepared for delivery. Census enumerators further updated the MAF in these areas when they delivered questionnaires in February–March 2000.

C.1.b Local Review

The Census Address List Improvement Act of 1994, which allowed the Postal Service to share the DSF with the Census Bureau, also permitted the Census Bureau to invite local governments to review the MAF for their areas and provide additions, deletions, and

Box C.1 Basic Steps to Develop the Master Address File Prior to Census Day, 2000 and 1990

2000 CENSUS MASTER ADDRESS FILE

City-Style Areas (mailout/mailback areas inside the "blue line")

1. Start with the 1990 Census Address Control File.
2. Refresh the 1990 list with periodic updates of the U.S. Postal Service Delivery Sequence File.
3. Conduct complete block canvass in the field in January–May 1999 (not in original plans).
4. Provide opportunity for counties, minor civil divisions, places, and tribal governments to review the MAF for their areas in the Local Update of Census Addresses Program, called LUCA 98. LUCA 98 spanned February 1998–March 2000; it included:
   a) local review of initial MAF and census maps;
   b) Census Bureau verification of address changes provided by localities (reconciliation);
   c) local review of feedback/final determination materials from the Bureau; and
   d) review of local appeals by Census Address List Appeals Office in the U.S. Office of Management and Budget.
5. Provide opportunity for localities to review the address list for dormitories, nursing homes, and other special places in December 1999–April 2000 (Special Places LUCA).
6. Incorporate updates from Postal Service's final intensive check of Delivery Sequence File prior to questionnaire delivery.
7. Provide opportunity for localities to supply addresses for newly constructed housing units in January–March 2000 to be enumerated in summer 2000 (New Construction LUCA).

Non-City-Style Areas (update/leave areas outside the "blue line")

1. Conduct an address list creation operation in July 1998–February 1999.
2. Provide opportunity for localities to review the list in the LUCA 99 Program. LUCA 99 spanned July 1998–March 2000. It was similar to LUCA 98 except that localities were asked to challenge housing unit counts for blocks in their initial review. They had to challenge and provide evidence for specific addresses in the appeals phase.
3. Provide opportunity for localities to review the address list for special places.
4. Instruct Census Bureau enumerators to update the MAF when they drop off questionnaires in February–March 2000.

Box C.1 (continued)

1990 ADDRESS CONTROL FILE

City-Style Areas

1. Purchase lists from two vendors; supplement with field listing operation.
2. Recheck the vendor-supplied lists in 1989 in a complete block canvass.
3. Provide the opportunity for localities to review housing unit counts by block in summer 1989.
4. Have the Postal Service conduct several checks in 1988–1990.

Non-City-Style Areas

1. Conduct an address list creation operation (called prelisting) in fall 1989.
2. Instruct Census Bureau enumerators to update the list when they drop off questionnaires.

corrections to the Bureau (Working Group on LUCA, 2001).[4] The Local Update of Census Addresses (LUCA) Program—covering counties, places, and minor civil divisions, over 39,000 jurisdictions in all—was conducted separately in areas inside the blue line (LUCA 98) and areas outside the blue line (LUCA 99). There was also a Special Places LUCA Program.

LUCA required participating local governments to sign a pledge to treat the address list as confidential. The program involved several steps of local review, field verification by the Bureau, and appeal to the U.S. Office of Management and Budget when localities disagreed with the Bureau's decision to reject local changes to the MAF. Due to time constraints, some planned LUCA operations were combined and rescheduled (see Chapter 4, Table 4.2). In response to local concerns, a New Construction LUCA Program was added to give localities inside the blue line an opportunity during January–March 2000 to identify newly constructed housing units. Addresses identified in the program were not mailed questionnaires; instead, they were visited by enumerators during the coverage improvement follow-up operation in summer 2000.

Of the 39,051 jurisdictions that were eligible for either or both LUCA 98 or LUCA 99, it is estimated that 25 percent participated fully in one or both programs by informing the Census Bureau of

[4]Local review procedures were used in the 1980 and 1990 censuses, but localities were not permitted to examine the list of individual addresses for their areas.

needed changes to the address list for their area (Working Group on LUCA, 2001:Ch.2). Participation varied by such characteristics as geographic region of the country, population size of jurisdiction, type of government, and city-style or non-city-style area (see Chapter 4, Table 4.4).

C.1.c Further Development of MAF

MAF was a dynamic file during the operation of the census. Not only were addresses added from each stage of census field operations, they were deleted in an effort to minimize duplicate and erroneous entries. In total, the Census Bureau estimates that about 4 million addresses were added to the MAF during census field operations—2.3 million addresses during questionnaire delivery in update/leave, update/enumerate, and list/enumerate areas (see Section C.2) and 1.7 million addresses during follow-up. About 10.4 million addresses were removed as duplicative of other addresses or nonexistent. About 5 million of these addresses were removed on the basis of two internal consistency checks, one of which was planned and the other of which was designed and implemented while the census data were being processed; the remaining addresses were deleted on the basis of field operations (see Section C.3). Whether the combination of internal checks and field checks reduced duplicate and erroneous addresses to a minimum or went too far or not far enough is a matter for evaluation (see Section 4–E.2). The final number of occupied and vacant housing units counted in 2000 was 115.9 million (Farber, 2001a:Tables 1, 2).

C.1.d Internal Checks for Duplicates

Reducing the NRFU Workload

In April 2000 the Census Bureau conducted an internal consistency check of the MAF prior to the beginning of nonresponse follow-up in order to remove from the NRFU workload as many addresses as possible that could clearly be identified as duplicative or nonexistent (Miskura, 2000a). At the conclusion of this operation, 3.6 million addresses were dropped or merged with another MAF address.

One source of potential duplicates and errors came about because LUCA—essentially, a new, untested program—did not run as smoothly as intended (Working Group on LUCA, 2001). Because of delays in providing materials to local governments to review, the Census Bureau agreed to include every address provided by a LUCA participant on the MAF that was used to label questionnaires in July 1999, even when there had not been time to verify the address in the field. LUCA-supplied addresses that the Bureau believed likely did not exist, based on field checks after July, were flagged. Processing specifications were developed to delete many of these addresses and other addresses of doubtful existence when no questionnaire was returned for them. In all, 2.5 million addresses that the Bureau had reason to believe did not exist were deleted from the MAF prior to nonresponse follow-up.

Also as part of this review, the Bureau attempted to identify duplicate addresses originating from LUCA or other sources. About 1.1 million addresses were merged with another address on the MAF when the addresses appeared to be exact duplicates. Follow-up was conducted either only for the one (merged) address or not at all if a questionnaire had been received for that address.[5]

Unduplication and Late Additions

Another important set of MAF internal checks, not previously planned, was put into place in summer 2000. From evaluations of MAF housing unit counts during January–June 2000 against estimates prepared from other sources, such as building permits, the Census Bureau determined that there were likely still a sizable number of duplicate addresses on the MAF (West and Robinson, 2001). Field verification carried out in June 2000 in a small number of localities substantiated this conclusion (Nash, 2000).

Consequently, the Bureau mounted a special operation to identify duplicate addresses and associated duplicate census returns to remove them from the MAF and the census. Software was written for this operation to match addresses and person records to identify potential duplicates. The flagged records were deleted from the

[5]If questionnaires were received for two addresses that were deemed to be exact duplicates, the Primary Selection Algorithm checked for duplicate enumeration and determined the census household (see Section C.5.c).

census file of valid, completed returns and further examined. After examination, it was decided that a portion of the potential duplicates were likely valid returns for addresses not already in the census, and they were restored to the census file (late additions). At the conclusion of the operation, 1.4 million housing units and all 3.6 million people in those units were permanently deleted from the census file, from a total of 2.4 million housing units and all 6.0 million people in those units that had been initially flagged as potential duplicates (Miskura, 2000b).

C.1.e Comparison: Address List Development in 1990

The procedures used to develop the 1990 Address Control File (ACF) differed in important respects from those used to develop the 2000 MAF (see Box C.1). Overall, the Census Bureau used fewer sources in developing the 1990 ACF than it used for the 2000 MAF; also, the 1990 local review operation was considerably less extensive than the 2000 LUCA Program (see National Research Council, 1995b:App.B).

For 1990 in areas with city-style addresses, the Census Bureau made no use of the 1980 census address list or the Postal Service DSF. Instead, the starting point for the ACF was two files of lists purchased from vendors, supplemented by a field listing operation carried out by census field staff in summer 1988 (precanvass). The Postal Service performed several reviews of the list in 1988–1990; Bureau staff also checked the part of the ACF that derived from commercial lists in a block canvass in 1989. Governmental jurisdictions in the city-style areas were given an opportunity for review in summer 1989; however, they could not review specific addresses but only counts of addresses at the block level. About 16 percent of eligible local governments responded, adding about 400,000 housing units to the ACF (Bureau of the Census, 1993:6-44). By comparison, twice as many eligible governments—36 percent—participated in the LUCA 98 Program in city-style areas.

In areas with non-city-style addresses, the development of the 1990 address list was similar to that in 2000; census field staff conducted an address listing operation in fall 1989. Census enumerators also checked the list in March 1990 when they delivered questionnaires in the areas in which the update/leave technique (new for the 1990 census) was used. However, there was no precensus local review program for the ACF in these areas.

C.2 QUESTIONNAIRE DELIVERY AND MAIL RETURN

The 2000 census, like the 1980 and 1990 censuses, was conducted primarily by delivering questionnaires to households and asking them to mail back a completed form. Procedures differed somewhat depending on such factors as type of addresses in an area and accessibility; in all, there were nine types of enumeration areas. Box C.2 provides brief descriptions of the nine types in 2000.

The two largest types of enumeration areas were: (1) mailout/mailback, covering almost 82 percent of the population, in which Postal Service carriers delivered questionnaires, and (2) update/leave/mailback (usually termed update/leave), covering almost 17 percent of the population, in which Census Bureau field staff delivered questionnaires and updated the MAF at the same time. These two types, together with small numbers of addresses in areas (6), (7), and (9), comprised the mailback universe, covering about 99 percent of the household population (calculated from Baumgardner et al., 2001). The remaining 1 percent of the household population was counted by census enumerators (see areas (3), (4), (5), and (8) in Box C.2). Separate enumeration procedures were used for such special populations as homeless people, residents of group quarters, and transients (see Citro, 2000c).

Approaches to boost mail response were to redesign the questionnaire and mailing package, adapt enumeration procedures to special situations (the reason for having nine types of enumeration areas), and allow multiple modes for response. Advertising and outreach efforts were also expanded from 1990 (see Section C.4).

The final mail response rate in 2000 (67 percent) was slightly higher than the rate in 1990 (65 percent); it was also considerably higher than the rate that was budgeted (61 percent), which reduced the burden of field follow-up. The mail *return* rate in 2000 (78 percent) was higher than the rate in 1990 (75 percent). This rate is a more refined measure of public cooperation than the mail response rate, which includes vacant and nonresidential addresses in the denominator in addition to occupied housing units (see Chapter 4, Box 4.1).

Box C.2 Types of Enumeration Areas (TEAs)

(1) **Mailout/mailback** In areas with predominantly city-style addresses (inside the blue line), U.S. Postal Service carriers delivered an address-labeled advance letter to every housing unit on the MAF the week of March 6. In mid-March the carriers delivered address-labeled questionnaires, followed 2 weeks later by a reminder postcard. Households were to fill out the questionnaire and mail it back.

(2) **Update/leave** In areas outside the blue line in which there were many rural route and post office box addresses that could not be tied to a specific location, census enumerators dropped off address-labeled questionnaires to housing units in their assignment areas. At the same time, they checked the address list and updated it to include new units not on the list, noting for each its location on a map (map spot), so that follow-up enumerators could find units that did not mail back a questionnaire. The update/leave effort began in late February in some areas and continued through March.

(3) **List/enumerate** In remote, sparsely populated, and hard-to-visit areas, census enumerators combined address listing and enumeration. There was no MAF for these areas created in advance. The enumerators searched for housing units, listed each unit in an address register (also its map spot), and enumerated the household at the same time. This operation was conducted in March–May 2000.

(4) **Remote Alaska** The enumeration procedure in remote areas of Alaska was similar to list/enumerate. It was conducted earlier (in February) before ice breakup and snow melt.

(5) **Rural update/enumerate** It was determined in some instances that blocks originally planned to be enumerated by update/leave would be better handled by a procedure in which address list updating and enumeration were conducted concurrently. “Rural” refers to the source of the address list—the address listing and LUCA 99 operations conducted outside the blue line.

Box C.2 (continued)

(6) **Military** Mailout/mailback procedures were used for all residential blocks on military bases (excluding group quarters). Such blocks in type 2 enumeration areas (but not those in type 1 enumeration areas) were assigned an enumeration area code of 6 because there was no need to update the address list or provide map spots.

(7) **"Urban" update/leave** It was determined that some blocks originally planned to have questionnaire delivery by the Postal Service would be better handled by having census enumerators follow an update/leave procedure. Such blocks contained older apartment buildings that lacked clear apartment unit designators, or they had many residents, despite having city-style addresses, who elected to receive mail at post office boxes. "Urban" refers to the source of the address list—the 1990 list updated by the DSF, the LUCA 98 Program, and the Postal Service check in early 2000. No map-spotting was needed for these addresses.

(8) **"Urban" update/enumerate** Some American Indian reservations contained blocks in more than one TEA. In these instances, all blocks in the reservation were enumerated using update/enumerate methods (see TEA 5). However, those blocks for which the mailing list was developed using "urban" procedures and for which no map-spotting was required were assigned a TEA code of 8 and not 5.

(9) **Mailout/mailback conversion to update/leave** Some blocks originally in TEA 1 areas contained a significant number of non-city-style addresses. They were identified and converted to "rural" address listing procedures before the urban block canvassing operation was carried out in 1999; they were reviewed as a special component of the LUCA 99 Program.

NOTE: For details, see U.S. Census Bureau (1999b).

C.2.a Redesign of Mailings and Materials to Boost Response

To encourage mail response, a new questionnaire format was adopted for 2000. Based on extensive research (see National Research Council, 1995b:Ch.6), a design was chosen that appeared as attractive and easy to fill out as possible. The mailing package was also redesigned to distinguish the questionnaire from junk mail and to motivate response (e.g., the envelope noted that responses were required by law).

One design change for the questionnaire was to ask households to list all members but to limit the space for characteristics to six members—instead of seven, as in 1990—in order to make the questionnaire less intimidating. It was planned to follow up households with more than six members by telephone (see Section C.5).

In the mailout/mailback area, multiple mailings were used to increase response. The first mailing was an advance letter (a new approach for 2000). The purpose of the letter was to alert residents to watch for the questionnaire, to provide a means for them to request a questionnaire in a language other than English, and to inform them of employment opportunities in census local offices. The second and third mailings were the questionnaire and a reminder postcard.

In both mailout/mailback and update/leave areas, the Bureau originally planned to deliver a second questionnaire to households not returning a form. Early testing showed that the use of a second questionnaire could increase mail response rates by as much as 10 percent (National Research Council, 1995b:120). However, the Bureau determined that vendors could not process the list of nonresponding households quickly enough to be able to mail out a replacement questionnaire on the schedule required. Mailing a second questionnaire to all households, as was done in the 1998 dress rehearsal, was deemed too expensive and likely to lead to negative publicity and confusion.

The advance letter operation did not proceed as smoothly as hoped. A programming error resulted in the addition of an extra digit to the beginning of every street address in the mailing; however, the barcode included on the letter used the correct address information, so that the Postal Service's sorting machines processed the letters properly. Hence, the addressing error did not stop the letters from being delivered.[6] In addition, the final version of the letter was not fully tested when it was decided after the 1998 dress rehearsal to add to the letter a way to request a foreign-language questionnaire. There was considerable public confusion about what to do with the enclosed return envelope if one did not need a special questionnaire. However, there were no apparent untoward effects of these problems on the public's cooperation with the census, and the publicity may have been helpful in alerting people to the need to respond.

C.2.b Multiple Response Modes

Another innovation for 2000 to encourage response was to allow multiple response modes. Households that received a short-form questionnaire could fill out a short form on the Internet or by telephone. To answer questions and also permit telephone response, the Bureau contracted with a commercial phone center to operate a toll-free telephone questionnaire assistance system. This system provided assistance in English, Spanish, and several other languages. Individuals could also pick up "Be Counted" forms, which were made available in six languages at various local sites throughout the country just prior to Census Day.

Because multiple response modes might not only boost return rates but also result in more responses that would require address verification and unduplication, the Census Bureau did not promote the Be Counted Program vigorously. Also, it did not widely publicize the Internet response option because of concerns about being able to handle a large response and maintain security. As it turned out, of 76 million questionnaires that were returned by households, 99 percent arrived by mail and only 1 percent by other modes: 66,000 were Internet returns; 605,000 were Be Counted forms; and 200,000 were forms completed by telephone. Not all Be Counted and telephone forms were included in the census: they were not counted if they did not have a valid address or if they duplicated another return.

[6]See Prepared Statement of Kenneth Prewitt, Director, U.S. Bureau of the Census, before the Subcommittee on the Census of the House Committee on Government Reform, March 8, 2000; http://www.census.gov/dmd/www/3-8-00.html [9/20/03].

C.2.c Comparison: 1990 Questionnaire Delivery and Return

Questionnaire delivery procedures in the 1990 census differed in some respects from those used in 2000 (National Research Council, 1995b:App.B). In 1990 about 84 percent of total housing units were in mailout/mailback areas; 11 percent—less than in 2000—were in update/leave areas (update/leave was a new procedure in 1990); and 5 percent—more than in 2000—were in list/enumerate areas. The list/enumerate procedure in 1990 differed somewhat from that used in 2000: Postal Service carriers delivered unaddressed short-form questionnaires to housing units in 1990, and census enumerators then came by to pick up completed questionnaires or obtain the answers, list the housing units in an address register, and at a predesignated subset of units, collect responses to the sample (long-form) questions. In 2000 Census Bureau field staff took questionnaires with them as they listed housing units and enumerated residents at the same time.

The 1990 census mailout procedures had not included an advance letter; however, a reminder postcard was delivered to all addresses in both mailout/mailback and update/leave areas. Responding by the Internet (which did not exist) was not an option. The questionnaire was designed not to facilitate response as much as to facilitate ready data capture (see Section C.5).

Overall, the mailing strategies used in the 1990 census did not appear to help mail response. The mail response rate declined from 75 percent in 1980 to 65 percent in 1990; the mail return rate declined from 81 percent in 1980 to 75 percent in 1990.[7]

C.3 FIELD FOLLOW-UP

Because not all households will mail back a form and because many addresses to which questionnaires are delivered will turn out to be vacant or nonresidential, the 2000 census, like previous censuses, included a large field follow-up operation (see Thompson, 2000). Over 500 local census offices (LCOs) were set up across the country, which reported to 12 regional census centers. The LCOs

[7] Mail response and return rates for 1970 and 1980 were not based on carefully controlled tracking systems, so they may not be comparable to the rates for 1990 and 2000.

were responsible for hiring the temporary enumerators and crew leaders who would be needed to conduct follow-up operations. In update/leave areas, enumerators were hired to deliver questionnaires prior to Census Day and to return to follow up nonresponding households and vacant units. LCOs also carried out operations to enumerate special groups, such as group quarters residents, transients, and the homeless.

In anticipation of possible difficulties in hiring and also the possibility that the mail response rate would decline from 1990, LCOs were authorized to recruit aggressively in advance of Census Day, to hire more enumerators than they thought would be needed, and to pay above-minimum wages (which differed according to prevailing area wages). Most offices were successful in meeting their hiring goals before the first follow-up operations began in mid-April 2000.

Follow-up operations were carried out in two separate stages, discussed below. The first stage, conducted in April–June, was the nonresponse follow-up, designed to obtain a questionnaire from every nonresponding unit in the mailback universe, occupied and vacant (or to determine that an address was nonresidential). The second stage, conducted in June–August, was called coverage improvement follow-up (CIFU), which included specific operations designed to check and supplement NRFU. Several operations included in CIFU for 1990 were dropped for 2000.

C.3.a Nonresponse Follow-Up

Preparation for NRFU began in early April 2000. Lists of addresses for inclusion in the NRFU workload were provided to the LCOs the week of April 11; a week later, notification was sent of late mail returns, which the LCOs had to delete manually from their follow-up lists. Also, the LCOs had to add information about surnames to their follow-up lists. The surname information was intended to help enumerators collect data accurately in situations in which questionnaires were misdelivered in multiunit structures and rural areas with clustered mailboxes. Because of a programming error, the surname information had to be sent separately to the LCOs (U.S. General Accounting Office, 2000c:11).

The final NRFU workload totaled 41.7 million addresses. This total included addresses in the MAF for which a completed questionnaire was not checked in prior to April 18 and new addresses from DSF updates. It also included addresses marked for deletion in the update/leave operation and addresses for which postal carriers returned the questionnaires as not deliverable and no attempt was made to redeliver them by census staff. The purpose of NRFU for these addresses was to doublecheck their status and, if they were in fact housing units, to obtain an enumeration.[8]

In most LCOs, NRFU enumerators went into the field beginning April 27. Their first objective was to visit each nonrespondent housing unit in person to try to obtain an interview, even if the residents said they had already mailed back a form,[9] or to obtain selected housing characteristics for vacant housing units. If unsuccessful, the enumerators were to try up to five additional times to obtain an interview, unless the residents were known to be out of town for an extended period or the housing unit was verified to be vacant or nonexistent by a proxy respondent (someone not a member of the household, such as a neighbor or landlord). Three of the follow-up attempts could be made by telephone if the enumerator could obtain a phone number. In the case of refusals, field observations indicated that some offices adhered to the six-visit rule, sometimes using different enumerators, while others allowed the use of proxy respondents without making all six visits. If no interview was obtained after the specified number of visits, then enumerators were instructed to obtain information from a proxy respondent, noting the name and address on the interview form. When an office had obtained information for 95 percent of its workload, the best enumerators were to be given the remaining cases to make one last attempt to obtain information from the household or a proxy, even if fewer than six visits had been made to the household. Some offices required that at least three visits be made to a household before allowing a last attempt.

Conducted concurrently with the NRFU enumeration was a quality assurance program, in which selected cases were reinter-

[8]NRFU enumerators could also enumerate housing units they identified that were not on the address list; however, field observation suggested that LCOs and enumerators did not consider checking the completeness of the address list to be part of the enumerator's job.

[9]Enumerators were instructed to be very diligent in this regard and to assure households that duplicate responses would be handled in the census processing.

viewed to identify fabrication ("curbstoning"). A random sample of the workload of each enumerator was reinterviewed; also, cases were selected purposively for reinterview by identifying enumerators whose work did not match that of other enumerators in the area. About 6 percent of the workload was reinterviewed in all, and preliminary analysis found discrepant results in 3 percent of the reinterview batches. The quality assurance reinterview process was delayed in some LCOs. Also, some reinterview forms were lost or not filled out correctly, so that analysis of the reinterview results must be interpreted with caution (see Baumgardner et al., 2001:17; see also Morganstein et al., 2003, who describe flaws in the quality assurance efforts).

NRFU operations were completed in most LCOs by June 26, so that the entire operation took only 8 weeks, shortening the original schedule by 1 week. At the conclusion of NRFU, enumerators had classified 62.3 percent of the 41.7 million addresses in their workload as occupied, 23.4 percent as vacant, 14.3 percent as "delete" (e.g., because the address identified a demolished structure or was nonresidential), and a handful (0.01 percent) as "not resolved" (Baumgardner et al., 2001:Table 4).

C.3.b Coverage Improvement Follow-Up

The coverage improvement follow-up effort that followed NRFU included several operations that involved about 8.9 million housing units. The largest portion of the workload comprised 6.5 million housing units that had been classified as vacant or delete in NRFU. These units, which CIFU checked to determine if they might have been occupied on Census Day, were only 41 percent of total addresses identified as vacant in NRFU. If such an address had not already been marked vacant or delete in another operation, it was revisited, but not otherwise. Examples of vacant or deleted units *not* included in CIFU were those classified as vacant or delete by an update/leave enumerator and a NRFU enumerator and those marked as undeliverable by a postal carrier and classified as vacant by a NRFU enumerator.[10] In addition, vacant units that NRFU enumerators had classified as "seasonal" were not checked in CIFU.

[10]Broadly speaking, the Census Bureau used a "two strikes and you're out" rule for assigning a final status of vacant or delete: if two separate operations classified an address as vacant or delete, then it generally was not checked again in CIFU.

There were five other components of the CIFU workload to visit or revisit: (1) 775,000 addresses that were added to the MAF in update/leave and urban update/leave, but from which no questionnaire was mailed back; (2) 372,000 addresses that were added to the MAF from the New Construction LUCA Program (in city-style areas); (3) 540,000 blank mail returns (including a small number of forms that were lost in the process of data capture); (4) 547,000 addresses that were added to the MAF from late updates from the Postal Service DSF; and (5) 86,000 addresses that were visited for some other reason. The fifth category included 62,000 addresses that were reenumerated in Hialeah, Florida; 17,000 addresses from the LUCA appeals process, and 7,000 other addresses (Moul, 2003:Table 2). A separate field operation was conducted to verify addresses on Be Counted forms and those filled out by telephone questionnaire assistance staff.

Addresses initially classified by CIFU itself as vacant or delete that had not been visited in any previous operation (e.g., an address added from the New Construction Program) were revisited for quality control purposes. The entire NRFU workload for one district office, in Hialeah, Florida, was reenumerated because of problems that came to light in that office (content was not being collected and sometimes not even the population count). Selected housing units were reenumerated in seven other offices for which problems were identified. The operations in 15 local offices were questioned by the House Subcommittee on the Census, but the Census Bureau determined, on review, that only two of these offices warranted some reenumeration. (These two offices are included in the total of seven in which partial reenumeration occurred.)

Overall, CIFU determined that 27 percent of the 8.9 million housing units visited were occupied, 43 percent were vacant, and 30 percent should be deleted. (Almost no units had an unresolved status at the end of CIFU; Moul, 2003:Table 3.) CIFU enumerators were most likely to find occupied units among the addresses added in update/leave; they classified 45 percent of these addresses as occupied. Other categories had lower percentages of units classified as occupied: 30 percent for blank mail returns, 27 percent for new construction addresses (52 percent of these addresses were not yet completed and so were deleted); and 23 percent for addresses classified as vacant or delete in NRFU that were checked in CIFU (Moul, 2003:Table 5). The percentage of NRFU vacant and delete addresses that CIFU reclassified as occupied, however, was 2 to 3 times the percentage of vacant and delete units found to be occupied in previous censuses for which a vacancy recheck was carried out (see below). The reason may be that, as noted at the beginning of the section, CIFU rechecked less than half of the addresses that were classified as vacant or delete by NRFU.

C.3.c Comparison: 1990 Field Follow-Up and Coverage Improvement

NRFU procedures in 1990 were similar in broad outline to the procedures used in 2000 (see Bureau of the Census, 1993:Ch.6). The NRFU enumerators were instructed to visit each household in person. If an enumerator could not obtain an interview but was able to obtain a telephone number, then he or she was to make up to five additional attempts to interview the household—three telephone attempts and two more personal visits at different times of the day. If the enumerator did not have a telephone number, he or she was to make two additional personal visits. When all of these attempts failed to result in an interview or if the case was a refusal or the respondent was away for an extended period of time, the enumerator was instructed to talk to someone outside the household to obtain "last resort" information. Such information was defined as three of the four characteristics of relationship to head of household, sex, race, and marital status for each household member and the number of units in the structure for each housing unit. When 95 percent of the caseload had been completed, the remaining cases were given to the best enumerators who were to make one last visit to try to gather "closeout" data, defined as at least two characteristics for each household member.

Concurrently with NRFU enumeration, a reinterview program was carried out to detect falsification, similar to the program in 2000. The 1990 quality control program reinterviewed 4.8 percent of the NRFU workload of 34 million housing units and estimated a very low rate of falsification overall (0.09 percent; see Bureau of the Census, 1994:30–34).

In contrast to 2000, the 1990 NRFU operations fell considerably behind schedule, largely because of the Census Bureau's failure to forecast the extent of the decline in the mail response rate from 1980 to 1990—the Bureau projected a 70 percent response rate (down from 75 percent in 1980), but the actual rate at the time NRFU began was 63 percent (the rate subsequently rose to 65 percent). The Bureau had to obtain additional appropriations and scramble to hire sufficient workers for NRFU and other follow-up activities; it raised pay rates in 140 of the 449 district offices (equivalent to LCOs) and took other steps to increase productivity. The NRFU operation was planned to take 6 weeks from when it began in late April; however, only 72 percent of the workload was completed by that time (by June 6). Another 18 percent of the workload was completed in 2 more weeks, but it took another 6 weeks—until early August—to complete the remaining 10 percent of the workload (U.S. General Accounting Office, 1992:46).

A subsequent stage of follow-up in 1990 included several coverage improvement procedures (Bureau of the Census, 1993:6-37 to 6-38;6-53 to 6-56). An operation called field follow-up, carried out in June–August, rechecked most units classified as vacant or delete in NRFU. Units that were not rechecked included those in areas with high proportions of seasonal housing or boarded-up buildings, plus units classified as delete by two precensus address update operations and a NRFU enumerator (a more stringent criterion than that used in 2000). By August 1, 5.3 percent of deleted units and 7.1 percent of vacant units that were rechecked in field follow-up were converted to occupied. (The corresponding percentages in 1980 were 7.5 percent deleted units and 10 percent of vacant units converted to occupied.) These figures are considerably below the rate of conversion from vacant or delete to occupied in the 2000 CIFU (24 percent).

In addition to the recheck of vacant and delete units, the 1990 field follow-up operation revisited failed-edit mail returns. These cases were mail returns that failed computer or clerical review with regard to completeness of coverage and content (Bureau of the Census, 1995b:8-10) and for which telephone follow-up was not successful (see Section C.5). Because of backlogs in the telephone follow-up operation for questionnaires handled by processing offices (those from central city areas), after mid-June most questionnaires in these offices that failed the content review and were not resolved by telephone were not sent to field follow-up. The 1990 field follow-up also revisited a number of mailback cases for which there was no record of data capture.

Another 1990 coverage improvement operation was the "Were You Counted" campaign, in which people who thought they had been missed were encouraged by media announcements in June–July 1990 to send in a special form. Those forms with addresses that could be assigned to census geography and with complete content were put through a search and matching operation to determine if they duplicated other forms. There was no field verification of the address, except in the Detroit district office, from which an unusually large number of forms were received.

Another special operation was the recanvass, carried out in July–November 1990, in which selected blocks, including those in high growth areas and those identified by postcensus local review, were relisted. The households were then reenumerated, provided the enumerator determined that the unit existed as of April 1. In all, the Bureau recanvassed more than 650,000 blocks containing about 20 million housing units (20 percent of all units).

Blocks identified for recanvassing by localities came about because in 1990 (though not 2000), local jurisdictions nationwide were invited to review preliminary census counts of housing units by block for their areas (Bureau of the Census, 1993:6-45 to 6-46). The counts were provided in August 1990, and localities had 15 days to challenge them. Responses were received from about 25 percent of all jurisdictions, including all of the 51 largest cities. All challenged blocks in which the discrepancy between the census count and that provided by the locality exceeded a specified amount were added to the recanvass operation, for which additional funding had to be obtained.

As part of the coverage improvement effort in 1990, in 24 local offices, all households for which the questionnaires reported only one household member were reenumerated. This procedure was implemented in response to allegations in late summer 1990 that enumerators in some offices during the closeout phase of NRFU had recorded households as one-person households without actually obtaining an interview (i.e., they were curbstoning). In addition, seven local offices in New Jersey were identified in which it appeared that fabrication may have occurred; households in these offices were reinterviewed when the questionnaires indicated household size but recorded no members (Bureau of the Census, 1993:6-55).

Finally, a special program was implemented to improve the coverage of people who were on parole or probation (Bureau of the Census, 1993:6-55). The first step was to contact each state to ask its parole or probation officers to distribute census forms to their assignees to be filled out and mailed back. This operation had a very low response rate, so census enumerators were sent to correction departments in designated counties to obtain information for parolees and probationers from administrative records. No attempt was made to contact parolees or probationers unless their addresses could not be verified. The operation was not completed until late November-early December 1990. The forms obtained were processed through an unduplication operation (see Section C.5); however, subsequent analysis determined that many of the parolee/probationer forms that were accepted in the census count represented erroneous enumerations (Ericksen et al., 1991:43–46).

C.3.d Summary: 1990 and 2000

The description of 2000 and 1990 follow-up procedures makes it clear that they were large-scale, complex operations, similar in broad outline but sufficiently different in detail to make it difficult to compare results across years. It is difficult, for example, to compare results from the 2000 CIFU recheck of vacant and delete units with the 1990 field follow-up vacancy check because of differences in how the workload was defined. Also, it is not clear exactly how such terms as "proxy" (2000), "last resort" (1990), "closeout," and "non-data-defined" were similar or dissimilar, again complicating the task of comparative evaluation.

One can, however, conclude that the Census Bureau was more successful in 2000 than in 1990 in controlling field follow-up operations and keeping them on schedule. Coverage improvement operations were more focused, and programs that appeared problematic in 1990 (e.g., the parolee and probationers check) were not repeated in 2000.

C.4 OUTREACH EFFORTS

To supplement field operations and special programs to improve population coverage and cooperation with the census, the Census Bureau engaged in large-scale advertising and outreach efforts for 2000. For the first time, the Census Bureau budget included funds ($167 million) for a paid advertising campaign (recommended by National Research Council, 1978). In previous censuses, the Advertising Council arranged for advertising firms to develop ads and air them on a pro bono, public service basis (Anderson, 2000).

The 2000 advertising campaign was extensive, involving a major contractor, Young and Rubicam, which contracted with four other agencies to prepare ads targeted to particular population groups and communities. The advertising ran from November 1, 1999, to June 5, 2000, and included a phase to alert people to the importance of the upcoming census, a phase to encourage filling out the form, and a phase to encourage people who had not returned a form to cooperate with the follow-up enumerator. Ads were placed on television (including one during the 2000 Super Bowl), radio, newspapers, and other media, using multiple languages. Based on market research, the ads stressed the benefits to people and their communities from the census, such as better targeting of government funds to needy areas for schools, day care, and other services.

In addition to the ad campaign, the Census Bureau hired partnership and outreach specialists in local census offices, who worked with community and public interest groups to develop special initiatives to encourage participation in the census. The Bureau signed partnership agreements with over 100,000 organizations, including federal agencies, state and local governments, business firms, nonprofit groups, and others. The Bureau did not fund these groups, but it provided materials and staff time to help them encourage a complete count. A special program was developed to put materials on the census in local schools to inform school children about the benefits of the census and motivate them to encourage their adult relatives to participate.

The Census Bureau director and other staff made numerous public appearances throughout the census period to stress the importance of a complete count and respond to questions and concerns. The director also put into place a program to use the Internet to challenge communities to raise their mail response rates. The 1990 response rates were posted for local areas on the Bureau's Web site beginning in mid-March, and 2000 response rates were regularly updated on the site through mid-April. Communities were challenged to exceed their 1990 rates by 5 percent. Although few communities achieved this goal, the overall response rate did not continue its decline from previous censuses.

The 1990 census had also included advertising and outreach efforts; however, their extent was less than in 2000. The advertising was prepared by a firm selected by the Advertising Council, which conducted its work on a pro bono basis. Ads were placed as public service announcements, which meant that many ads ran in undesirable times. The partnership program was not as extensive as in 2000.

In both censuses, perhaps more so in 2000, advertising and outreach efforts varied in intensity across the country. Some localities were more active than others in coordinating and supplementing outreach and media contacts. Whether this variability narrowed or widened the difference in net undercount rates among major population groups depends on the extent to which outreach efforts were more (or less) effective in hard-to-count areas in comparison with other areas.

C.5 DATA PROCESSING

Data processing for the 2000 census was a continuing, high-volume series of operations that began with the capture of raw responses and ended with the production of voluminous data products for the user community, which were made available in 2001–2003.[11] Important innovations were adopted for 2000. For the first time, the Census Bureau contracted with outside vendors for major components of data processing. Also for the first time, data capture operations were carried out using optical character recognition technology in addition to optical mark recognition. A telecommunications network linked Census Bureau headquarters in Suitland, Maryland; 12 permanent regional offices; the Bureau's permanent

[11]Data processing also included a series of computer systems for management of operations, including payroll, personnel, and management information systems.

computer center in Bowie, Maryland; 12 regional census centers and the Puerto Rico Area Office; the Bureau's permanent National Processing Center in Jeffersonville, Indiana; 3 contracted data capture centers in Phoenix, Arizona, Pomona, California, and Baltimore County, Maryland; 520 local census offices; and contracted telephone centers for questionnaire assistance (U.S. Census Bureau, 1999b:XI-1).

Five operations in 2000 are described in this section: data capture, coverage edit and telephone follow-up, unduplication, editing and imputation, and other data processing. Data processing operations for 1990 are also summarized.

C.5.a Data Capture

The first step in data processing was to check in the questionnaires and capture the data on them in computerized form. The return address on mailback questionnaires directed them to one of four data capture centers—the Bureau's National Processing Center and three run by contractors. Each questionnaire had a bar code that was scanned to record its receipt. The questionnaires were then imaged electronically, check-box data items were read by optical mark recognition (OMR), and write-in character-based data items were read by optical character recognition (OCR). Clerks keyed data from images in cases when the automated technology could not make sense of the data. Keying of the additional long-form-sample items was deferred until fall 2000 to permit the fastest possible processing of the basic (complete-count) data from short and long forms.

C.5.b Coverage Edit and Telephone Follow-Up

The data on mailed-back questionnaires were reviewed by computer to identify those returns that failed coverage edit specifications. These failed-edit cases were reinterviewed by telephone, using contractor-provided clerical telephone staff. The workload for the coverage edit and telephone follow-up operation totaled about 2.3 million cases. It included returns that reported more, or fewer, household members in question one ("How many people were living or staying in this house, apartment, or mobile home on April 1, 2000?") than the number of members for which individual information (e.g., age, race, sex) was provided; returns in which question one was left blank and individual information was provided for exactly six people (the limit of the space provided on the mail questionnaires); and returns that reported household counts of seven people or more.

The purpose of the edit and telephone follow-up was to reduce coverage errors in the households selected for follow-up and to obtain basic characteristics for household members for whom the household had no room to report their characteristics on the form. No characteristics were obtained for missing responses for household members for whom only some characteristics were reported. There was no field follow-up for failed-edit households for which telephone follow-up was unsuccessful. Because of computer problems, the start of the coverage edit and telephone follow-up operation was delayed. Originally planned to be conducted in April–June 2000, it was carried out in May through mid-August.

C.5.c Unduplication of Households and People

Two major, computer-based unduplication operations were carried out subsequent to field follow-up. One of those operations, the use of the primary selection algorithm (PSA) to unduplicate multiple returns for the same address, was planned from the outset and is described below. The other operation, the use of special software and procedures to reduce duplication of addresses in the MAF, was designed and implemented in summer 2000 to respond to evidence of duplicate addresses not eliminated by previous processing (described in Section C.1.d). The special unduplication operation used the results of the PSA; final determination of which returns to delete from the census because they duplicated a return from another MAF address was not made until after the PSA had processed multiple returns for the same address.

The purpose of the PSA was to identify unique households and people to include in the census when more than one questionnaire was returned with the same census address identification number. Such duplication could occur in a number of ways: when a respondent mailed back a census form after the cutoff date for determining the NRFU workload and the enumerator then obtained a second form from the household (or perhaps identified the household as vacant); when someone was enumerated in a group quarters but provided another "usual" address to which his or her information was assigned; or when a respondent filled out a Be Counted form, thinking that he or she had been missed, but another member of the household also mailed back a questionnaire for the household (which might or might not contain information for the individual).

For each housing unit, returns with one or more persons in common were combined to form a single PSA household, retaining only one response for each household member who was reported on more than one return, as well as the responses for household members who were reported on only one return. All vacant returns for a housing unit were also combined to form a PSA household. In some cases more than one PSA household might exist for a unit. For each PSA household, the algorithm selected which return best represented the Census Day household ("basic" return) and which people from the other returns were part of that household.[12]

In all, 9 percent of census housing units had two returns that were eligible for the PSA operation, and 0.4 percent had three or more eligible returns. (Extra returns for an address that had no useful information were not included in the operation.) In most instances, the operation of the PSA discarded duplicate household returns or extra vacant returns. Less often, the PSA found additional people to assign to a basic return or identified more than one household at an address (see Baumgardner et al., 2001:22–27).

C.5.d Editing and Imputation

Editing and imputation were carried out for all data-captured questionnaires that were retained in the census after the PSA operation. The editing and imputation process included whole-household imputation, called substitution, when there was minimal or no information for the housing unit; editing content items for consistency and to fill in (assign) values for missing items on the basis of related items (e.g., to calculate age when only date of birth was provided); and imputation of content items using values reported for another person or household, called allocation, when values were missing for one or more items. See Box 4.2 in Chapter 4 for types of whole-household imputation; Table 4.1 for whole-household imputation

[12]The Census Bureau does not make public the criteria for the PSA.

rates by type in the 1980–2000 censuses; and Chapter 7 for basic (complete-count) and long-form-sample item imputation rates.

All editing and imputation were computer based; there was no clerical review or editing of any items as in past censuses. When it was not possible to perform an edit that used other information for the same person or housing unit, imputation was performed using hot-deck methods that made use of information for other people and households, generally in the immediate neighborhood. First used in processing the 1960 census, the Census Bureau's computerized hot-deck procedures have been refined and elaborated. The donor pool is geographically restricted to take advantage of common characteristics among small-area populations (see Appendix G; see also Citro, 2000b).

C.5.e Other Data Processing

A number of other data processing steps were carried out to generate data files and publications from the 2000 census records. Such steps for the complete-count records included tabulating the data on various dimensions and modifying the data appropriately on files that were to be released to the public in order to protect the confidentiality of individual responses. For the long-form-sample records, there were the added steps of coding such variables as occupation and industry and weighting the records to complete-count control totals on several dimensions.

C.5.f Comparison: 1990 Data Processing

The 1990 census data processing system was more decentralized than in 2000 and made more use of clerical review (see Bureau of the Census, 1995b:Ch.8; National Research Council, 1995b:App.B). There were seven processing offices and 559 district offices. Mail-back questionnaires in district offices in hard-to-enumerate areas in central cities went directly to a processing office for check-in by scanning bar codes, data capture by using the Census Bureau's Film Optical Sensing Device for Input to Computers (FOSDIC), and computerized review to identify cases that failed to meet the edit specifications for completeness of coverage and content. Failed-edit cases went to telephone follow-up, and those cases that could not be contacted were sent to a district office for field follow-up. However, backlogs in the telephone follow-up operation necessitated curtailment of field follow-up for cases that could not be contacted by telephone, and, for cost reasons, only a 10 percent sample of mailed-back short forms that failed the content review (and not also the coverage review) were sent for telephone follow-up. Enumerator returns for central city offices were checked-in at the district office and then sent to the processing office for data capture, computerized review of coverage and content, and telephone follow-up as needed. Enumerator returns were not eligible for field follow-up. Once any further data had been received from follow-up, computerized editing, whole-household imputation, and item imputation routines were used to fill in remaining missing or inconsistent data.

Mailback and enumerator returns in other district offices went first to the district office for check-in, clerical review of coverage and content, telephone follow-up as needed, and field follow-up of failed-edit mail returns for which telephone follow-up was unsuccessful. After completion of follow-up, the questionnaires were sent to the processing offices for data capture and computerized editing and imputation.

Another step in data processing included the search/match operation, in which forms received from various activities were checked against microfilm images of questionnaires for the same address to determine which people should be added to the household roster and which were duplicates. This operation was carried out for "Were You Counted" forms, for parolee/probationer forms, and for people who sent in a questionnaire from one location with an indication that their usual home was elsewhere. Such people might have two homes, such as people who spend the winter in a southern state and the summer in a northern state. There was no way on the 2000 form to indicate usual home elsewhere.

The FOSDIC technology used for data capture was originally developed by Census Bureau staff for the 1960 census and reengineered and enhanced in 1970, 1980, and 1990 (see Salvo, 2000). It involved two main stages: microfilming the questionnaires and using the FOSDIC equipment to scan the microfilm and read the filled-in answer circles for each item and output responses to a computer file (the answer dots showed up as light images on a dark background). In 1990, FOSDIC processed over 130 million questionnaires; about

900,000 forms had to be "repaired" by clerks and remicrofilmed before they could be read (e.g., because the forms were torn or folded improperly and so were out of alignment for scanning). The FOSDIC equipment could read answer dots and sense the presence of write-in entries but not capture such entries directly. Write-in responses were keyed by clerks using the paper questionnaires for long-form-sample items and a microfilm access device for keying of write-in responses to the race question. After keying, the write-in responses were coded by a combination of computer and clerical review.

Appendix D

Completeness of Census Returns

This appendix summarizes research from the 1990 and 2000 censuses documenting the generally more complete population coverage on mail returns compared with enumerator-obtained returns. See Chapter 7 and Appendixes G and H for documentation that mail returns generally have lower missing data rates for content items on the short and long forms compared with enumerator-obtained returns.

D.1 COVERAGE COMPLETENESS: 1990

Research from the 1990 census, based on a match of P-sample and E-sample records in the 1990 Post-Enumeration Survey (PES), found that mail returns were substantially more likely than returns obtained by enumerators to cover all people in the household. Only 1.8 percent of mail returns had within-household misses, defined as cases in which a mail return in the E-sample matched a P-sample housing unit but the P-sample case included one or more people who were not present in the E-sample unit. In contrast, 11.6 percent of returns obtained by enumerators had within-household misses (Siegel, 1993; see also Keeley, 1993). These rates did not vary by type of form: within-household misses were 1.9 percent and 1.8

percent for short-form and long-form mail returns and 11.7 percent and 11.3 percent for short-form and long-form enumerator-filled returns.

In an analysis of the 1990 PES for 1,392 poststrata, Ericksen et al. (1991:Table 1) found that both the gross omission rate and the gross erroneous enumeration rate were inversely related to the "mailback rate" (equivalent to the mail response rate) for PES cases grouped by mailback rate category.[1] The relationship was stronger for omissions than for erroneous enumerations—the omission rate was 3 percent in the highest mailback rate category and 19 percent in the lowest mailback rate category, compared with 4 percent and 10 percent, respectively, for the erroneous enumeration rate. Consequently, the net undercount rate also varied inversely with the mailback rate.

D.2 COVERAGE COMPLETENESS: 2000

Using data from the original Accuracy and Coverage Evaluation (A.C.E.) P-sample and E-sample, we carried out several analyses of the relationship between mail returns and population coverage for 2000. The analyses are not as comparable as we would have liked to the 1990 analyses summarized above: not only are there differences between the PES and the A.C.E., but also it is difficult a decade later to determine exactly how the 1990 analyses were performed. Nonetheless, the work is sufficiently similar that we are confident that the findings, which largely confirm the 1990 results, are valid.

D.2.a Within-Household Omissions and Erroneous Enumerations by Type of Return

We linked P-sample and E-sample records in the same housing units to provide a basis for calculating rates of within-household omissions for 2000 that could be compared to the 1990 rates from Siegel (1993). We also developed other classifications of linked P-sample and E-sample households.

Table D.1 shows our results: E-sample mail returns received before the cutoff for determining the nonresponse follow-up workload

[1]Ericksen et al. (1991) defined 10 mailback rate categories: one for under 55 percent, eight intervals of 5 percentage points from 55–59.9 percent to 80–84.9 percent, and one for 85 percent and over.

included proportionally fewer cases with one or more omissions or possible omissions (2.8 percent) than did E-sample returns that were obtained by enumerators in nonresponse follow-up (7.0 percent). The difference was in the same direction as in 1990, but it was not as pronounced. Perhaps more striking, enumerator-obtained returns in 2000 included a much higher proportion with one or more erroneous or unresolved enumerations (15.5 percent) than did mail returns (5.3 percent) (comparable data are not available for 1990). Such enumerations included duplicates, geocoding errors, people lacking enough reported data for matching, and other erroneous and unresolved enumerations.

If the corrections for additional duplications and other A.C.E. measurement errors, which were first reported by the Census Bureau in October 2001 (Executive Steering Committee for A.C.E. Policy, 2001b), and corrections for other measurement errors could be incorporated in the analysis, they would affect the percentages cited in the text. However, they would not likely affect the relationship in which erroneous enumerations occur more frequently in enumerator-obtained than mail returns. Indeed, the A.C.E. Revision II results showed lower estimated correct enumeration rates for enumerator return poststrata compared with mailback poststrata (U.S. Census Bureau, 2003c:Table 6).

By housing tenure, both owner-occupied households and renter households showed the same patterns: mail returns included proportionally fewer cases of within-household omissions or cases with one or more erroneous or unresolved enumerations than enumerator returns. Consistently, renter households had higher proportions of these kinds of households than owner households (comparable data are not available for 1990).

D.2.b Omissions and Erroneous Enumerations by Mail Return Rate Deciles

In an analysis similar to Ericksen et al. (1991), Table D.2 classifies P-sample cases and E-sample cases in the 2000 mailback universe into 10 mail return rate categories, with each category defined to include 10 percent of the total. (The decile cutoffs are very similar between the two samples.) The mail return rate associated with each case is the rate for the census tract in which the P-sample or

Table D.1 Composition of 2000 Census Households, as Measured in the Original A.C.E. E-Sample, by Enumeration Status, Mail and Enumerator Returns, and Housing Tenure (weighted)

	Total		Owner		Renter	
Type of Census (E-Sample) Household	Mail Return (%)	Enumerator Return (%)	Mail Return (%)	Enumerator Return (%)	Mail Return (%)	Enumerator Return (%)
No Valid Omissions from P-Sample	97.1	93.0	97.5	93.8	96.3	92.3
All Members Match with P-Sample	81.5	62.9	84.6	69.4	73.0	56.1
All Matches or Correct (nonmatched) Enumerations	10.3	14.6	8.7	12.7	14.7	16.6
At Least One Correct and One Erroneous or Unresolved Enumeration	2.1	4.2	1.9	3.6	2.8	4.9
All Erroneous or Unresolved Enumerations	3.2	11.3	2.3	8.1	5.8	14.7
One or More Omissions or Possible (Unresolved) Omissions	2.8	7.0	2.5	6.2	3.7	7.6
One or More Omissions	2.3	5.7	2.2	5.3	2.9	6.0
One or More Possible Omissions	0.5	1.3	0.3	0.9	0.8	1.6

NOTES: Households in all categories may contain P-sample cases classified as unresolved or removed. Mail returns are those received before the cutoff for determining the nonresponse follow-up universe (variable NRU = 1); enumerator returns are those obtained in nonresponse follow-up (variable NRU = 3). Enumeration status is before imputation for unresolved cases.

SOURCE: Tabulations by panel staff from P-Sample and E-Sample Person Dual-System Estimation Output Files (U.S. Census Bureau, 2001b), provided to the panel February 16, 2001. Data files matched by household and weighted using the median value of TESFINWT within households.

Table D.2 Rates of P-Sample Omissions, E-Sample Erroneous Enumerations, and P-Sample and E-Sample Unresolved Cases in the Original 2000 A.C.E., by Mail Return Rate Decile of Census Tract (weighted)

	P-Sample Rates (%)		E-Sample Rates (%)	
Census Tract Decile (Return Rate Range)[a]	Omissions[b]	Unresolved Cases[c]	Erroneous Enumerations[d]	Unresolved Cases[e]
10th (82.8–100.0)	3.8	1.2	2.5	1.3
9th (79.7–82.7)	4.8	1.7	2.7	1.7
8th (77.3–79.6)	5.2	1.9	3.2	1.9
7th (74.9–77.2)	5.8	2.0	3.6	1.9
6th (72.6–74.8)	6.8	2.2	3.9	2.1
5th (69.9–72.5)	7.6	2.3	4.6	2.8
4th (66.8–69.8)	8.3	2.4	4.6	2.6
3rd (63.2–66.7)	9.4	2.5	5.0	3.5
2nd (57.7–63.1)	11.6	3.0	5.7	4.3
1st (19.9–57.6)	14.8	3.6	7.2	4.5

[a] The return rate ranges shown are for the P-sample; the ranges for the E-sample are almost identical.

[b] The omission rate is omissions divided by the sum of omissions plus matches (excluding unresolved cases, cases removed from the P-sample as not appropriately in the sample of Census Day household residents, and inmovers who were not sent through the matching process).

[c] The unresolved rate for the P-sample is unresolved cases divided by the sum of omissions, matches, and unresolved cases.

[d] The erroneous enumeration rate is erroneous enumerations divided by the sum of erroneous enumerations, matches, and other correct enumerations.

[e] The unresolved rate for the E-sample is unresolved cases divided by the sum of unresolved cases, erroneous enumerations, matches, and other correct enumerations.

SOURCE: Tabulations by panel staff from P-Sample and E-Sample Dual-System Estimation Output Files (U.S. Census Bureau, 2001b), provided to the panel, February 16, 2001; weighted using TESFINWT.

E-sample person resided. Within each P-sample decile, the omission rate is calculated as the ratio of valid P-sample cases that did not match an E-sample person to the total of nonmatches plus matches. Within each E-sample decile, the erroneous enumeration rate is calculated as the ratio of erroneous enumerations (duplicates, fictitious persons, etc.) to the total of erroneous enumerations plus correct enumerations.

The omission rate ranges from 3.8 percent in the highest mail return rate decile to 14.8 percent in the lowest decile. The erroneous enumeration rate ranges from 2.5 percent in the highest return rate decile to 7.2 percent in the lowest mail return rate decile. The differences are in the same direction as those estimated for 1990, although they are not as pronounced.[2]

The rate of unresolved cases in the P-sample and E-sample (cases whose match or enumeration status could not be resolved even after field follow-up) also shows a relationship to mail return rate deciles (see Table D.2). The unresolved rate (unresolved cases as a percentage of unresolved plus matches and nonmatches for the P-sample, and as a percentage of unresolved plus correct and erroneous enumerations for the E-sample) ranges from 1.2 percent in the highest P-sample mail return rate decile to 3.6 percent in the lowest decile and from 1.3 percent in the lowest E-sample mail return rate decile to 4.5 percent in the highest decile. These results indicate that it was easier to determine match or enumeration status in areas with higher mail return rates.

D.2.c Erroneous Enumerations by Domain and Tenure

In an analysis of 2000 data—for which we have not seen comparable findings for 1990—we examined erroneous enumeration rates for E-sample people (including unresolved cases) by whether their household mailed back a return or was enumerated in the field, looking separately at cases in the mailout/mailback and update/leave/mailback universes. This analysis finds considerable overall differences and for domain (race/ethnicity; see Appendix E, Table E.3) and tenure groups. The analysis differs from that reported above for mail return rate deciles for 1990 and 2000 in that it uses the mail return status of the individual household to classify E-sample people and not the return rate of either the poststratum (as in 1990) or the census tract (as in 2000).

As shown in Table D.3, the rate of erroneous and unresolved enumerations for people on mail returns in the mailout/mailback universe is 4.3 percent, compared with 12.6 percent for people on returns obtained by a nonresponse follow-up enumerator in that universe. For most race/ethnicity groups, people on mail returns have lower erroneous and unresolved enumeration rates than do

[2]We cannot determine the comparability of the 1990 and 2000 analysis due to missing details on how mail return, omission, and erroneous enumeration rates were calculated for 1990 by Ericksen et al. (1991). Hence, the comparison between the 1990 and 2000 results should be limited to order of magnitude.

people on enumerator-obtained returns, and owners have lower rates than renters. There are two exceptions: American Indian and Alaska Native owners living off reservations and Native Hawaiian and other Pacific Islander owners, for whom the rates of erroneous and unresolved enumerations are similar between mail and enumerator-obtained returns; and American Indian and Alaska Native renters living on reservations, for whom the rates of erroneous and unresolved enumerations are higher for mail returns than for enumerator-obtained returns.

For the update/leave cases, the patterns are similar to mailout/mailback cases: the rate of erroneous and unresolved enumerations for people in update/leave areas is 3.1 percent, compared with 9.2 percent for people on enumerator-obtained returns, with rates for owners lower than those for renters (see Table D.3). The exceptions are American Indians and Alaska Natives living on reservations, for whom people on mail returns have higher (not lower) rates of erroneous enumerations than do people on enumerator-obtained returns, and for whom owners have higher rates than renters.

There were very few list/enumerate cases, but they have a high rate of erroneous and unresolved enumerations—18.3 percent (data not shown). In contrast, the rates are relatively low for rural update/enumeration (5.2 percent) and urban update/enumeration (2.9 percent). For all other enumerator-obtained returns (e.g., those obtained from new construction), the rates of erroneous and unresolved enumerations is high—14.3 percent.

Table D.3 Rates of E-Sample Erroneous Enumerations and Unresolved Cases, in Mailout/Mailback and Update/Leave Types of Enumeration Area (TEA), by Mail or Enumerator Return, Race/Ethnicity Domain, and Housing Tenure, Original 2000 A.C.E. (weighted)

	Percent Erroneous and Unresolved Cases of Total E-Sample Enumerations			
	Mailout/Mailback TEA		Update/Leave TEA	
Race/Ethnicity Domain and Tenure Category	Mail Return	Enumerator Return	Mail Return	Enumerator Return
American Indian and Alaska Native on Reservation				
Owner	0.0	6.5	13.5	9.0
Renter	5.1	1.9	10.2	3.9
American Indian and Alaska Native off Reservation				
Owner	7.2	7.7	2.2	8.7
Renter	9.1	12.8	5.5	11.3
Hispanic				
Owner	3.4	8.2	4.0	7.3
Renter	7.0	15.2	7.1	15.1
Black				
Owner	4.6	11.7	3.7	7.3
Renter	8.4	16.8	5.6	10.6
Native Hawaiian and Pacific Islander				
Owner	6.6	6.9	7.4	6.7
Renter	5.3	13.8	10.8	13.5
Asian				
Owner	4.0	8.8	3.5	8.5
Renter	8.4	17.5	8.4	23.2
White and Other Non-Hispanic				
Owner	3.0	8.5	2.6	8.0
Renter	7.2	16.6	5.2	12.9
Total	4.3	12.6	3.1	9.2

NOTES: See Table E.3 in Appendix E for definitions of race/ethnicity domains. Mail returns are those received before the cutoff for determining the nonresponse follow-up universe (variable NRU = 1). Enumerator returns are those obtained in nonresponse follow-up (variable NRU = 3).

SOURCE: Tabulations by panel staff from E-Sample Person Dual-System Estimation Output Files (U.S. Census Bureau, 2001b), provided to the panel February 16, 2001; weighted using TESFINWT.

Appendix E

A.C.E. Operations

This appendix describes the operations of the original 2000 Accuracy and Coverage Evaluation (A.C.E.) Program, which produced population estimates in March 2001.[1] Differences from the analogous 1990 Post-Enumeration Survey (PES) are summarized in Chapter 5, which also describes the dual-systems estimation (DSE) method used to develop population estimates for poststrata from the A.C.E. results. Chapter 6 describes the differences in estimation methods used for the A.C.E. Revision II results, which were made available in March 2003. This appendix covers six topics: sampling, address listing, and housing unit match (E.1); P-sample interviewing (E.2); initial matching and targeted extended search (E.3); field follow-up and final matching (E.4); weighting and imputation (E.5); and poststrata estimation (E.6).

E.1 SAMPLING, ADDRESS LISTING, AND HOUSING UNIT MATCH

The 2000 A.C.E. process began in spring 1999 with the selection of a sample of block clusters for which an independent listing of addresses was carried out in fall 1999. The selection process was designed to balance such factors as the desired precision of the DSE

[1]See Childers and Fenstermaker (2000) and Childers (2000) for detailed documentation of A.C.E. procedures.

estimates, not only for the total population, but also for minority groups, and the cost of field operations for address listing and subsequent interviewing. In addition, the A.C.E. selection process had to work within the constraints of the design originally developed for integrated coverage measurement (ICM).

E.1.a First-Stage Sampling and Address Listing of Block Clusters

Over 3.7 million block clusters were formed that covered the entire United States, except remote Alaska.[2] Each cluster included one census collection block or a group of geographically contiguous blocks, in which the block(s) were expected to be enumerated using the same procedure (e.g., mailout/mailback) and to contain, on average, about 30 housing units on the basis of housing unit counts from an early version of the 2000 Master Address File (MAF). The average cluster size was 1.9 blocks.

Next, clusters were grouped into four sampling strata: small (0–2 housing units), medium (3–79 housing units), large (80 or more housing units), and American Indian reservations (in states with sufficient numbers of American Indians living on reservations). Systematic samples of block clusters were selected from each stratum using equal probabilities, yielding about 29,000 block clusters containing about 2 million housing units, which were then visited by Census Bureau field staff to develop address lists.

The sample at this stage was considerably larger than that needed for the A.C.E. The reason was that the Census Bureau had originally planned to field a P-sample of 750,000 housing units for use in ICM, and there was not time to develop a separate design for the planned A.C.E. size of about 300,000 housing units. So the ICM block cluster sample design was implemented first and then block clusters were subsampled for A.C.E., making use of updated information from the address listing about housing unit counts.[3]

[2]A.C.E. operations were also conducted in Puerto Rico; the Puerto Rico A.C.E. is not discussed here.

[3]Our panel reviewed this decision and found it satisfactory because the development of direct dual-systems estimates for states was not necessary in the A.C.E. as it would have been under the ICM design (National Research Council, 1999a, reprinted in Appendix A.4.a).

E.1.b Sample Reduction for Medium and Large Block Clusters

After completion of the address listing and an update of the MAF, the number of medium and large block clusters was reduced, using differential sampling rates within each state. Specifically, medium and large clusters classified as minority on the basis of 1990 data were oversampled to improve the precision of the DSE estimates for minority groups. Also, clusters with large differences in housing unit counts from the P-sample address list and the January 2000 version of the MAF were oversampled in order to minimize their effect on the variance of the DSE estimates.

E.1.c Sample Reduction for Small Block Clusters

The next step was to stratify small block clusters by size, based on the current version of the MAF, and sample them systematically with equal probability at a rate of 1 in 10. However, all small block clusters that were determined to have 10 or more housing units and all small block clusters on American Indian reservations, in other American Indian areas, or in list/enumerate areas were retained. After completion of the cluster subsampling operations, the A.C.E. sample totaled about 11,000 block clusters.

E.1.d Initial Housing Unit Match

The addresses on the P-sample address listing were matched with the MAF addresses in the sampled block clusters. The purpose of this match was to permit automated subsampling of housing units in large blocks for both the P-sample and the E-sample and to identify nonmatched P-sample and E-sample housing units for field follow-up to confirm their existence. Possible duplicate housing units in the P-sample or E-sample were also followed up in the field. When there were large discrepancies between the housing units on the two samples, indicative of possible geocoding errors, the block clusters were relisted for the P-sample.

E.1.e Last Step in Sampling: Reduce Housing Units in Large Block Clusters

After completion of housing unit matching and follow-up, the final step in developing the P-sample was to subsample segments of housing units on the P-sample address list in large block clusters in order to reduce the interviewing workload. The resulting P-sample contained about 301,000 housing units. Subsequently, segments of housing units in the census were similarly subsampled from large block clusters in order to reduce the E-sample follow-up workload. For cost reasons, the subsampling was done to maximize overlapping of the P-sample and E-sample. Table E.1 shows the distribution of the P-sample by sampling stratum, number of block clusters, number of housing units, and number of people.

E.2 P-SAMPLE INTERVIEWING

The goal of the A.C.E. interviewing of P-sample households was to determine who lived at each sampled address on Census Day, April 1. This procedure required that information be obtained not only about nonmovers between Census Day and the A.C.E. interview day, but also about people who had lived at the address but were no longer living there (outmovers). In addition, the P-sample interviewing ascertained the characteristics of people who were now living at the address but had not lived there on Census Day (inmovers).

The reason for including both inmovers and outmovers was to implement a procedure called PES-C, in which the P-sample match rates for movers would be estimated from the data obtained for outmovers, but these rates would then be applied to the weighted number of inmovers. The assumption was that fewer inmovers would be missed in the interviewing than outmovers, so that the number of inmovers would be a better estimate of the number of movers. PES-C differed from the procedure used in the 1990 PES (see Section 5–D.1).

It was important to conduct the P-sample interviewing as soon as possible after Census Day, so as to minimize errors by respondents in reporting the composition of the household on April 1 and to be able to complete the interviewing in a timely manner. However, independence of the P-sample and E-sample could be compromised if A.C.E. interviewers were in the field at the same time as census nonresponse follow-up interviewers. An innovative solution for 2000 was to conduct the first wave of interviewing by telephone, using a computerized questionnaire. Units that were eligible for

Table E.1 Distribution of the 2000 A.C.E. P-Sample Block Clusters, Households, and People, by Sampling Stratum (unweighted)

	Block Clusters		Households		People		
Sampling Stratum	Number	Percent	Number	Percent	Number	Percent	Average Households per Block Cluster
Small Block Clusters (0–2 housing units)	446	4.4	3,080	1.2	7,233	1.1	6.9
Medium Block Clusters (3–79 housing units)	5,776	57.6	146,265	56.7	386,556	57.8	25.3
Large Block Clusters (80 or more housing units)	3,466	34.6	102,286	39.6	253,730	38.0	29.5
Large and Medium Block Clusters on American Indian Reservations	341	3.4	6,449	2.5	21,018	3.1	18.9
Total	10,029	100.0	258,080	100.0	668,537	100.0	25.7

NOTES: Block clusters are those in the sample after all stages of sampling that contained one or more P-sample cases; households are those that contain at least one valid nonmover or inmover; people are valid nonmovers and inmovers. Outmovers are not included, nor are people that were removed from the sample.

SOURCE: Tabulations by panel staff from P-Sample Person Dual-System Estimation Output File (U.S. Census Bureau, 2001b), provided to the panel, February 16, 2001.

telephone interviewing included occupied households for which a census questionnaire (either a mail or an enumerator-obtained return) had been captured that included a telephone number, had a city-style address, and was either a single-family home or in a large multiunit structure. Units in small multiunit structures or with no house number or street name on the address were not eligible for telephone interviewing. Telephone interviewing began on April 23, 2000, and continued through June 11. Fully 29 percent of the P-sample household interviews were obtained by telephone, a higher percentage than expected.

Interviewing began in the field the week of June 18, using laptop computers. Interviewers were to ascertain who lived at the address currently and who had lived there on Census Day, April 1. The computerized interview—an innovation for 2000—was intended to reduce interviewer variance and to speed up data capture and processing by having interviewers send their completed interviews each evening over secure telephone lines to the Bureau's main computer center, in Bowie, Maryland.

For the first 3 weeks, interviewers were instructed to speak only with a household resident; after then, they could obtain a proxy interview from a nonhousehold member, such as a neighbor or landlord. (Most outmover interviews were by proxy.) During the last two weeks of interviewing, the best interviewers were sent to the remaining nonrespondents to try to obtain an interview with a household member or proxy. Of all P-sample interviewing, 99 percent was completed by August 6; the remaining 1 percent of interviews were obtained by September 10 (Farber, 2001b:Table 4.1).

E.3 INITIAL MATCHING AND TARGETED EXTENDED SEARCH

After the P-sample interviews were completed, census records for households in the E-sample block clusters were drawn from the census unedited file; census enumerations in group quarters (e.g., college dormitories, nursing homes) were not part of the E-sample. Also excluded from the E-sample were people with insufficient information (*II*s), as they could not be matched, and late additions to the census whose records were not available in time for matching. People with insufficient data lacked reported information for at least two characteristics (among name, age, sex, race, ethnicity, and household relationship); computer imputation routines were used to complete their census records. Census terms for these people are "non-data-defined," "whole person-allocations," and "substitutions;" we refer to them in this report as "whole-person imputations." In 2000, there were 5.8 million people requiring imputation, as well as 2.4 million late additions (reinstated records) due to the special operation to reduce duplication in the MAF in summer 2000 (see Section 4–E).

For the P-sample, nonmovers and outmovers were retained in the sample for matching, as were people whose residence status was not determined. Inmovers or people clearly identified from the interview as not belonging in the sample (e.g., because they resided in a group quarters on Census Day) were not matched.

E.3.a E-Sample and P-Sample Matching Within Block Cluster

Matching was initially performed by a computer algorithm, which searched within each block cluster and identified clear matches, possible matches, nonmatches, and P-sample or E-sample people lacking enough reported data for matching and follow-up. (For the A.C.E., in addition to meeting the census definition of data defined, each person had to have a complete name and at least two other characteristics). Clerical staff next reviewed possible matches and nonmatches, converting some to matches and classifying others as lacking enough reported data, as erroneous (e.g., duplicates within the P-sample or E-sample, fictitious people in the E-sample), or (when the case was unclear or unusual) as requiring higher-level review.[4] The work of the clerical staff was greatly facilitated by the use of a computerized system for searching and coding (see Childers et al., 2001).

On the P-sample side, the clerks searched for matches within a block cluster not only with E-sample people, but also with non-E-sample census people. Such people may have been in group

[4]Duplicates in the E-sample were classified as erroneous enumerations; duplicate individuals in a P-sample household with other members were removed from the final P-sample; whole-household duplications in the P-sample were treated as household noninterviews.

quarters or in enumerated housing units in the cluster that were excluded when large block clusters were subsampled.

E.3.b Targeted Extended Search

In selected block clusters, the clerks performed a targeted extended search (TES) for certain kinds of P-sample and E-sample households (see Navarro and Olson, 2001). The search looked for P-sample matches to census enumerations in the ring of blocks adjacent to the block cluster; it also looked for E-sample correct enumerations in the adjacent ring of blocks. The clerks searched only for those cases that were whole household nonmatches in certain types of housing units. The purpose was to reduce the variance of the DSE estimates due to geocoding errors (when a housing unit is coded incorrectly to the wrong census block). Given geocoding errors, it is likely that additional P-sample matches and E-sample correct enumerations will be found when the search area is extended to the blocks surrounding the A.C.E.-defined block cluster.

Three kinds of clusters were included in TES with certainty: clusters for which the P-sample address list was relisted; 5 percent of clusters with the most census geocoding errors and P-sample address nonmatches; and 5 percent of clusters with the most weighted census geocoding errors and P-sample address nonmatches. Clusters were also selected at random from among those clusters with P-sample housing unit nonmatches and census housing units identified as geocoding errors. About 20 percent of block clusters were included in the TES sample. Prior to matching, field work was conducted in the TES clusters to identify census housing units in the surrounding ring of blocks.

Only some cases in TES block clusters were included in the extended clerical search. These cases were P-sample nonmatched households for which there was no match to an E-sample housing unit address and E-sample cases identified as geocoding errors. When an E-sample geocoding error case was found in an adjacent block, there was a further search to determine if it duplicated another housing unit or was a correct enumeration.

Following the clerical matching and targeted extended search, a small, highly experienced staff of technicians reviewed difficult cases and other cases for quality assurance. Then a yet smaller analyst staff reviewed the cases the technicians could not resolve.

E.4 FIELD FOLLOW-UP AND FINAL MATCHING

Matching and correct enumeration rates would be biased if there were not a further step of follow-up in the field to check certain types of cases. On the E-sample side, almost all cases that were assigned a nonmatch or unresolved code by the computer and clerical matchers were followed up, as were people at addresses that were added to the MAF subsequent to the housing unit match. The purpose of the person follow-up was to determine if these cases were correct (nonmatching) enumerations or erroneous.

On the P-sample side, about half of the cases that were assigned a nonmatch code and most cases that were assigned an unresolved code were followed up in the field. The purpose was to determine if they were residents on Census Day and if they were a genuine nonmatch. Specifically, P-sample nonmatches were followed up when they occurred in: a partially matched household; a whole household that did not match a census address and the interview was conducted with a proxy respondent; a whole household that matched an address with no census person records and the interview was conducted with a proxy; or a whole household that did not match the people in the E-sample for that household. In addition, P-sample whole household nonmatches were followed up when: an analyst recommended follow-up; when the cluster had a high rate of P-sample person nonmatches (greater than 45 percent); when the original interviewer had changed the address for the household; and when the cluster was not included in the initial housing unit match (e.g., list/enumerate clusters, relisted clusters).

The field follow-up interviews were conducted with a paper questionnaire, and interviewers were instructed to try even harder than in the original interview to speak with a household member. After field follow-up, each P-sample and E-sample case was assigned a final match and residence status code by clerks and, in some cases, technicians or analysts.

E.5 WEIGHTING AND IMPUTATION

The last steps prior to estimation were to:[5]

- weight the P-sample and E-sample cases to reflect their probabilities of selection;
- adjust the P-sample weights for household noninterviews;
- impute missing characteristics for P-sample persons that were needed to define poststrata (e.g., age, sex, race); and
- impute residence and/or match status to unresolved P-sample cases and impute enumeration status to unresolved E-sample cases.

Weighting is necessary to account for different probabilities of selection at various stages of sampling. Applying a weight adjustment to account for household noninterviews is standard survey procedure, as is imputation for individual characteristics. The assumption is that weighting and imputation procedures for missing data reduce the variance of the estimates, compared with estimates that do not include cases with missing data, and that such procedures may also reduce bias, or at least not increase it.

For the P-sample weighting, an initial weight was constructed for housing units that took account of the probabilities of selection at each phase of sampling. Then a weighting adjustment was performed to account for household noninterviews. Two weight adjustments were performed, one for occupied households as of the interview day and the other for occupied households as of Census Day. The adjusted interview day weight was used for inmovers; the adjusted Census Day weight, with a further adjustment for the targeted extended search sampling, was used for nonmovers and outmovers. E-sample weighting was similar but did not require a household noninterview adjustment. Table E.2 shows the distribution of P-sample and E-sample weights.[6]

Item imputation was performed separately for each missing characteristic on a P-sample record. The census editing

[5]Cantwell et al. (2001) provide details of the noninterview adjustment and imputation procedures used.

[6]The weights were trimmed for one outlier block cluster.

Table E.2 Distribution of Initial, Intermediate, and Final Weights, 2000 A.C.E. P-Sample and E-Sample

Sample and Mover Status	Number of Non-Zeros	Percentile of Weight Distribution										
		0	1	5	10	25	50	75	90	95	99	100
P-Sample												
Initial Weight[a]												
Total	721,734	9	21	48	75	249	352	574	647	654	661	1,288
Nonmovers	631,914	9	21	48	76	253	366	575	647	654	661	1,288
Outmovers	24,158	9	21	48	69	226	348	541	647	654	661	1,288
Inmovers	36,623	9	21	47	67	212	343	530	647	654	661	1,288
Intermediate Weight[b]												
Total with Census Day Weight	712,442	9	22	49	78	253	379	577	654	674	733	1,619
Total with Interview Day Weight	721,426	9	21	48	76	249	366	576	651	660	705	1,701
Final Weight[c]												
Census Day Weight												
Total	640,795	9	22	50	83	273	382	581	654	678	765	5,858
Nonmovers	617,390	9	22	50	83	274	382	581	654	678	762	5,858
Outmovers	23,405	9	23	50	77	240	363	577	655	682	798	3,847
Inmovers	36,623	9	21	47	67	214	345	530	651	656	705	1,288
E-Sample												
Initial Weight[d]	712,900	9	21	39	55	212	349	564	647	654	661	2,801
Final Weight[e]	704,602	9	21	39	56	217	349	567	647	654	700	4,009

[a] P-sample initial weight, PWGHT, reflects sampling through large block subsampling; total includes removed cases.

[b] P-sample intermediate weight, NIWGT, reflects household noninterview adjustment for Census Day; NIWGTI reflects household noninterview adjustment for A.C.E. interview day.

[c] P-sample final weight, TESFINWT, applies to confirmed Census Day residents, including nonmovers and outmovers (reflects targeted extended search sampling); NIWGTI applies to inmovers.

[d] E-sample initial weight, EWGHT, reflects sampling through large block subsampling.

[e] E-sample final weight, TESFINWT, reflects targeted extended search sampling.

SOURCE: Tabulations by panel staff from P-Sample and E-Sample Person Dual-System Estimation Output Files (U.S. Census Bureau, 2001b), provided to the panel, February 16, 2001.

and imputation process provided imputations for missing basic (complete-count) characteristics on the E-sample records (see Appendix G). Finally, probabilities of being a Census Day resident and of matching the census were assigned to P-sample people with unresolved status, and probabilities of being a correct enumeration were assigned to E-sample people with unresolved enumeration status.

E.6 POSTSTRATA ESTIMATION

Estimation of the DSE for poststrata and the variance associated with the estimates was the final step in the A.C.E. process. The poststrata were specified in advance on the basis of research with 1990 census data (see Griffin and Haines, 2000), and each E-sample and P-sample record was assigned to a poststratum as applicable. Poststrata that had fewer than 100 cases of nonmovers and outmovers were combined with other poststrata for estimation. In all, the originally defined 448 poststrata, consisting of 64 groups defined by race/ethnicity, housing tenure, and other characteristics cross-classified by seven age/sex groups (see Table E.3), were reduced to 416, by combining age/sex groups as needed within one of the other poststrata.

Weighted estimates were prepared for each of the 416 poststrata for the following:

- P-sample total nonmover cases (NON), total outmover cases (OUT), and total inmover cases (IN) (including multiplication of the weights for nonmovers and outmovers by residence status probability, which was 1 for known Census Day residents and 0 for confirmed nonresidents);
- P-sample matched nonmover cases (MNON) and matched outmover cases (MOUT) (including multiplication of the weights by match status probability, which was 1 for known matches and 0 for confirmed nonmatches);
- E-sample total cases (E); and
- E-sample correct enumeration cases (CE) (including multiplication of the weights by correct enumeration status probability).

Also tabulated for each poststratum was the census count (C) and the count of IIs (people with insufficient information, including people requiring imputation and late additions). The DSE for each poststratum was calculated as the census count minus IIs, times the correct enumeration rate (CE/E), times the inverse of the match rate, or

$$(C - II)\left(\frac{CE}{E}\right)\left(\frac{P}{M}\right).$$

The match rate (M/P) was calculated for most poststrata by applying the outmover match rate (MOUT/OUT) to the weighted number of inmovers (IN) to obtain an estimate of matched inmovers (MIN), and then solving for

$$\frac{M}{P} = \frac{\text{MIN} + \text{MNON}}{\text{IN} + \text{NON}}.$$

However, for poststrata with fewer than 10 outmovers (63 of the 416), the match rate was calculated as

$$\frac{M}{P} = \frac{\text{MOUT} + \text{MNON}}{\text{OUT} + \text{NON}}.$$

Procedures were implemented to estimate the variance in the DSE estimates for poststrata. Direct variance estimates were developed for the collapsed poststrata DSEs that took account of the error due to sampling variability from the initial listing sample, the A.C.E. reduction and small block subsampling, and the targeted extended search sampling. The variance estimates also took account of the variability from imputation of correct enumeration, match, and residence probabilities for unresolved cases. Not included in the variance estimation were the effects of nonsampling errors, other than the error introduced by the imputation models. In particular, there was no allowance for synthetic or model error; the variance calculations assume that the probabilities of being included in the census are uniform across all areas in a poststratum (see Starsinic et al., 2001).

Table E.3 Poststrata in the Original 2000 A.C.E., 64 Major Groups

Race/Ethnicity Domain	Other Characteristics
1. American Indian or Alaska Native on Reservation[a]	❑ 2 groups: owner, renter
2. American Indian or Alaska Native off Reservation[b]	❑ 2 groups: owner, renter
3. Hispanic[c]	❑ 4 groups for owners: ❖ High and low mail return rate ❖ By type of metropolitan statistical area (MSA) and enumeration area ➤ Large and medium-size MSA mailout/mailback areas ➤ All other ❑ 4 groups for renters (see Hispanic owners)
4. Non-Hispanic Black[d]	❑ 4 groups for owners (see Hispanic owners) ❑ 4 groups for renters (see Hispanic owners)
5. Native Hawaiian or Other Pacific Islander[e]	❑ 2 groups: owner, renter
6. Non-Hispanic Asian[f]	❑ 2 groups: owner, renter
7. Non-Hispanic White or Some Other Race[g]	❑ 32 groups for owners: ❖ High and low mail return rate ❖ By region (Northeast, Midwest, South, West) ❖ By type of metropolitan statistical area and enumeration area: ➤ Large MSA, mailout/mailback areas ➤ Medium MSA, mailout/mailback areas ➤ Small MSA and non-MSA, mailout/mailback areas ➤ Other types of enumeration area (e.g., update/leave) ❑ 8 groups for renters: ❖ High and low mail return rate ❖ By type of metropolitan statistical area and enumeration area ➤ (See owner categories)

All 64 groups were classified by seven age/sex categories (below) to form 448 poststrata; in estimation, some age/sex categories were combined (always within one of the 64 groups) to form 416 strata.
Under age 18
Men ages 18–29; women ages 18–29
Men ages 30–49; women ages 30–49
Men age 50 and older; women age 50 and older.

NOTES: Large metropolitan statistical areas (MSAs) are the largest 10 MSAs in the United States; medium MSAs are other MSAs with 500,000 or more population; small MSAs are MSAs with less than 500,000 population.

The description of race/ethnicity domains is simplified somewhat; see Haines (2000) for complete set of classification rules (see also Farber, 2001a).

[a] All people on a reservation with American Indian or Alaska Native as their single or one of multiple races.

[b] All people in Indian Country not on a reservation with American Indian or Alaska Native as their single or one of multiple races; all non-Hispanic people not in Indian Country with American Indian or Alaska Native as their single race.

[c] All Hispanic people in Indian Country not already classified in Domain 2; all Hispanic people not in Indian Country except those living in Hawaii with Native Hawaiian or Other Pacific Islander as their single or one of multiple races.

[d] All non-Hispanic people with Black as their only race; all non-Hispanic people with Black and American Indian or Native Alaska race not in Indian Country; all non-Hispanic people with Black and another single race group, except those living in Hawaii with Black and Native Hawaiian or Other Pacific Islander race.

[e] All non-Hispanic people with Native Hawaiian or Other Pacific Islander as their only race; all non-Hispanic people with Native Hawaiian or Other Pacific Islander and American Indian or Alaska Native race not in Indian Country; all non-Hispanic people with Native Hawaiian or Other Pacific Islander and Asian race; all people in Hawaii with Native Hawaiian or Other Pacific Islander as their single or one of multiple races.

[f] All non-Hispanic people with Asian as their only race; all non-Hispanic people with Asian and American Indian or Alaska Native race not in Indian Country.

[g] All non-Hispanic people with White or some other race as their only race; all non-Hispanic people with White or some other race in combination with American Indian or Alaska Native not in Indian Country; or in combination with Asian; or in combination with Native Hawaiian or Other Pacific Islander not in Hawaii; all non-Hispanic people with three or more races (excluding American Indian or Alaska Native) in Indian Country or outside of Indian Country (excluding Native Hawaiian or Other Pacific Islander in Hawaii).

Appendix F

Methods for Treating Missing Data

F.1 OVERVIEW

The 2000 census, continuing with a pattern typical of all recent censuses, experienced an appreciable amount of nonresponse for characteristics information from households, especially on the long form. Overall, the rate of nonresponse for most items on the long form in 2000 was typically at least comparable to, and for many items was considerably greater than, that of previous censuses. Details on the frequency of item nonresponse are provided in Chapter 7. The Census Bureau, faced with this degree of nonresponse, again used a version of sequential hot-deck imputation, as described in Appendix G, and summarized below. The effectiveness of this treatment of nonresponse, in comparison with other methods currently available, and current techniques that might be used to estimate the variance of statistics derived from long-form-sample data in the presence of nonresponse, are examined in this appendix.

Nonresponse is usefully separated into two broad types: (1) *item* nonresponse, where a household provides responses for some but not all items on the questionnaire, and (2) *unit* nonresponse, where a household does not provide any information on a questionnaire.[1] Item nonresponse is often amenable to more sophisticated missing data methods because the responses that are available may be used to help predict the missing values. In contrast, for unit nonresponse, the only information available in the decennial census is the geographic location of the residence, and while geographic information is useful in somewhat the same way as responses to other items, it generally has limited value in comparison to information that is more specific to the household. The lack of information limits the techniques available to the Census Bureau, and as a result censuses have addressed unit nonresponse in the long-form-sample weighting process. This appendix is focused primarily on techniques that address item nonresponse, though some of the discussion will also apply to techniques for addressing unit nonresponse.[2]

Untreated, nonresponse can cause two problems. First, nonresponse can result in statistical bias. (Statistical bias is a measure of the difference in the expected value of a statistic and its true value.) It is generally the case that the data that one receives from respondents are different distributionally from the data that would have been provided by nonrespondents. This is why so-called complete-case analysis is problematic, since the restriction of the analysis to those cases that have a complete response fails to adequately represent the contribution from those that have missing data. These distributional differences may be present not just unconditionally, but (often) also conditionally given responses to certain items. For example, data for nonrespondents may be different from data for respondents because the respondents have a different demographic or socioeconomic profile. Even within demographic or socioeconomic groups or other conditioning variables that one might choose, data for nonrespondents may still have a different distribution than that for respondents. Methods that do not take these differences into account can introduce a statistical bias, which can be appreciable

[1]In many applications the unit is defined as a person, but in this application, the unit is defined as a household. Thus, a person may have no data provided for them, but as long as other members of the household provided information, that would still be referred to in this application as item nonresponse.

[2]The process of weighting the long-form sample to the complete-count totals relies on the assumption that cases of unit nonresponse are missing at random (defined below).

with substantial amounts of nonresponse. For example, if nonrespondents tend to commute greater distances than respondents, methods to treat nonresponse that are not sensitive to this will produce estimates of average commuting distance that are generally too low.

Second, nonresponse often is either ignored or not fully represented in estimates of the variances of statistics computed using data sets with missing data, which will result in variance estimates that are statistically biased low. In other words, improper treatment of missing data in the estimation of variances will result in statistics that are represented as being more reliable than they in fact are, which will cause hypothesis tests and confidence intervals based on the (treated) data to understate the variation in the data. As will be discussed below, the current treatment of nonresponse in the long-form sample does result in estimated variances that are likely too low. This is a challenging problem with no easy solutions, but there are new approaches that are extremely promising that are discussed below.

Generally speaking, there is considerably more nonresponse to the additional long-form items than to the basic items asked of everyone. In addition, the relative lack of useful correlations among the basic variables reduces, though it does not eliminate, the benefits one may derive from some of the techniques described here. Therefore, the panel believes that the selection of treatments to address nonresponse for basic items does not have as much urgency as that for the additional long-form items, and as a result, we focus in this appendix on missing value treatments for the long-form sample. However, techniques that are found useful for treatment of nonresponse in long-form-sample items also should be examined for their potential benefits in treating nonresponse in the basic items on the short and long forms.

Before continuing, it is important to stress that the bias due to missing data is often not measurable, or only measurable using relatively costly efforts at additional data collection, including reinterviews, or through use of alternative sources of information such as administrative records. Even state-of-the-art techniques for treatment of missing data are ultimately limited in their usefulness either by increased amounts of nonresponse or by lack of information concerning the nature of the nonresponse. Therefore, the highest priority is to reduce the extent of missing data. There is a great deal of work by cognitive survey scientists in motivating people to provide data, and efforts should obviously continue along these lines, possibly more so due to the higher degree of nonresponse to the long form in the 2000 census. (Fortunately, the proposed American Community Survey, a continuous survey designed as a replacement for long-form-type information, being continuous is more amenable to this type of research.) The techniques in the literature are meant to address at most moderate amounts of nonresponse; they are not intended for situations in which the nonresponse is on the order of 30 percent or more. (There are examples in the literature in which these techniques have been applied to more extreme cases; however, in those cases the reliability of the resulting estimates was not intended to be of the quality typically expected for census output.)

F.1.a Mechanisms for Nonresponse

When considering nonresponse, it is important to try to understand the likely mechanism underlying the nonresponse, since this has important implications for the methods used to treat it. This is obviously more difficult for unit nonresponse due to the lack of available information, but any progress in this direction is useful. Mechanisms underlying nonresponse can be roughly classified into three categories:

(1) missing data are *missing completely at random*—in this case the indicator variable for whether or not a response is provided is independent of all other data for the household, whether collected or not, and therefore, cases with missing values have the identical distribution as those with complete responses;

(2) missing data are *missing at random*—in this case the indicator variable for response is independent of missing values, but could be dependent on responses, and therefore the missing values are, *conditional on responded information*, distributed the same as responses; and

(3) missing data are *not missing at random*—in this case the indicator variable for response is dependent on missing values, and therefore, even conditionally on collected data, the distribution of missing values will depend on the values of uncollected data.

This third category is clearly the most difficult situation to be in. It is important to examine which situation is likely to be in effect for missing values for various items on the census questionnaire in order to identify the proper treatments to minimize bias. For instance, are people at the high end of the income spectrum likely to omit their responses to income questions, or is this more typical of those at the low end of the spectrum? Is nonresponse for income more or less common for those in various demographic groups, or in various occupations, or in various geographic regions?

Consider a situation in which responses are missing for some households for commuting distance. If these missing values are missing completely at random, using a hot-deck procedure to substitute random commuting distances from other respondents would not result in a statistical bias for assessment of average commuting distance. On the other hand, if these missing values are missing at random but are not missing completely at random—in particular, that conditional on income, nonresponse for commuting distance is independent of other items on the census questionnaire—then using a hot deck procedure to substitute random commuting distances from other respondents with the same income level would not result in a statistical bias. Relevant to the current Census Bureau missing data procedure, if it is the case that many variables are somewhat homogeneous over small geographic distances, and the data are missing at random by conditioning on possibly unknown covariates, forcing a hot deck to make use of only nearby respondents for donor values can be a sensible procedure since it implicitly conditions on all other data for the household. However, this assumption does not always even approximately obtain.

If the missing values are not missing at random, which is referred to as *nonignorable nonresponse*, while there are approaches that can be tried, they are very dependent on assumptions and much more difficult to implement in a large, multidimensional data set like the decennial census. So, if those people with very long commuting distances fail to respond as often as those with shorter commutes and nothing else is known about the relationship of commuting distance to other variables, then there may be nothing that can be done to address the resulting bias in the context of decennial census output.

Of course, in real applications, cases do not fall neatly into the above categories. One could imagine a response indicator variable that had a distribution that was predominantly missing at random but that had a relatively small contribution to its distribution that was nonignorable. In that case, making use of a method that was based on the assumption that the data were missing at random would provide a substantial reduction in bias, though it would fail to eliminate it.

A short summary is as follows: (1) if the missing values are missing completely at random, any sensible missing data method will be effective; (2) if the missing values are missing at random, taking into consideration the variables that differentiate the respondents from the nonrespondents is the key to proper treatment of the nonresponse; and (3) if the nonresponse mechanism is not ignorable, then accounting for variables that explain some nonresponse is likely to reduce bias as much as possible without the use of arbitrary assumptions.

With respect to variance estimation, it is still typical for government agencies, based on survey data, to publish variance estimates that understate the variance present in the data due to nonresponse. Specifically, standard complete-data-set variance estimates applied to data sets with missing data values filled in by imputation obviously fail to include the contribution to variance from the variability in the imputations. In addition, failure to correctly model the nonresponse, say by using the assumption that the missing data are missing completely at random when they are only missing at random, or by conditioning on the wrong variables when assuming that missing values are missing at random, not only causes bias in the estimates, but also causes bias in estimates of the variance.

F.1.b Implementation Considerations

A key constraint concerning treatment of missing values in the decennial census is that the Census Bureau needs to provide data products (population counts, cross-tabulations, averages, and public use microdata samples) that can be used by the wide variety of census data users for their varied purposes. Certainly, if one was only interested in the treatment of missing data for a single response to estimate a national average, one could make use of a method optimized for that situation, and that method might be very sophisticated and computationally intensive, given the focus of the problem. However, the Census Bureau cannot do this for each of the myriad potential uses of decennial census data, and it cannot ask its data users to implement their own treatments for nonresponse. This is because it is the job of the Census Bureau to provide official products from the decennial census and, more pragmatically, because access to individual-level data—needed to implement most procedures—is precluded by Title 13 of the U.S. Code. An exception is that of public use microdata sample files (PUMS), but the ability of PUMS users to carry out their own missing data treatment is severely complicated by the reduced sample size that is available to them. What is needed is a treatment that can be used for a large, multidimensional data set that can accommodate the needs of a wide variety of analysts. (Nevertheless, the Census Bureau should continue the current practice of flagging data values that result from the use of treatments for nonresponse.) Understanding this, the Census Bureau has used, in the most recent censuses, an omnibus treatment of nonresponse that is effective for a large database with a considerable number of response variables (and variable types, i.e., continuous, discrete, and categorical) that supports a wide range of analyses. Any new approach must retain this characteristic.

The multidimensional character of the decennial census data set raises an important issue. There are many situations in which the values provided (or that would have been provided, absent nonresponse) for variable *A* are correlated with the values for variable *B*, and vice versa. Often, many more than two variables are involved in these dependencies. In order to minimize bias, assuming the nonresponse is missing at random, it is important to use all the available information that is relevant for imputing values. A multivariate approach to treatment of missing values therefore can have a distinct advantage over a univariate approach, since then the dependence between variables jointly missing can be properly taken into account.

In addition to the multivariate aspect of missing data, especially in the long-form sample, there are computational considerations that have historically helped to select the missing data treatments. Certainly, given the memory size and processing speed typical of the best widely available main frame computers in the 1980s, keeping the entire decennial census data base in active memory, or even the database for large states, and repeatedly accessing individual data elements, was outside the bounds of what was generally feasible. Therefore, what was needed was a methodology for the treatment of nonresponse that kept a large fraction of the database in remote memory while processing only small pieces of the database sequentially. This was one of the arguments in support of the use of sequential hot-deck imputation. However, computational power and speed have changed dramatically in the last decade. Assuming Moore's Law is even approximately true, the entire 2010 decennial census database (even if the long-form sample is included in the 2010 census) will almost certainly be able to be manipulated in 2011 in the active memory of the more powerful desktop computers. Therefore, for future censuses, the treatment for missing data does not need to be sequential anymore, and treatments can be considered that require multiple access to the database.

F.2 OUTLINE OF THE CURRENT METHODOLOGY

For the past five decennial censuses, the Census Bureau has used variants of the following treatment for nonresponse in the long-form sample (see Appendixes G and H for further details). With respect to unit nonresponse, as mentioned above, the Census Bureau treats this as an additional sampling mechanism. Iterative proportional fitting is used to weight long-form-sample population counts to complete population counts as a variance reduction technique that also treats unit nonresponse.

With respect to item nonresponse, the Census Bureau makes use of sequential hot-deck imputation. (The quick description here ignores the issue of how this technique is initialized and other details.) In this procedure, the long-form data set for each state is processed in one pass through the data. Assume that a given household has missing values for variables X_1, X_2, and X_3, and assume also that the Census Bureau, using its accumulated imputation rules, believes that to predict a household's responses for variables X_1, X_2, and X_3 it is useful to condition on the values for variables C_1, C_2, C_3, and C_4 consistent with a missing-at-random assumption. In processing the long-form data file, a large number of one-way, two-way, three-way, etc., matrices are constantly updated. The matrices are made up of cells that contain the data from the most recently processed household with complete information that match the specific combination of values for the variables defining the matrices' dimensions. (In addition, the census file is sorted by some characteristics not included in the matrix, providing data that are for households close geographically as well as similar for their characteristics.) This essentially provides a geographically proximate donor housing unit with identical information for these matching variables for a current household with item nonresponse. Then the values for variables X_1, X_2, and X_3 from the donor household occupying that cell are used to fill in the missing values for the current household.

This description is somewhat inadequate in two ways. First, there are also imputation rules that are used in the imputation of numeric responses. For example, there may be additional variables not used in the matching, say, A_1, A_2, and A_3, that have values that, when combined with the imputed values for the variables X_1, X_2, X_3, and with the matching variables C_1, C_2, C_3, and C_4, result in unusual or impossible combinations. Adjustments are therefore made to handle these situations. Second, the choice of variables to condition on can depend on the values of other collected variables. That is, there may be a variable Z for a household, which may take on some values that would require matching on C_1, C_2, and C_3 to identify the donor, and other values that would require matching on C_1, C_2, and C_4. So the imputation can become fairly elaborate.

Besides the benefits of using a geographically (and other characteristics) proximate responding household that is likely to share many nonmatched characteristics with the current household, this process is quick to implement since it only requires a single pass through the data file (barring initialization). It also permits conditioning on important variables, thus avoiding the assumption that nonresponses are missing completely at random and only relying on a specific missing-at-random assumption.

The U.S. Census Bureau's sequential hot-deck imputation methodology is somewhat related to that used in the Canadian Census, referred to either as nearest-neighbor imputation or new imputation methodology (NIM). As described in Bankier (1999) and Bankier et al. (2002), and originally based on the work by Fellegi and Holt (1976), missing values are filled in through use of a nearest neighbor donor, where nearest neighbor is defined by (1) the donor household lies within the same geographic region as the household in need of imputation, and (2) given households satisfying (1), the donor household must minimize a metric measuring closeness of the donor and the current household with respect to a number of selected matching variables. As above, there is the complication that variables not used in the matching in conjunction with both the variables used for matching and the imputations may collectively fail a number of edits that households are required to pass. An imputation that permits a household to pass all edits is referred to as a feasible imputation, and among the nearest neighbors, a random feasible imputation is selected for use. In a very rough sense, NIM trades off the computational benefits of the sequential hot-deck procedure by looking around locally for a better match. However, this "better" match may be more distant geographically, in which case it is also poorer with respect to an important variable, geography.

Finally, with respect to variance estimation, as mentioned above, the weighting resulting from unit nonresponse is subsumed into the sampling weights. Therefore, the sampling variances that are published for population counts and other summary statistics, such as means, appropriately represent the contribution to variance from unit nonresponse, assuming the missing-at-random assumption is reasonable. However, no data products from the decennial census long-form sample include an estimate of the contribution to variance from item nonresponse.

F.3 PROBLEMS WITH THE CURRENT METHODOLOGY

The current methodology has two basic deficiencies that can be at least partially addressed by techniques that have recently been proposed in the statistical research literature and, to some extent, implemented in practice. First, the current methodology of sequential hot-deck imputation is somewhat inefficient due to its reliance on a single donor for imputation instead of using more of the local information, including households that do not match. The basic problem is that data from an individual donor household can be odd. Second, the failure to represent the variance from item nonresponse in data products is a serious problem for census data users because it fails to provide them with accurate information concerning the quality of the information they are using.

More specifically, with respect to the first deficiency, a single-donor imputation is subject to the vagaries of individual data elements. Certainly, a particularly unusual household would be a poor donor because it would provide in essence a very unlikely prediction for the household with missing values. Along the same lines, a very typical donor would not necessarily provide a good imputation for an unusual household with missing data. Through use of more of the information in an area, a better sense of the relationships between variables and the remaining uncertainty can be attained, providing a better imputation. Using the information from a local area has been precluded until recently by the computational demands that such use would entail. However, advances in computing make a number of methods for using this information quite feasible.

A related issue is the necessarily multivariate nature of imputation. We separate this into two issues. First, even a single-variable imputation is dependent on other values provided by the household. The sequential hot-deck procedure addresses this need for multivariate dependence through use of the matching variables and the various imputation rules. For dependencies that are essentially deterministic, this approach is likely reasonable. Certainly, the computer code that specifies which variables should condition the imputation of which other variables, and which rules are useful for insuring that imputations do not violate any understanding of how variables interrelate, represents an important knowledge base that needs to be preserved (though we are curious about how this knowledge base grows—and is pruned—over time and whether that process can be improved).[3] On the other hand, some of these specifications for conditioning and some of these rules are probably not uniformly obeyed and should probably be applied more probabilistically through use of a statistical model (see, e.g. Chiu et al., 2001). Further, it would be interesting to examine whether the selection of matching variables and the use of imputation rules would benefit from being applied subnationally rather than nationally.

[3]In addition, there is a software engineering issue as to how these computer-based rules, created relative to questionnaires that change from census to census, can be quickly and accurately translated to rules for subsequent censuses.

The second multivariate issue concerns simultaneous imputation of more than one variable (X_1, X_2, and X_3 in the notation above). While the current procedure used by the Census Bureau does impute some variables simultaneously, there are situations in which the sequential hot deck does not impute in a way that would provide proper estimates of the correlation between variables. Roughly speaking, relationships for variables that are closely related (e.g., income questions) may be handled in a manner that preserves correlations, but correlations between, say, income and education questions will not likely not be properly accounted for. This is because the hot deck uses strata based on cross-classification, which sharply limits the number of characteristics that can be considered to a small number, when it is known that there is information in other variables. This motivates the methods described below.

Finally, the most important deficiency with the current method that the Census Bureau uses to address item nonresponse is the failure to represent this nonresponse in its estimates of the variance of its data products. This failure is probably negligible for much complete-count-based statistical output, but will not be negligible for many long-form-sample-based statistics nationally, and for many other long-form-sample statistics for local areas or subpopulations. For the purposes of this discussion, the decennial census data products can be categorized into four types: (1) population counts, (2) cross-tabulations, (3) means and other summary statistics, and (4) public use microdata sample files. Until recently, the only technique that was available for estimating the contribution of nonresponse to variance was probably multiple imputation (see Little and Rubin, 1987). Multiple imputation could have been applied in 2000 to estimate the variance for population counts and various summary statistics. However, providing all users with, say, five replicate cross-tabulations or five PUMS files was probably not feasible. However, large amounts of disk space are becoming so cheap and widely available and this technique is sufficiently easy to apply that this option will certainly be more acceptable in the near future.

F.4 NEW APPROACHES TO IMPUTATION AND THEIR ADVANTAGES

While the current methodology used to treat missing data in the long-form sample was quite effective, given the various requirements and implementation constraints discussed above, for the 1980 and 1990 and possibly the 2000 censuses, there remains room for improvement through use of a number of techniques that have been recently proposed for use in this area.

The fundamental idea is to try to make greater use of data from households in the local area to assist in the imputation while retaining the benefits currently expressed in the choice of variables for conditioning and the imputation rules. A first step away from the current technique would be to retain the notion of the use of matching donors, but make use of information from more than the last household processed. An idea that would be relatively easy to implement is that of fractionally weighted imputation, proposed by Fay (1996), in which for some small integer m, m matching donors provide imputations, which are averaged, each receiving weight $1/m$. One benefit of this technique is that the Rao and Shao (1992) variance estimation procedure can be directly applied to the resulting database for a large class of estimators (though this variance estimator has limitations described below). Further, most of the software currently used could be retained. The choice of m involves an interesting trade-off between geographic proximity and stable estimation. To implement this technique, it would probably not be reasonable to simply expand the imputation matrices to retain m donors, since this would likely restrict the donors to "one side" of the current household. Instead, it would be preferable to search the area for the best m donors.

A second step would be to try to use data from local households that do not match through use of statistical models. The introduction of statistical models into long-form item imputation is probably the direction that needs to be pursued for substantial improvement (see Box F.1). One attempt at use of statistical modeling for imputation has been carried out by Chiu et al. (2001). More directly relevant for the issues considered here is the work by Thibaudeau (1998)) and Raghunathan et al. (2001).

In Thibaudeau (1998), basic-item imputations, which are simply categorical variables indicating membership in demographic subgroups and whether the residence is owned or rented, are generated for data collected in the 1998 census dress rehearsal in Sacramento, California. Thibaudeau uses a hierarchical log-linear model, fit at the tract level, which is composed of several two-way interactions of variables for a household and its immediate "neighbor" on the

Box F.1 Model-Based versus Model-Free Treatments for Imputation

Decisions as to which of many algorithms to use for imputation will involve the trade-off between an algorithm such as sequential hot-deck imputation in which no explicit model is used to identify a possible value for imputation, versus an algorithm that uses an explicit model. It is useful to examine some of the trade-offs and related considerations comparing and contrasting these two approaches.

Sequential hot-deck imputation selects donor households by requiring them to match recipient households on several variables, in addition to being closely proximate geographically. The matching implicitly assumes that the matching variables, and all of their interactions, are important for predicting the variable (variables) with missing values. Unfortunately, the requirement of a match may result in use of a donor household not as proximate geographically. Further, as mentioned above, reliance on data from one household for the imputation is inefficient, whereas model-based methods generally provide fairly well-behaved predictions. However, inefficiency of a single donor approach can be readily addressed by using a procedure such as Fay's fractionally weighted imputation.

A model-based algorithm would propose that non-matching households within some geographic area be used to help fit parameters from some statistical model that would specify how the various variables are related to a variable with missing data. The hope would be that these models would be able to rely on a small number of parameters, since they may only have main effects and interactions, not using the higher-order interactions implicit in the matching.

With respect to geography, an empirical question is which type of algorithm would use households less proximate to the household of interest. A reasonable guess is that even with fractionally weighted imputation, one will usually need to use households further away with a model-based approach, though with the use of non-matching donors, it is conceivable that one may be able to fit the parameters using households no more distant than the matching donor used in a sequential hot deck.

The obvious benefit from the use of a model rather than a proximate matched donor (or donors) is that models borrow more information to help reduce variance. However, there is a real opportunity for introduction of bias. So we have a standard bias versus variance trade-off. With a proximate matched donor, the imputation is likely to have minimal bias—unless the matching variables are poorly chosen—but it can have a substantial variance. The reason for the opportunity for additional bias with the model-based approach is that a model can provide poor imputations not only if the covariates are chosen poorly (in the same way as if matching variables are chosen poorly), but also if the model relating how the variables for imputation are related to the covariates is seriously wrong. Therefore, some efforts at model validation are needed, but unfortunately there are a lot of models here. For example, Thibaudeau fit his models at the tract level, and the United States has approximately 65,000 tracts.

Box F.1 (continued)

There are two final smoothing possibilities that this problem raises. First, it would be bothersome if all of the local models used for imputation had strongly different parameter values. One would expect to see slow changes in the parameters as one moved from one area to another. Therefore, one interesting possibility would be to smooth the estimated parameters to each other after fitting, by using some combining information approach. Another smoothing idea is to try to blend the benefits of a sequential hot-deck imputation with a model-based one, by taking some linear combination of the two approaches.

census file. Very roughly speaking, this model has similar structure to sequential hot-deck imputation conditioned on matching variables, since these two-factor interactions are similar to the use of the same variables in setting up the various matrices containing donor records for neighbors. Imputations and parameter estimates are produced jointly using a form of the EM algorithm (see Box F.2), specifically, the data augmentation Bayesian iterative proportional fitting algorithm (DABIPF) found in Schafer (1997). The parameter estimates are interesting to examine to understand the local nature of various main effects and interactions in producing the imputations. Thibaudeau's framework also provides immediate variance estimation through use of the posterior predictive distribution from his model, which also provides a basis for evaluating his procedure in comparison with sequential hot-deck imputation. The entire fitting process took 12 hours on a computer that can currently be emulated by many standard desktop computers; however, the population size of the dress rehearsal was only one-two thousandths that of the United States. Still, it is very reasonable to believe that this approach will be feasible before the 2010 census if the long form is used, and almost immediately for the American Community Survey. Schafer (1997) provides a more general approach that could be used to provide imputations for discrete and continuous data as well as categorical data. The main question is whether a reasonably robust model could be identified that would provide good imputations for long-form items.

In Raghunathan et al. (2001), an imputation procedure is described that has been implemented recently in the SAS software package as part of the IVEware system. This method is also closely related to the methods described in Schafer (1997). Using a Bayesian

Box F.2 EM Algorithm

The EM algorithm (E is for expectation, M is for maximization of likelihood) is a broadly applicable method for providing maximum-likelihood estimates in the presence of missing data. The EM algorithm was identified and initially investigated in its general form by Dempster et al. (1977). (It had been applied in several different missing-data situations for decades without the common structure of the techniques being recognized.) Begin with initial values for the parameters of the distribution generating the data and the sufficient statistics for this distribution. (Sufficient statistics are often small-dimensional functions of the data that allow one to estimate the parameters of a distribution.) Then the EM algorithm can be informally described, for well-behaved data distributions, as an iterative application of the following two steps, continuing until convergence:

> **E-Step:** Fill in missing portions of the sufficient statistics due to the missing data with their expectation given the observed data and the current estimated values for the parameters of the data distribution
>
> **M-Step:** Using these estimated sufficient statistics, carry out a standard maximum-likelihood calculation to update the estimates of the parameters of the data distribution.

So, very crudely, the parameters help one to identify the contribution of missing data to the sufficient statistics, and then the parameters are reestimated given the updated estimated sufficient statistics.

framework, explicit models for selected parameters and data are used, relating an individual response variable to missing values, conditional on the fully observed values and unknown parameters. Bayesian simulation is used to update a noninformative prior to form a posterior distribution for both the parameters and the missing values. Imputations are therefore draws from the posterior predictive (marginal) distribution for the missing values. The models are regression-type models, including logistic, Poisson, and generalized logit, to be able to fit categorical, discrete, and continuous data-type variables. In addition to the regression models, IVEware can accommodate restrictions of the regression models to relevant subpopulations, especially including local areas, and the imposition of logical bounds or constraints for the imputed values.

In the fitting process, variables are imputed individually, conditionally on all other observed and imputed variables at that point in the computation. However, Raghunathan et al. (2001) argue that by cycling through the fitting process variable by variable, the imputations can properly represent their dependence structure. As will be discussed below, Raghunathan et al. (2001) propose repeating the process of randomly drawing from the posterior predictive distribution for the missing values to form multiple imputations for variance estimation. The computational demands of IVEware are considerable, and efforts are being made to implement modifications that could permit use on data sets of the size of the ACS.

It is clear that the area of imputation has a number of promising research avenues well underway. Given that, it would be reasonable to predict that some form of EM imputation mechanism that was feasible computationally for either the long-form sample or the ACS and preferable to sequential hot-deck imputation could be identified well before the 2010 census with concentrated research efforts.

F.5 NEW APPROACHES TO VARIANCE ESTIMATION AND THEIR ADVANTAGES

Along with the recent work on alternatives to sequential hot-deck imputation, there has been promising work done recently on estimating variances for statistics computed from data with missing values.

The first suggestion for an omnibus approach to the estimation of variances was Rubin's proposal of multiple imputation. (For details, see, e.g., Little and Rubin, 1987.) The basic idea is to repeat the process for producing imputations for missing values in a data set m times, where the imputation mechanism can be represented as draws from a posterior predictive distribution for the missing values. The m data sets, "completed" through use of one of the m imputations, are each used to estimate a statistic S and its variance, where the variance is computed assuming that the imputations were collected data. The variance of the statistic S is estimated by separately estimating the within-imputation and the between-imputation contributions to variance, as follows:

$$\widehat{\mathrm{Var}(S)} = \sum_{i=1}^{m} \frac{\mathrm{WithinVar}_i}{m} + \left(\frac{m+1}{m}\right)\mathrm{BetweenVar}_i,$$

where $\mathrm{WithinVar}_i$ is the standard complete-data variance of S_i, S_i is the estimate for S computed from the ith completed-data set, and $\mathrm{BetweenVar}_i$ is the variance of the S_i's, (namely $\frac{1}{m-1}\sum_{i=1}^{m}(S_i - \bar{S})^2$

where $\bar{S}$ is the average of the S_i). Multiple imputation has been implemented in IVEware and in the software described in Schafer (1997).

A great deal of research has been carried out examining the properties of multiple imputation (see Rubin, 1996). However, to date it is not widely used by federal statistical agencies, and in particular, it has not been used to estimate variances for data output from the long-form sample. One primary reason for this is the need to provide users with *m* versions of a large data base (or cross-tabulation). However, the greater availability of computer memory and the wide acceptance by users of acquiring data sets electronically rather than in paper version reduce the relevance of this criticism.

Another objection to the use of multiple imputation was raised by Fay (1992). Fay discovered that when the assumptions for the model underlying the imputation mechanism and the assumptions underlying the model used by the analyst on the data set with the imputations are different, multiple imputation can produce erroneous estimates of variance. A number of people have contributed to this discussion to determine how likely the situation identified by Fay is to occur in practice and what the magnitude of the bias would be, including Binder (1996); Eltinge (1996); Fay (1996); Judkins (1996); Meng (1994); Rubin (1996). We believe that imputation models exist that are suitably general to accommodate many analyses of the data without biasing the variance estimation substantially. However, this is both a theoretical and an empirical question that is worth further investigation in the census context.

Several approaches have been proposed that might avoid these two primary objections to multiple imputation. A leading alternative approach has been proposed by Rao and Shao (1992), which is based on a modification to the leave-one-out jackknife that accounts for the effect of each collected data value on the mean of the imputed values, when the objective is to estimate the variance of a mean (though this technique works for more general types of estimators). As pointed out by Judkins (1996), though Rao-Shao avoids the Fay difficulty, the algorithm is inherently univariate and therefore is not a candidate for an omnibus approach to variance estimation for the long-form sample.

Kim and Fuller (2002) have a specific proposal that is also very promising for directly estimating the variance of statistics computed from data sets with missing values, in particular for long-form-type applications. Their approach has a great deal of appeal, but it seems somewhat limited in its ability to handle highly multivariate patterns of nonresponse. This would exclude, for example, most small-area estimators. Also, it is not clear what the computing demands are for Kim and Fuller's algorithm and therefore whether it could be implemented on as large a problem as the decennial census long-form sample.

One may wonder why the posterior predictive distribution for the imputations created from application of the EM algorithm could not be used to estimate variances due to each imputation and thus the variances for statistics computed using the imputations (as was done by Thibaudeau, 1998). In fact, this could be carried out, at least approximately, for most situations, for specific statistics of interest. However, as mentioned above, the objective is an omnibus approach to variance estimation that can be used for any statistics of interest to a census data user. For example, suppose a public use microdata sample user wanted to fit a small-area model of average income by area, explained by other area averages, and was particularly interested in the significance of percent female head of household as a covariate in this model. Use of the posterior predictive variances would still require that a considerable amount of analytic work be carried out by this user in order to compute variances of the parameter estimates in this model. Similarly, there is an issue as to how posterior predictive variances should be communicated to users interested in fitting log-linear models to census cross-tabulations.

Given the degree of item nonresponse in the 2000 census long-form sample, the issue of variance estimation to incorporate the variance due to missing values is the key missing value problem facing the Census Bureau heading into either the implementation of a 2010 long-form sample or the ACS. As discussed here, there are some very exciting research efforts now ongoing that might be feasible by 2010. Even if none of the newer possibilities become fully capable of being implemented on either a 2010 long-form sample or ACS, multiple imputation, if used carefully, should be strongly preferable to the current practice of ignoring the contribution to variances from item nonresponse.

F.6 SUGGESTIONS FOR WORK FOR THE AMERICAN COMMUNITY SURVEY AND THE 2010 CENSUS

To conclude the above discussion, a number of suggestions are proposed for work leading up to either the 2010 census or to the full implementation of the American Community Survey, depending on how long-form-sample information is to be collected in the immediate future. The Census Bureau should:

1. Examine patterns of nonresponse in the 2000 long-form sample to see the extent to which assumptions about data either missing completely at random or missing at random are justified. This can be accomplished through use of reinterview studies (a 2000 reinterview study to measure response variance was carried out; see Singer and Ennis, 2003) and through use of matching to other data sources, including administrative records.
2. Examine the current quality of the sequential hot-deck imputation. This can be accomplished in two steps. First, to see whether hot-deck imputation accurately mimics the data generation mechanism for respondents, simulate with 2000 census files the random omission of collected data and examine the quality of the resulting imputations in comparison to the values omitted. Second, to see whether the mechanism for census long-form-item nonresponse is ignorable, use reinterview studies and matching studies to compare the sequential hot-deck imputations to collected data through these sources. If this is not feasible, late-arriving census forms might be considered to be intermediate between responses and missing values, and imputations could be simulated for and compared to those values. Such analyses should distinguish between data provided by household members and proxy responses.
3. Implement a comprehensive procedure for validation of imputation rules. As mentioned previously, the Census Bureau has constraints on imputations to observe various relationships with other collected and imputed information. The question is whether collected data observe these rules, and whether there are new rules not currently in place that should be added. This is essentially a data mining problem (though obviously some field work would be beneficial to decide whether data that failed edits were valid), and there is a growing body of techniques, such as

classification and regression trees, and other forms of machine learning that, when applied to the decennial census long-form sample, could be used to examine the validity of current rules and the benefits of additional ones.

4. Initiate a comprehensive program pointed toward 2010 to examine whether some modification of the work of Thibaudeau (1998) or other EM algorithm-type approaches to long-form-item imputation are feasible and whether they provide superior imputations in comparison to sequential hot-deck imputation.

5. Initiate a comprehensive program pointed toward 2010 to examine whether some type of multiple imputation process would be feasible to incorporate the variance due to item nonresponse in the long-form sample or in the ACS. The goal would be to publish estimated population counts and other summary statistics (e.g., means and frequencies) with standard confidence intervals computed using variances resulting from implementation of multiple imputation, and also to release in electronic form multiply imputed versions of cross-tabulations and PUMS for the same purpose.

Finally, with respect to shifting long-form-type data collection from the decennial census to the ACS, it is useful to point out that the missing data problem is somewhat different for that survey. The ACS, having a lower sampling rate than the long form, has fewer neighbors in close proximity to a household with item nonresponse (and the rate of item nonresponse is likely to be lower given experience with the Census 2000 Supplementary Survey; see Section 7–C.2). This argues more strongly for a model-based approach. Also, the data file is somewhat smaller, putting fewer computational demands on an algorithm. Further, there is the possibility of stratifying responses for imputation by mode of response (telephone vs. mail). Given this, the algorithms that are feasible and optimal for these two problems may not be identical. But a first guess would be that the problems are sufficiently similar so that procedures useful for the long-form sample would be useful for ACS. An early decision on which approach is to be taken for collection of long-form information in 2010 would help focus research. Lastly, the continuous nature of the ACS would make it easier to carry out research and experimentation on possible imputation methods.

APPENDIX G

2000 Census Basic (Complete-Count) Data Processing

The 2000 census, like every census since 1940, asked some basic items of every household and person in the nation, with other items asked only of a sample. In 2000, these basic data items were:

- age and date of birth,
- sex,
- race,
- ethnicity (Hispanic origin),
- relationship to reference person (first person listed on the questionnaire), and
- housing tenure (own or rent).

In addition, name (first, last, and middle initial) was captured as a basic item in 2000. (Vacant units were also included in the census and a few items collected for them.) The basic census items make up the short-form questionnaire and are included on the long-form questionnaire administered to a fraction of households and residents of group quarters. The basic items are also called 100 percent or complete-count items.

This appendix describes the processing of the basic data items for households and persons, focusing on procedures to supply values for missing items (see Titan Corporation, 2003, Sheppard, 2003, and Alberti, 2003, respectively, for detailed information on the steps in data capture, coverage edit and telephone follow-up, and data processing). The Census Bureau distinguishes between "imputation" (or "allocation" in the Bureau's terminology), in which information from another person or household is used to supply a missing value, and "assignment," in which a value is assigned on the basis of other information for the same person (e.g., supplying a value for sex on the basis of first name). There is also "editing," in which inconsistent reported values are reconciled. Sometimes an inconsistent value is deleted and imputation or assignment is used to supply a value for the item. Finally, there is whole-household imputation (or "substitution"), in which an entire household is duplicated for another enumerated household that lacks sufficient information to be termed "data-defined." For the complete count, a household is data-defined if at least one member has reported values for at least two basic items (including name).

G.1 DATA CAPTURE AND COVERAGE EDIT

For the 2000 census, data capture of information on questionnaires was performed by optical mark and optical character recognition (OMR/OCR) after the questionnaires were scanned into computer files. Names were captured, along with write-in and checked values. Clerks keyed data items from images when the automated technology could not read the responses. Keying of long-form-sample information was set aside in the processing to permit the fastest possible keying of the basic information, which was captured from both short-form and long-form questionnaires to obtain complete-count records.

After data capture, computer routines reviewed the complete-count records for mail returns (including the small number of Internet and Be Counted returns) to identify cases for telephone follow-up. The workload for the coverage edit follow-up operation totaled

2.5 million cases, or about 3 percent of mail return forms (Sheppard, 2003:vii). These cases included mail returns that reported household counts of seven people or more (55 percent of the workload); mail returns that reported more household members in question one ("How many people were living or staying in this house, apartment, or mobile home on April 1, 2000?") than the number of members for which at least two basic items were provided (27 percent of the caseload); and mail returns that reported fewer household members in question one than the number for which at least two basic items were provided (18 percent of the caseload).

The purpose of the edit and telephone follow-up was to obtain basic items for all members of large households and to resolve discrepancies in the household count for other households. However, there was no attempt to obtain responses for those members of a contacted household for whom some but not all items were filled out on the questionnaire; the only data collected were for additional people identified in the follow-up and people for whom there had not been room on the form to provide basic data.

The telephone effort was successful in obtaining information from only 54 percent of the workload. There was no field follow-up when the telephone follow-up was unsuccessful nor for cases that lacked telephone numbers. As a result of the telephone follow-up, 153 thousand people were added to the census count and 258 thousand were deleted because they duplicated another person or for another reason should not have been included in the count (Sheppard, 2003:viii).

Following data capture and coverage edit follow-up, census complete-count records for households and their members could fall into one of three categories—categories A, B, and C, as shown in Box 4.2 in Chapter 4. Records for group quarters residents could fall into categories A or B, but not C.

G.2 ITEM IMPUTATION AND EDITING

G.2.a Imputation Methodology

Census records that meet the criteria for being data-defined (at least two reported items, counting name as an item for complete-count records) can still contain missing responses and responses that are inconsistent with other reported data. The Census Bureau first attempts to provide values for missing data and reconcile inconsistent responses by edit and assignment processes that use other information for the same person. Often, however, there is no other information on which to base a reasonable assignment. In these cases, the Bureau uses hot-deck imputation, which supplies values for missing or irreconcilable responses from reported information from a neighboring household (sometimes the imputation can use reported information from other members of the same household). The imputation process begins with a "cold-deck" value; then, in a procedure first used in 1990, passes through a "warm deck;" and then uses a "hot deck" (see Stiller and Dalzell, 2003).

As a matter of history, the term "cold deck" derives from early data processing technology in which a set of punched cards (deck) contained numeric values to represent a known distribution of answers to a question in a previous census or survey. For example, using a long-form item as an example, a cold deck might contain a random sequence of values for veteran status such that a certain percentage of people who did not report veteran status would have the value for "served in the Armed Forces" assigned to them. A collection of such decks, with specific distributions for, say, men and women of different ages, would form an imputation matrix. The values from the appropriate cell of the cold deck matrix would be assigned sequentially to nonreporters and would be reused as many times as necessary without change.

Cold decks were first used to impute age in the 1940 census. In the 1960 census computerized routines employed both cold-deck and hot-deck imputation routines. The cold-deck process did not vary the distribution of values that was used for imputation. The hot-deck process, in contrast, continually updated the values in the distribution from the census data themselves—imputation for a missing entry was made from the latest stored value that fit other known characteristics of the person or housing unit. Because the census records were stored in a geographic hierarchy (block, census tract, county, state), the hot-deck method generally ensured that a donor record would be in the same small neighborhood as the record requiring imputation. It also reproduced the variability in the reported data.

Over the decades, hot-deck imputation matrixes (also edit matrixes) have been refined and enhanced. The matrixes are much more complex for long-form than for short-form items because of the availability of so many additional variables to work with on the long form. Some hot-deck matrixes apply to specific items; other matrixes jointly impute values for groups of items. Subject specialists in the Population Division and the Housing and Household Economic Statistics Division specify the edit and imputation matrixes that apply to their area of expertise. Unfortunately, while written specifications for edits and imputations are available, there is no documentation that makes the specifications readily interpretable.

In 2000, the hot-deck process operated separately for each item or group of items for housing units, household members, and group quarters residents. First, programmers implemented a cold-deck matrix of starting values, chosen to be the most likely distribution of valid responses for the item in question. Then, within each state, the records with valid reported values for an item were processed to find the first four valid values for each cell of the matrix, replacing the cold-deck values (in 1990, each cell of an imputation matrix could contain up to 8 values). This procedure was called "warming the deck." The reason to process the entire state was that some cells of a matrix could contain very few cases. Then all of the records in each state were run again. If the first record required imputation, then a value would be obtained from the first of the four values in the appropriate cell of the matrix. If the second through fifth records also required imputation, then the second, third, and fourth values, followed by the first value again, would be used. If the sixth record had a reported value, that value would be entered into the matrix and the value most recently used as a donor would be discarded from the matrix, and so on. There was no limit on how often a donor record could be used to provide a value for imputation. A simple illustration is given in Box G.1, using marital status (a short-form item in 1990, moved to the long form in 2000; see also the diagrams in Stiller and Dalzell, 2003). Very similar procedures were followed in 1990.

G.2.b Example of Edit and Imputation Specifications: Housing Tenure

The edit and imputation specifications for housing tenure in 2000 (U.S. Census Bureau, 2002a) illustrate both the simplicity and com-

Box G.1 Simple Illustration of 2000 Census Hot Deck Imputation Process for a Single Cell of an Imputation Matrix

"Cold Deck" cell values (e.g., for imputing married = 1 or not married = 2 to women of a specified age)
1, 2, 1, 1

"Warm Deck" cell values
(reported value and record number)
1 (rec. 3), 1 (rec. 4),
2 (rec. 9), 1 (rec. 10)

Record No.	Missing?	Value Assigned
1	Yes	1 (rec. 3)
2	Yes	1 (rec. 4)
3	No	Keep reported value (1)
4	No	Keep reported value (1)
5	Yes	2 (rec. 9)
6	Yes	1 (rec. 10)
7	Yes	1 (rec. 3)
8	Yes	1 (rec. 4)
9	No	Keep reported value (2)
10	No	Keep reported value (1)
11	No	Keep reported value (2)

"Hot Deck" cell values
2 (rec. 11), 2 (rec. 9),
1 (rec. 10), 1 (rec. 3)

12	No	Keep reported value (1)

"Hot Deck" cell values
1 (rec. 12), 2 (rec. 11),
2 (rec. 9), 1 (rec. 10)

13	Yes	1 (rec. 12)
14	Yes	2 (rec. 11)
15	Yes	2 (rec. 9)
16	No	Keep reported value (1)

"Hot Deck" cell values
1 (rec. 16), 1 (rec. 10),
2 (rec. 12), 2 (rec. 11)

The resulting distribution at this point in the process would contain 11 values of "1" and 5 values of "2" for women in the specified age range.

plexity of the process, depending on the extent of available information. Because housing tenure was the only short-form housing item in 2000, the imputation specifications for short-form records were simple:

- accept a reported value for tenure from an occupied housing unit;

- fill in a missing value by using the reported value for the preceding household that falls in the same cell as the nonreporting household.

The imputation matrix included 5 cells: household size 1 person; household size 2 people (household with spouse present, other household); household size 3 or more people (household with spouse present, other household). There were no consistency edits for housing tenure on short forms because there were no other housing variables (e.g., reported rent) to employ.

In contrast, the edit and imputation specifications for housing tenure on long-form records were complex because additional relevant information was available. Variables used included not only household size and type, but also property value, monthly rent amount, whether there is a mortgage, amount of monthly mortgage payment, whether there is a second mortgage or home equity loan, amount of monthly second mortgage payment, type of building (e.g., mobile home, detached one-family, apartment building with 50 or more units), and whether there is a mobile home installment loan. The edit and imputation matrix had over 50 cells. Some cells specified edits for reported values—for example, change a reported value for tenure of 3 (rented for cash rent) to 1 (owned with a mortgage) when monthly rent is blank and a mortgage value is reported. Some cells specified edits for blank values—for example, set tenure to 1 (owned with a mortgage) when tenure is blank but there is a reported value for a first or second mortgage. Other cells specified hot-deck imputations when tenure is blank and other variables (e.g., mortgage) are blank, too, or when tenure is blank and other reported values may be contradictory. The imputation cells were the same as for the short-form records (5 cells based on household size and type).

Type of return (mail, enumerator) was not used as an imputation cell for housing tenure or any other basic (or long-form-sample) item. This omission may have biased the distributions of values if enumerator (mail) returns with missing values were more like other enumerator (mail) returns than total returns.

G.3 PERSON IMPUTATION

Whole-person or type 1 imputation (see Box 4.2 in Chapter 4) refers to instances when one or more people in a household, but not all household members, are not data-defined. In 2000, a large fraction of whole-person imputations (called "totally allocated people" by the Census Bureau) resulted from the decision to limit space on the mailback questionnaire for recording basic items to only 6 members. The coverage edit and telephone follow-up operation attempted to obtain characteristics for additional household members but was not always successful (see Section G.1 above). Editing and imputation were employed to construct values for the basic items for non-data-defined people in enumerated households. The items were edited and imputed one at a time by making use of information about the other household members in order to construct a household that made sense in terms of the relationships, ages, sex, race, and ethnicity of all of the household members.

In 2000, there were 2.33 million whole-person imputations for household members, or 0.9 percent of the household population (Schindler, 2001). By contrast, in 1990, there were only 373,000 such imputations, or 0.2 percent of the total population (including some imputations for people in group quarters; Love and Dalzell, 2001). In 1990, the questionnaire had room to report characteristics of 7 household members, and field follow-up was used in addition to telephone follow-up to obtain data for members of enumerated households who lacked basic information. In 1980, the number of type 1 imputations was lower yet: about 152,000 persons, or less than 0.1 percent of the total population (including some imputations for people in group quarters; calculated from Love and Dalzell, 2001).

Whole-person imputations in 2000 were most likely to involve children. Such imputations accounted for 1.9 percent of children in households, compared with 0.6 percent of people aged 18–29, and 0.4 percent of older people. By race/ethnicity domain and housing tenure, whole-person imputations were most common among the following groups, accounting for 2.1 percent to 2.3 percent of each: American Indian and Alaska Native owners and renters on reservations, black owners and renters, and Native Hawaiian and Pacific Islander owners. Whole-person imputations were least common among white and other owners and renters, accounting for 0.6 percent and 0.4 percent of these two groups, respectively.[1]

One question about the success of the imputation methodology is whether it reproduced family living patterns appropriately for different groups. For example, large multigenerational Asian families may have listed elderly parents rather than children last on the questionnaire and therefore not have reported characteristics for them. Table G.1 shows the distribution of whole-person imputations by domain/tenure group for four age categories: 0–17, 18–29, 30–49, 50 and older, and the ratio of whole-person imputations for children under age 18 to those for adults aged 50 and older. These ratios are lower for renters than owners in all race/ethnicity domains, indicating a greater propensity to impute people of older ages in large renter households than in large owner households. The lowest ratios are for black and Native Hawaiian and Pacific Islander renters.

It is difficult to know what to make of these patterns without information on the age distributions of large households with characteristics reported for all members versus those lacking data for some members. Data are available at the census tract level that could be analyzed to compare the age distribution by domain and tenure of data-defined people with the age distribution for whole-person imputations. However, these data do not permit direct analysis of households that had whole-person imputations versus comparable households that did not.

G.4 HOUSEHOLD IMPUTATION

Four types of situations can occur in the census that require whole-household imputation (what the Census Bureau terms "substitution") because nothing is known about the basic characteristics

[1]These and other characteristics of whole-person imputations were obtained from tabulations by panel staff of U.S. Census Bureau, File of Census Imputations by Postratum, provided to the panel July 30, 2002 (Schindler, 2001).

Table G.1 Percent Whole-Person Imputations (Type 1) by Age and Domain/Tenure Category, Household Members, 2000 (Percent)

Domain/Tenure Category	Age in Years (Percent Whole-Person Imputations of Category)				Ratio, 0–17 Percent to 50 and Older Percent
	0–17	18–29	30–49	50 and Older	
American Indian/Alaska Native on Reservation					
Owner	3.8	2.1	0.7	0.5	7.6
Renter	3.3	1.6	0.6	0.8	4.1
American Indian/Alaska Native off Reservation					
Owner	1.8	1.1	0.4	0.4	4.5
Renter	2.3	0.9	0.5	0.7	3.3
Black, non-Hispanic					
Owner	4.1	2.6	0.9	1.0	4.1
Renter	3.6	1.7	1.0	1.5	2.4
Hispanic					
Owner	2.3	1.7	0.6	0.5	4.6
Renter	2.3	1.1	0.6	0.8	2.9
Native Hawaiian and Other Pacific Islander					
Owner	4.0	2.7	0.9	1.1	3.6
Renter	3.2	1.4	0.8	1.4	2.3
Asian, non-Hispanic					
Owner	3.2	2.1	0.6	0.8	4.0
Renter	3.2	1.1	0.6	0.9	3.6
White and Other Race, non-Hispanic					
Owner	1.1	0.6	0.2	0.2	5.5
Renter	1.5	0.5	0.3	0.4	3.8
Total					
Owner	1.6	1.1	0.3	0.3	5.3
Renter	2.3	0.9	0.5	0.6	3.8

NOTE: Domain/tenure categories are those defined for the 2000 A.C.E. (see Table E.3).

SOURCE: Tabulations by panel staff of U.S. Census Bureau, File of Census Imputations by Poststratum, provided to the panel July 30, 2001 (Schindler, 2001).

of any of the household members (see Box 4.2 in Chapter 4). They are:

(a) household size (number of persons) is known for an occupied unit, but the characteristics of the household members are not known (characteristics or type 2 imputation);

(b) a housing unit is known to be occupied, but household size is not known (count or type 3 imputation);

(c) a housing unit is known to exist, but its status as occupied or vacant is not known (occupancy or type 4 imputation);

(d) an address is recorded, but its housing unit status (occupied, vacant, or not a housing unit) is not known (housing status or type 5 imputation).

These situations can occur in field follow-up when repeated interview attempts are not successful in finding a respondent at home or in obtaining adequate information from a landlord or neighbor. These situations can also result from processing problems that result in lost or corrupted data.

The procedure used in 2000 (and in 1990) for type 2 imputations was to duplicate the basic information from another housing unit record in the nearby area that had the same household size. For type 3 imputations, the imputation process first categorized them as units at single-unit or multiunit addresses. Then, household size and basic items were imputed from an occupied unit at a single-unit or multiunit address with a reported population count from an enumerator-completed form. (In 1990, mail returns were also included in the donor pool.) A similar process was followed for type 4 imputations, for which occupancy status, and, if need be, occupied household size and basic items had to be imputed (the donor pool consisted of occupied and vacant units from enumerator-completed forms). The same type of process was also followed for type 5 imputations, for which housing status had to be imputed first, followed by occupancy status, and, for occupied housing units, their size and characteristics (the donor pool consisted of occupied, vacant, and deleted units from enumerator-completed forms). A potential donor record could only be used once and, in general, was selected from the same census tract as the unit requiring imputation (see Griffin, 2001).

Table 4.1 in Chapter 4 provides statistics on the numbers of people for whom 2000 census records were imputed in each of imputation types (2) through (5), with corresponding statistics for 1990 and 1980, when available. Table G.2 shows the distribution of people in whole-household imputation situations in 2000 among race/ethnicity and housing tenure groups by type of imputation required. It also shows the total percentage of people requiring whole-household imputation for each group. Several patterns stand out. Renters have twice as high a whole-household imputation rate as owners—2 percent compared with 0.9 percent. Most of the difference is due to a higher proportion of type 2 imputations (when household size is known but not characteristics of members)—72 percent of all whole-household imputations are type 2 for renters compared with 60 percent for owners. American Indians and Alaska Natives on reservations have the highest whole-household imputation rate of any race/ethnic group—3 percent. This result is largely due to very high proportions of imputations for status as a housing unit (type 5). Thus, 61 percent and 55 percent, respectively, of whole-household imputations for American Indian owners on reservations and American Indian renters on reservations are type 5 imputations, compared to 12 percent type 5 imputations among whole-household imputations for the nation as a whole.

Analysis of the geographic distribution of whole-household imputations finds wide variations in the geographic location of most imputation types. Type 2 imputations (count is known but not characteristics) are clustered in large cities, such as New York and Chicago. However, they are not prominent in other large cities, notably Los Angeles. Type 3 imputations (size is not known for an occupied housing unit) follow the same general pattern as type 2 imputations. Type 4 imputations are the least common and do not show a particular geographic pattern. Type 5 imputations, which are the least well-founded, are clustered in rural areas that were enumerated by list/enumerate techniques in which the enumerator developed the address list and obtained responses to the census questions at the same time.

Table G.2 Distribution of People Requiring Whole-Household Imputation by Type of Imputation, by Race/Ethnicity Domain and Housing Tenure, 2000 Census

Domain and Tenure Group	Percent of People Requiring Imputation, by Type: Charac-teristics (2)	Count (3)	Occu-pancy (4)	Housing (5)	Percent of Total Household Population Requiring Imputation: Types 2–5	Types 3–5
American Indian/ Alaska Native on Reservation						
Owner	15.7	20.5	3.2	60.6	3.0	2.5
Renter	17.8	25.4	2.2	54.6	2.8	2.3
American Indian/ Alaska Native off Reservation						
Owner	56.5	15.1	7.3	21.1	1.4	0.6
Renter	68.1	13.1	6.1	12.7	1.7	0.5
Hispanic Origin						
Owner	59.2	18.2	5.8	16.8	1.5	0.6
Renter	70.1	15.7	5.1	9.2	1.9	0.6
Black (Non-Hispanic)						
Owner	69.2	16.6	6.0	8.2	1.6	0.5
Renter	74.4	15.8	4.9	4.9	2.5	0.6
Native Hawaiian and Other Pacific Islander						
Owner	65.2	16.6	3.9	14.3	1.4	0.5
Renter	71.6	14.9	3.5	10.0	2.0	0.6
Asian (Non-Hispanic)						
Owner	67.5	13.6	6.3	12.6	0.8	0.3
Renter	75.7	13.3	4.9	6.2	2.0	0.5
White and Other Races (Non-Hispanic)						
Owner	58.8	13.6	10.8	16.9	0.8	0.3
Renter	71.2	12.5	7.0	9.4	1.7	0.5
Total						
Owner	60.4	14.7	9.1	15.7	0.9	0.3
Renter	71.9	14.1	5.9	8.2	2.0	0.6
Grand Total	66.0	14.4	7.6	12.1	1.3	0.5

NOTES: See Box 4.2 in Chapter 4 for definition of imputation types; type 1 imputation is not included because it involves imputation for one or more people in a household with at least one data-defined person.

SOURCE: Tabulations by panel staff of U.S. Census Bureau, File of Census Imputations by Poststratum, provided to the panel July 30, 2001 (Schindler, 2001).

Appendix H

2000 Census Long-Form-Sample Data Processing

Similar to the companion Appendix G on basic data processing, this appendix describes the processing of the items that were asked of the long-form sample in the 2000 census (see Appendix B for a listing of the long-form-sample items). It covers data capture, weighting, and imputation, assignment, and editing of items for missing and inconsistent responses. About 1 in 6 households received the long form in 2000. Sampling rates were 1 in 2 for governmental areas (counties, towns, townships, and school districts) with fewer than 800 occupied housing units (fewer than about 2,100 people); 1 in 4 for governmental areas with 800–1,200 occupied housing units (about 2,100–3,100 people); 1 in 6 for census tracts with fewer than 2,000 occupied housing units (fewer than about 5,200 people); and 1 in 8 for larger census tracts. (Estimates of occupied housing units were those developed by the Population Division as part of the intercensal estimates program.) The 1990 census long-form sampling scheme was similar, except that there was no 1 in 4 sampling rate and school districts were not among the governmental areas that were eligible to be oversampled.

H.1 DATA CAPTURE

For the 2000 census, data capture of information on questionnaires was performed by scanning short-form and long-form returns into computer files and using optical mark and optical character recognition (OMR/OCR) to record the information. Clerks keyed data items from images when the automated technology could not read the responses. Keying of long-form-sample information was carried out in a second, separate process in order to permit the fastest possible completion of data capture for the basic items on all returns.

After data capture, long-form-sample records for households and their members could fall into one of two categories:

(A) *Long-Form Data-Defined* At least one member of a household in the long-form sample was "long-form data-defined;" that is, at least one member had at least two long-form data items reported. All records for long-form data-defined households were retained in the sample. Any long-form housing or person items not reported, or reported inconsistently, had missing or consistent values supplied through item imputation, assignment, and editing. Imputations for any missing complete-count items that were performed during the basic data processing were retained (i.e., they were not reimputed during the long-form-sample processing).

(B) *Whole-Household Nonresponse* Households that lacked any long-form data-defined persons were dropped from the sample. Weights were developed for long-form data-defined households and their members so that long-form-sample estimates agreed with complete-count totals on basic items. The weighting effectively adjusted for whole-household nonresponse.

H.2 WEIGHTING

H.2.a Initial Weighting Areas

In a procedure similar to that used in 1990, 2000 long-form-sample weights were developed to produce estimates for specified groups and geographic areas that agreed with estimates from the basic (complete-count) data records. A goal of the weighting was to minimize the variation in weights, which, in turn, would minimize the variation in estimated sampling error across population groups and geographic areas. Adjusting the weights to match complete-count controls would also reduce the variance in estimates.[1] The weighting was specified by the Decennial Statistical Studies Division (see Hefter, 2000; see also U.S. Census Bureau, 2003d:Ch.8).

Initial weighting steps included defining initial weighting areas and computing for each the ratio of basic records (complete-count data-defined people and whole-person and whole-household imputations) to long-form data-defined records. The purpose for the ratios was to determine if data augmentation would be needed before proceeding to calculate final weights.

Initial weighting areas (IWAs) were defined within counties. They comprised all of the records in a tabulation block group with the same expected sampling rate (1/2, 1/4, 1/6, or 1/8). Ratios were calculated of long-form data-defined records to the total records in each IWA separately for housing units, persons in group quarters, and persons counted in service-based enumeration.

H.2.b Data Augmentation

In order to achieve the target IWA ratios of sample to total records for housing units, group quarters enumerations, and service-based enumerations so that weights would not be too large, an appropriate number of records that were not previously long-form data-defined were selected to be augmented—that is, have values for their long-form data items supplied through editing, assignment, and imputation. After this first round of augmentation, IWAs were combined into final weighting areas (FWAs), and a second round of augmentation was performed as needed.

As it turned out, augmentation was rarely required in 2000: only 1,477 occupied housing units were selected for augmentation out of 16.4 million occupied units in the long-form sample (2,412 vacant housing units were also selected for augmentation). Less than 0.001 percent of long-form-sample household member records (unweighted) were augmented people (4,090 of 42.6 million records);

[1]See Fairchild (2001) for draft specifications for the estimation of direct variances for the 2000 long-form sample. These variance estimates do not account for the variance due to item imputation.

however, 4.2 percent of group quarters long-form-sample records were augmented people. On a weighted basis, augmented household members represented only 0.002 percent of the household population estimated from the long-form sample, while augmented group quarters residents represented 11.9 percent of the group quarters population estimated from the long-form-sample. This percentage is disturbingly high. It would be useful to know how it varied by type of group quarters (e.g., college dormitory or prison).

H.2.c Final Weighting Areas

The specifications for combining IWAs into FWAs called for no FWA to have fewer than 400 long-form data-defined persons, while also observing the constraint that IWAs were never combined across counties. If an IWA had more sample persons than the threshold, it was an FWA on its own. If it had fewer sample persons, then it was combined with one or more IWAs in the same tabulation block group, or in the same census tract if necessary, or in the same county if necessary, to reach the threshold. When possible, IWAs were combined that had the same expected sampling rate.

H.2.d Construction of Weights

Initial weights were calculated separately for five groups: persons in households, group quarters residents, service-based enumerations, occupied housing units, and vacant housing units. These weights were the ratio of the complete count to the long-form-sample count for each group in an IWA.

Final Person Weights

Next, person weighting matrixes were computed for each FWA. The matrix for household members included the complete count, unweighted sample count, and weighted count (based on the initial person weights) for each cell of a 4-dimensional matrix (39,312 cells). Marginal counts (complete count, unweighted, initially weighted) were also produced for each cell of each dimension separately.

The four dimensions for household members were as follows (there were simpler matrices for group quarters residents and service-based enumerations):

1. three categories of household type by seven categories of household size (21 cells);
2. three categories of sampling type (1 in 2, 1 in 4, 1 in 6, or 1 in 8) (3 cells);
3. two categories of householder status (householder or not) (2 cells); and
4. two categories of Hispanic origin, by six categories of race (American Indian or Alaska Native, Black, Asian, Native Hawaiian or Other Pacific Islander, White, Other, with multiple-race persons assigned to the largest nonwhite single-race category in the FWA), by 13 categories of age, by two categories of sex (312 cells).

Type of return (mail, enumerator) was not included as a dimension in the weighting matrix, although almost all non-sample-data-defined records (whole-household nonrespondents) were enumerator returns and not mail returns. Thus, members of non-sample-data-defined occupied housing units were 17.7 percent of total enumerator long-form household member records, compared with only 0.5 percent of total mail long-form household member records (from tabulations by Census Bureau staff provided to the panel spring 2003).

The cells of the household member weighting matrix were collapsed as necessary. The entire matrix was collapsed to a single cell if the complete person count in the FWA was more than 40 times the uninflated sample count. One or more cells in a single dimension were collapsed, following the process below (see Hefter, 2000:Section V.I, for the rules for combining categories within a dimension):

- first, determine if one or more categories of the household type/size dimension needed to be collapsed (each category must have at least 10 sample persons and a ratio of complete-count persons to initially inflated sample persons of less than 3.5);
- next, determine if the Hispanic origin classification failed (both the Hispanic and non-Hispanic categories must have at least 150 complete-count persons in addition to the criteria above for household type/size);

- next, determine if the race categories (within Hispanic origin, if still present) required collapsing (each category must meet the criteria specified for Hispanic/non-Hispanic);
- next, determine if the age/sex classifications within Hispanic/race categories must be collapsed (each category must have at least 10 sample persons and a ratio of complete-count persons to initially inflated sample persons of less than 4);
- next, determine if the householder/nonhouseholder dimension failed (each category must have at least 10 sample persons and a ratio of complete-count persons to initially inflated sample persons of less than 3.5);
- finally, determine if the sampling type categories required collapsing (each must have at least 10 sample persons and a ratio of complete-count persons to initially inflated sample persons of less than 3.5).

After the final person matrix was determined for an FWA, then an iterative proportional fitting (raking) procedure was conducted, in which the initially inflated sample counts in each cell were adjusted so that the marginal cell totals for each dimension were practically equal between the complete counts and the inflated sample counts. This result was accomplished by first adjusting the initially inflated cell counts to equal the complete-count marginals for one dimension, then a second dimension, and, sequentially, through all the dimensions, followed by additional iterations as needed until a specified stopping point was reached. (See National Research Council, 1985:App.3.2, for a general description of iterative proportional fitting, which has been used in every census since 1970.)

The last step in constructing person weights was to use a controlled rounding procedure in order to produce integer weights within each state. Before approving these weights, they were tested to be sure they did not exceed specified size criteria. If they did, then a procedure was used to force additional collapsing of the person-weighting matrix, by successively lowering the maximum ratio of complete-count persons to initially inflated sample persons that was permitted.

The distribution of final long-form-sample person weights for 2000 is concentrated in the range of 3–30, with a longer upper tail for group quarters residents compared with all people, as seen below. It would be useful to analyze weight distributions and average weights for small geographic areas.

	Cumulative Percent	
Weight Value	All Persons	Group Quarters
1	0.45	0.02
2	4.33	0.12
3	7.34	0.44
4–10	70.58	56.76
11–20	98.06	80.23
21–30	99.60	91.70
31–40	99.91	97.89
41–50	99.97	99.22
51–max.	100.00	100.00
(Max =)	(320)	(180)

SOURCE: Tabulated by panel staff from U.S. Census Bureau, Edit Tallies for Long-Form Population Records (variable WT; see Philipp, 2001).

Final Occupied Household Weights

The procedure for developing occupied household weights was similar to that for persons. The only difference was in the definition of the weighting matrixes and the criteria for collapsing.

The occupied household weighting matrix consisted of three dimensions (1,512 cells):

1. three categories of household type by seven categories of household size (21 cells);
2. three categories of sampling type (3 cells);
3. two categories of tenure by two categories of Hispanic/non-Hispanic origin of householder by six categories of race of householder (as defined for the person weighting) (24 cells).

The entire matrix was collapsed to a single cell if the ratio of the complete count to the unweighted sample count was more than 40 to 1 for an FWA. One or more categories of a dimension were collapsed if the marginal unweighted sample count was less than 5 or the ratio of complete-count persons to initially inflated sample persons was greater than 3.5 (for tenure, Hispanic origin, and race, collapsing also occurred if there were fewer than 50 sample cases in a cell). The collapsing proceeded by testing, sequentially, household type/size categories, tenure categories, Hispanic origin categories within tenure, race categories within tenure/Hispanic origin, and sampling type categories.

Final Vacant Housing Unit Weights

The process for developing vacant housing unit weights was not iterative. Vacant units were classified into three categories: vacant for rent, vacant for sale, other vacant. These categories were collapsed as necessary, and weights were calculated for the vacant units in each category in an FWA by inflating the initially inflated sample counts to equal the complete counts.

H.3 ITEM IMPUTATION

Imputation (using reported values from another person or household in a hot-deck procedure), assignment (using reported values for the same person or household), and editing (changing values according to specified rules for consistency) were used in 2000 for all instances of missing and inconsistent values for members of long-form-sample data-defined households (whether or not the person record was sample data-defined). These procedures were also used for all long-form-sample data-defined group quarters residents. The imputations made during the complete-count processing of basic items were retained and not reimputed in the long-form-sample data processing. The hot-deck imputation procedure is described in broad outline in Appendix G.1; the Census Bureau uses the term "allocation" for item imputation.

The edit and imputation specifications for the long-form were quite complex. Generally, related variables (e.g., the set on education, see below) were imputed sequentially so that responses to a specific question would be consistent with responses to a logically preceding question. In the case of income and employment variables, there was a "joint economic edit," which was the most complex procedure of all (see U.S. Census Bureau, 2002b). It was carried out after all other editing and imputation had been performed and applied to year last worked, industry, occupation, class of worker, work experience in 1999, earnings, and all other income types. For yet another example of long-form-sample imputation and edit specifications, see Appendix G.2.b which describes procedures for editing and imputing housing tenure. The long-form procedures were much more complex than the short-form procedures because of the availability of related variables on the long form, such as mortgage payment and rent. For most long-form person variables, there were somewhat different procedures for household members and group quarters residents.

H.3.a Example of Edit and Imputation Specifications: Education Variables

The three education variables—school enrollment, grade attending, and educational attainment—were edited jointly (see U.S. Census Bureau, 2001a). Variables that were used for the education edits and imputations included age, race, ethnicity, whether worked last week or was on layoff or temporarily absent from work, occupation, and employment status recode. Starting ("cold-deck") values were specified for 17 different matrices, although these values were superceded by the "warm deck," and then were continuously updated through the "hot-deck" process (see Appendix G.1).

The first steps involved a large number of edits based on age. For example, all three education variables were set to zero (not in universe) if age (which may have been imputed in the complete-count processing) was less than 3 years. One or more of the education variables were also set to zero if the reported or imputed age was not consistent with the educational data (e.g., if age was 18 or more and school enrollment was no, then any reported value for grade attending was set to zero). These edits assumed that age reporting and imputation were reliable.

Next were edits to make educational attainment consistent with reported grade attending. For example, if grade attending was grade 1 to grade 4 and age was 8–10, but educational attainment was greater than grade 4, then educational attainment was blanked and imputed at a later step. Sometimes, it was grade attending that was blanked depending on age and reported educational attainment.

After the edits were completed, then blank values because of nonresponse or editing were imputed using the specified imputation matrix. Some matrices were simple; for example, when educational attainment was missing but school enrollment and grade attending were reported, then the imputed value for educational attainment was the hot-deck value in the appropriate cell of a matrix of categories of age by grade attending. More complex matrices handled situations when all three education variables were missing: matrix 7A imputed all three education variables for unemployed people according to the hot-deck values in the appropriate cell formed by age and race/ethnicity; matrix 7B imputed all three education variables for employed people according to the hot-deck values in the appropriate cell formed by age and occupation group.

H.3.b Analysis

Chapter 7 analyzes imputation rates for the 2000 and 1990 census long-form samples and the Census 2000 Supplementary Survey. Eight tables supplement those provided in Chapter 7. Tables H.1 through H.7 are for the household population (see National Research Council, 1995b:App.L, for similar tables for 1990): imputation rates for selected population and housing items by self versus enumerator form for 2000 and 1990 (Table H.1); imputation rates for selected population and housing items for 2000 by race and Hispanic origin of the reference person or householder (Table H.2); imputation rates for the worst 10 percent census tracts for selected population and housing items for 2000 by race and Hispanic origin of the reference person (Table H.3); imputation rates for selected population and housing items for 2000 by geographic aggregations (Table H.4); imputation rates for the worst 10 percent census tracts for selected population and housing items for 2000 by geographic aggregations (Table H.5); 2000 imputation rates, 2000 imputation and assignment rates (1990-comparable), and 1990 imputation rates for population items (Table H.6); and 2000 imputation rates, 2000 imputation and assignment rates (1990-comparable), and 1990 imputation rates for housing items (Table H.7). Table H.8 provides 2000 imputation and assignment rates and 1990 imputation rates for population items for group quarters residents by type of group quarters.

Figure H.1 graphs comparable imputation rates for 2000 and 1990 for housing items. It shows (as would a similar graph for population items) that 2000 imputation rates are higher than 1990 rates for most items. One reason relates to the fact that the percentage of long forms included in sample processing was 2 percentage points higher in 2000 than in 1990 (93.2 percent and 91.2 percent, respectively). It was easier for a household to be sample data-defined in 2000 because of the layout of the questionnaire (see discussion in Chapter 7), but, as a consequence, a larger proportion of sample data-defined forms in 2000 were only minimally completed, which produced higher imputation rates. Another and more important reason for higher imputation rates in 2000 is that the design, in contrast with 1990, precluded telephone and field follow-up for missing content (see discussion in Chapter 4). Users should examine both weighting factors and imputation rates to assess the effects of nonresponse on the variability and possible bias of estimates from the 2000 and 1990 long-form samples.

Table H.1 Imputation/Assignment Rates for Selected Population and Housing Items, 2000 and 1990 Census Long-Form Sample, Household Members, by Type of Response: Household Respondent (Self) vs. Enumerator-Filled (Enum) (weighted)

	2000			1990		
Item	Total	Self (75.5%)	Enum (24.5%)	Total	Self (73.1%)	Enum (26.9%)
Relationship	2.7	2.6	3.0	1.9	1.7	2.4
Sex	1.6	1.5	2.0	0.8	0.9	0.7
Race	3.2	3.7	2.0	1.1	1.0	1.6
Age	2.6	1.9	4.3	0.9	0.8	1.2
Marital Status	3.4	2.3	6.2	0.9	0.8	0.9
Hispanic Origin	4.0	4.7	2.4	3.4	4.0	1.7
Place of Birth	9.2	7.8	12.5	5.1	4.3	7.1
Educational Attainment	7.2	5.2	12.0	4.5	3.8	6.1
English-Speaking Ability	7.6	7.3	7.9	8.5	8.3	9.0
Veteran Status	8.2	6.8	11.8	4.8	4.0	7.0
Work Disability	11.4	12.2	9.3	7.4	7.4	7.6
Mobility Disability	10.0	10.5	8.6	5.1	4.7	6.4
Self-Care Disability	7.9	7.4	9.1	5.8	5.5	6.7
Employment Status Recode	11.1	10.2	13.4	3.8	3.0	6.2
Place of Work (Place)	10.6	8.0	16.8	9.7	8.7	12.5
Occupation Last Year	16.1	14.3	20.4	9.1	7.9	12.5
All Income Imputed	24.5	18.9	38.5	11.7	9.1	19.1
Some Income Imputed	29.7	25.5	40.3	13.4	10.9	20.5
Housing Tenure	8.0	7.6	8.9	1.4	1.4	1.4
Number of Rooms	6.2	6.2	6.4	0.4	0.4	0.5
Complete Plumbing	3.4	3.5	3.1	1.7	1.7	1.8
Complete Kitchen	3.4	3.5	3.1	1.8	1.8	1.8
Fuel Used for Heating	7.4	6.3	10.1	2.9	2.7	3.4
Annual Electric Cost	18.5	15.3	26.9	5.5	4.4	8.5
Monthly Rent	15.6	13.2	19.2	1.3	1.1	1.6
Property Taxes	32.0	27.0	49.6	12.2	10.3	19.4
Value of Property	13.3	12.3	16.6	3.3	3.3	3.4

NOTES: 2000 rates include both assignments and imputations and are therefore comparable to 1990. Work and mobility disability items differ in wording between 2000 and 1990. Imputation rates for employment status recode include only cases for which none of the half-dozen relevant questions were answered. Some income imputed includes 100 percent and less than 100 percent imputed.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 Sample Census Edited File (SCEF) and the 1990 Sample Edited Data File (SEDF), provided to the panel spring 2003.

Table H.2 Imputation Rates for Selected Population and Housing Items, 2000 Census Long-Form Sample, Household Members, by Race and Hispanic Origin of Household Reference Person (weighted)

		Non-Hispanic							
Variable	Total (273,643)	White (191,994)	Black (32,380)	Am. Ind. (1,987)	Asian (9,630)	Pacific Isl. (339)	Other (336)	Multi (3,579)	Hispanic (33,397)
Relationship	2.3	1.5	3.8	3.2	3.2	4.1	3.3	2.9	4.6
Sex	0.9	0.6	1.7	1.6	1.4	2.1	1.3	1.2	2.0
Race	3.2	1.6	2.7	3.0	3.3	6.7	5.5	2.7	13.1
Age	2.6	2.0	4.2	3.4	3.5	5.0	4.8	3.4	4.5
Hispanic Origin	3.6	2.7	7.5	5.3	6.0	8.2	6.0	4.7	4.4
Number of Rooms	6.2	5.2	8.8	7.0	11.0	11.1	9.8	7.0	9.6
Housing Tenure	4.3	4.0	6.1	4.2	3.1	4.8	6.9	4.2	4.4
Value of Property	13.3	12.9	17.6	15.2	10.4	14.8	21.7	14.1	13.0
Monthly Rent	15.6	14.3	20.2	14.4	14.1	15.5	20.7	16.6	15.5
Complete Plumbing	3.4	3.1	4.9	3.4	2.9	4.9	5.3	3.5	4.2
Complete Kitchen	3.4	3.1	4.8	3.4	2.9	4.8	5.3	3.4	4.1
Fuel Used for Heating	7.4	6.6	10.4	9.9	7.8	9.3	11.9	9.1	9.6
Annual Electric Cost	17.1	16.1	23.4	15.6	13.9	18.6	26.3	18.7	18.1
Property Taxes	32.0	29.7	47.3	55.2	31.5	43.4	49.4	39.6	40.6
Place of Birth	9.2	8.2	14.1	9.1	7.8	12.0	13.7	10.2	10.2
Educational Attainment	7.2	5.6	10.8	8.2	9.0	11.2	13.3	9.1	12.4
English-Speaking Ability	7.6	10.0	11.7	7.1	5.2	9.3	8.1	7.1	6.8
Veteran Status	7.5	6.3	11.6	7.4	8.6	10.6	11.7	8.7	10.5
Work Disability	11.4	10.7	15.5	9.9	10.5	12.7	13.5	10.6	12.4
Physical Activity Disability	7.6	6.7	11.1	7.7	7.8	10.3	11.5	8.3	9.6
Self-Care Disability	7.9	6.9	11.2	8.0	8.1	10.6	12.0	8.6	10.0
Employment Status Recode	11.1	9.7	16.4	10.5	11.8	13.8	15.5	12.0	15.3
Place of Work (Place)	10.6	9.3	16.0	11.6	11.0	14.8	19.1	13.1	14.5
Occupation Last Year	14.9	12.9	22.5	14.8	15.2	17.8	22.0	16.4	20.3
All Income Imputed	24.5	22.7	33.2	24.6	20.8	28.4	36.8	28.4	28.1
Some Income Imputed	29.7	27.8	39.5	28.8	25.4	32.2	40.5	32.5	33.5

NOTES: Number at the top of each column in parentheses is the population in thousands.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 Sample Census Edited File (SCEF) and the 1990 Sample Edited Data File (SEDF), provided to the panel spring 2003.

Table H.3 Imputation Rates for Selected Population and Housing Items, 2000 Census Long-Form Sample, Household Members, 10% Worst Census Tracts, by Race and Hispanic Origin of Reference Person (weighted)

		Non-Hispanic							
Variable	Total (23,517)	White (6,312)	Black (6,557)	Am. Ind. (271)	Asian (1,178)	Pacific Isl. (52)	Other (68)	Multi (436)	Hispanic (8,643)
Relationship	4.9	2.6	5.0	7.0	5.5	5.6	4.4	4.6	6.3
Sex	2.4	1.2	2.4	4.5	2.9	3.0	1.8	2.1	3.1
Race	8.1	2.9	3.7	6.0	5.4	8.3	6.4	4.1	15.9
Age	5.0	3.2	5.4	7.0	5.7	6.8	5.5	5.5	5.9
Hispanic Origin	6.3	4.0	9.3	9.4	9.1	8.7	6.5	7.0	5.3
Number of Rooms	9.6	6.7	10.9	10.8	12.3	13.7	11.5	9.9	11.4
Housing Tenure	5.9	5.0	7.6	6.3	4.5	5.3	7.4	5.9	5.3
Value of Property	17.4	15.9	21.7	20.2	15.1	18.9	26.9	19.0	15.7
Monthly Rent	20.3	17.3	24.5	18.5	17.9	19.8	24.0	21.6	18.3
Complete Plumbing	5.3	4.1	6.5	5.7	4.7	5.3	7.0	5.4	5.5
Complete Kitchen	5.2	4.1	6.4	5.7	4.6	5.1	7.0	5.4	5.3
Fuel Used for Heating	11.3	8.7	13.3	14.0	11.6	10.6	13.9	13.2	12.3
Annual Electric Cost	21.2	17.7	26.8	17.5	17.4	18.8	26.5	22.4	20.3
Property Taxes	42.3	35.3	50.4	78.0	40.0	47.4	57.0	48.4	45.1
Place of Birth	13.5	11.1	17.8	13.6	11.4	14.6	15.5	14.3	12.3
Educational Attainment	12.9	8.3	14.0	14.1	13.8	14.1	15.0	14.0	15.3
English-Speaking Ability	8.7	9.0	14.3	10.3	8.0	10.8	10.3	10.0	8.4
Veteran Status	11.8	8.5	14.3	11.2	12.5	12.7	13.3	12.7	12.7
Work Disability	14.8	12.5	18.3	12.5	13.9	13.3	13.5	13.9	14.2
Physical Activity Disability	11.5	8.7	13.9	12.0	11.4	11.7	12.2	11.8	11.8
Self-Care Disability	11.8	9.0	14.1	12.3	11.9	11.7	12.6	12.3	12.2
Employment Status Recode	16.6	12.3	19.7	14.4	16.5	16.3	17.3	16.6	18.1
Place of Work (Place)	16.5	11.5	20.4	15.4	16.6	17.4	21.5	19.6	18.5
Occupation Last Year	22.3	15.9	27.2	18.7	21.7	20.7	25.3	22.7	24.5
All Income Imputed	30.8	25.4	36.9	28.1	25.8	31.7	36.9	33.8	31.2
Some Income Imputed	36.5	30.6	43.6	32.5	31.0	35.5	40.8	37.9	37.0

NOTES: Number at the top of each column in parentheses is the population in thousands. Worst 10 percent census tracts are those tracts with the highest number of imputations of basic (complete-count) items.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 Sample Census Edited File (SCEF) and the 1990 Sample Edited Data File (SEDF), provided to the panel spring 2003.

Table H.4 Imputation Rates for Selected Population and Housing Items, 2000 Census Long-Form Sample, Household Members, by Geographic Aggregations (weighted)

Variable	Total (273,643)	In CMSA Cent. City (38,064)	In CMSA Other (69,012)	In MSA Cent. City (44,423)	In MSA Other (68,740)	Non-MSA Urban (21,271)	Non-MSA Rural (32,133)
Relationship	2.3	3.6	2.2	2.4	1.8	1.9	1.9
Sex	0.9	1.6	0.9	1.0	0.7	0.7	0.7
Race	3.2	5.8	3.4	3.5	2.3	2.4	1.8
Age	2.6	3.8	2.5	2.8	2.2	2.4	2.3
Hispanic Origin	3.6	5.1	3.3	3.9	3.0	3.4	3.6
Number of Rooms	6.2	7.8	6.2	6.1	5.7	5.5	6.3
Housing Tenure	4.3	4.7	3.7	3.9	4.3	4.1	5.5
Value of Property	13.3	13.9	11.1	12.0	13.2	13.4	18.0
Monthly Rent	15.6	18.2	14.8	14.8	14.3	14.6	16.3
Complete Plumbing	3.4	4.2	3.1	3.2	3.3	3.1	3.8
Complete Kitchen	3.4	4.1	3.1	3.2	3.3	3.1	3.8
Fuel Used for Heating	7.4	9.3	6.5	7.0	6.8	6.8	9.3
Annual Electric Cost	17.1	19.5	16.5	16.8	16.4	16.7	17.7
Property Taxes	32.0	35.2	27.8	32.5	32.1	31.6	36.9
Place of Birth	9.2	11.9	8.7	9.8	8.5	8.5	8.5
Educational Attainment	7.2	10.5	6.9	7.4	6.3	6.3	6.4
English-Speaking Ability	7.6	7.5	6.8	7.3	8.5	8.1	11.2
Veteran Status	7.5	10.1	6.9	7.7	6.7	7.0	7.1
Work Disability	11.4	13.4	10.8	11.5	10.8	11.2	11.4
Physical Activity Disability	7.6	9.7	7.1	7.9	7.1	7.2	7.5
Self-Care Disability	7.9	10.0	7.4	8.0	7.3	7.3	7.7
Employment Status Recode	11.1	14.3	10.3	11.3	10.2	10.7	11.1
Place of Work (Place)	10.6	13.6	10.6	9.8	9.9	9.1	10.8
Occupation Last Year	14.9	18.9	14.0	15.1	13.8	14.0	15.1
All Income Imputed	24.5	27.3	23.2	24.0	23.8	23.7	26.4
Some Income Imputed	29.7	33.0	28.2	29.3	28.7	29.2	31.7

NOTES: Number at the top of each column in parentheses is the population in thousands. CMSA: Consolidated Metropolitan Statistical Area; MSA: Metropolitan Statistical Area; Cent. City: Central City.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 Sample Census Edited File (SCEF) and the 1990 Sample Edited Data File (SEDF), provided to the panel spring 2003.

Table H.5 Imputation Rates for Selected Population and Housing Items, 2000 Census Long-Form Sample, Household Members, Worst 10% Census Tracts, by Geographic Aggregations (weighted)

Variable	Total (23,517)	In CMSA: Cent. City (11,136)	In CMSA: Other (3,839)	In MSA: Cent. City (3,591)	In MSA: Other (2,978)	Non-MSA: Urban (659)	Non-MSA: Rural (1,314)
Relationship	4.9	5.2	5.5	4.5	4.1	3.3	3.8
Sex	2.4	2.5	2.7	2.1	2.0	1.5	2.1
Race	8.1	8.6	9.9	7.5	7.0	5.0	4.1
Age	5.0	5.3	5.4	4.9	4.3	3.8	4.1
Hispanic Origin	6.3	7.1	5.8	6.3	4.9	4.9	5.6
Number of Rooms	9.6	10.3	10.2	8.7	8.7	7.3	8.0
Housing Tenure	5.9	6.3	5.2	5.3	5.5	5.0	6.8
Value of Property	17.4	18.7	15.1	16.3	16.7	15.2	19.5
Monthly Rent	20.3	22.0	17.9	18.3	16.5	17.8	17.9
Complete Plumbing	5.3	5.9	4.9	4.5	4.7	3.9	4.7
Complete Kitchen	5.2	5.8	4.8	4.5	4.7	3.9	4.7
Fuel Used for Heating	11.3	12.9	10.0	9.9	9.5	8.6	11.2
Annual Electric Cost	21.2	23.7	19.1	20.2	17.8	16.9	18.8
Property Taxes	42.3	44.5	39.6	43.4	40.4	39.1	42.2
Place of Birth	13.5	15.2	12.2	13.1	11.3	10.3	10.8
Educational Attainment	12.9	14.4	12.8	11.9	11.0	9.0	9.3
English-Speaking Ability	8.7	9.0	8.5	8.5	8.1	7.3	9.6
Veteran Status	11.8	13.3	11.2	11.4	9.9	8.8	8.7
Work Disability	14.8	16.2	13.8	14.4	12.7	11.7	11.6
Physical Activity Disability	11.5	12.7	10.8	11.1	10.1	8.9	9.4
Self-Care Disability	11.8	13.1	11.1	11.4	10.2	9.1	9.6
Employment Status Recode	16.6	18.5	15.9	16.1	14.0	12.5	12.3
Place of Work (Place)	16.5	18.7	16.6	14.4	14.1	11.6	12.2
Occupation Last Year	22.3	24.8	21.7	21.6	18.8	16.4	16.4
All Income Imputed	30.8	32.7	29.3	30.0	27.8	27.1	29.3
Some Income Imputed	36.5	38.9	34.8	35.8	32.7	32.1	34.2

NOTES: Number at the top of each column in parentheses is the population in thousands. Worst 10 percent census tracts are those tracts with the highest number of imputations of basic (complete-count) items. CMSA: Consolidated Metropolitan Statistical Area; MSA: Metropolitan Statistical Area; Cent. City: Central City.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 Sample Census Edited File (SCEF) and the 1990 Sample Edited Data File (SEDF), provided to the panel spring 2003.

Table H.6 Imputation Rates for Population Items, 2000 and 1990 Census Long-Form Sample, Household Members (weighted)

Variable	2000a	2000b	1990
Relationship	2.3	2.7	1.9
Sex	0.9	1.6	0.8
Age	2.6	2.6	0.9
Hispanic Origin	3.6	4.0	3.4
Race	3.2	3.2	1.1
Marital Status	2.2	3.4	0.9
Attending School	6.2	6.2	4.2
Grade Level Attending	9.0	9.0	—
Educational Attainment	7.2	7.2	4.5
Non-English Language	5.2	5.8	4.8
Language Spoken	11.4	11.4	11.9
English-Speaking Ability	7.6	7.6	8.5
Place of Birth	9.2	9.2	5.1
Citizenship	0.8	5.2	4.2
Year of Entry	14.7	14.7	8.9
Mobility Status	5.8	8.6	5.2
Migration (State)	8.6	8.6	5.7
Migration (County)	8.6	8.6	7.5
Migration (MCD)	—	—	7.6
Migration (Place)	8.8	8.8	9.6
Sensory Disability	6.9	6.9	—
Physical Activity Disability	7.6	7.7	—
Mental Disability	7.5	7.5	—
Self-Care Disability	7.9	7.9	5.8
Outside Difficulty	9.9	10.0	5.1
Work Disability	11.4	11.4	7.4
Grandchildren	4.5	5.1	—
Responsible for Grandchildren	15.3	15.3	—
How Long Responsible?	17.8	17.8	—

Variable	2000a	2000b	1990
Veteran Status	7.5	7.5	4.8
Active Duty Periods	9.8	10.3	6.1
Years of Active Duty	9.1	9.1	17.5
Employment Status Recode	11.1	11.1	3.8
Place of Work (State)	9.7	9.7	7.2
Place of Work (County)	10.1	10.1	7.9
Place of Work (MCD)	10.8	10.8	10.3
Place of Work (Place)	10.6	10.6	9.7
Place of Work (Tract)	10.2	10.2	—
Place of Work (Block)	10.2	10.2	—
Means to Work	7.6	8.2	4.6
Vehicle Occupancy	10.0	10.0	4.9
Departure Time	15.0	15.8	10.8
Travel Time	11.8	12.3	6.9
When Last Worked	11.5	14.8	6.6
Industry Last Year	14.9	15.2	8.0
Occupation Last Year	14.9	16.1	9.1
Class of Worker Last Year	17.0	17.6	9.0
Worked Last Year	9.4	13.6	13.5
Weeks Worked	19.3	20.2	14.7
Hours per Week Worked	17.4	18.1	14.5
Wages or Salary Income	20.0	20.0	10.0
Self Employment Income	9.9	9.9	6.4
Interest Income	20.8	20.8	8.1
Social Security Income	8.7	8.7	8.0
Supplemental Security Income	19.0	19.0	6.4
Public Assistance Income	18.2	18.2	7.5
Retirement Income	18.8	18.8	7.7
Other Income	18.3	18.3	7.6
All Income Imputed	24.5	24.5	9.1

NOTES: 2000a rates include imputations only, 2000b rates include imputations and assignments (comparable to 1990). —; not available.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 Sample Census Edited File (SCEF) and the 1990 Sample Edited Data File (SEDF), provided to the panel spring 2003.

Table H.7 Imputation Rates for Housing Items, 2000 and 1990 Census Long-Form Sample, Household Members (weighted)

Variable	2000a	2000b	1990
Housing Tenure	4.3	8.0	1.4
Building Type	4.4	4.4	1.6
Year Built	11.7	11.7	23.0
Year Moved In	6.2	6.2	2.9
Number of Rooms	6.2	6.2	0.4
Number of Bedrooms	8.9	10.2	7.5
Complete Plumbing	3.4	3.4	1.7
Complete Kitchen	3.4	3.4	1.8
Telephone Service	4.3	4.3	1.9
Fuel Used for Heating	7.4	7.4	2.9
Automobiles	6.2	6.2	2.2
Business on Property	8.2	8.2	2.4
How Many Acres?	10.6	10.6	4.4
Agricultural Sales	14.3	14.3	13.7
Annual Electric Cost	17.1	18.5	5.5
Annual Gas Cost	23.9	24.7	10.7
Annual Water Cost	19.6	21.8	7.3
Other Fuel Cost	28.7	31.9	17.5
Monthly Rent	15.6	15.6	1.3
Rent Includes Meals?	7.9	7.9	5.1
Mortgage?	6.0	18.6	5.7
Mortgage Payment	19.6	22.4	5.5
Payment Includes Taxes?	16.0	17.1	6.0
Payment Includes Insurance?	17.2	17.4	6.2
Second Mortgage?	11.8	16.0	5.1
Second Mortgage Payment	23.9	23.9	8.1
Property Taxes	32.0	32.0	12.2
Insurance Cost	36.6	36.6	16.8
Value of Property	13.3	13.3	3.3
Mobile Home Installment Loan	0.2	44.5	—
Mobile Home Costs	63.5	63.5	41.8

NOTES: 2000a rates include imputations only, 2000b rates include imputations and assignments (comparable to 1990). —; not available.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 Sample Census Edited File (SCEF) and the 1990 Sample Edited Data File (SEDF), provided to the panel spring 2003.

Table H.8 Imputation/Assignment Rates (percents) for Selected Population Items for Group Quarters Residents, 2000 and 1990 Long-Form Samples, by Type of Group Quarters (weighted)

Type of Group Quarters and Year	Sex	Age	Hispanic Origin	Race	Marital Status	Attend School	Grade Attend	Highest Grade
Total								
2000	3.0	3.8	8.0	4.5	18.0	31.9	30.0	39.3
1990	0.6	1.5	7.6	1.8	4.2	15.3	—	17.9
Prisons								
2000	2.7	5.5	11.8	5.4	30.9	47.7	70.1	53.8
1990	1.1	2.1	16.8	2.7	11.1	28.9	—	24.6
Juvenile Institutions								
2000	2.9	3.7	8.2	5.2	21.4	37.4	38.3	43.6
1990	0.6	3.3	5.0	2.0	2.6	12.6	—	13.7
Nursing Homes								
2000	3.4	1.8	5.1	1.5	17.7	33.1	89.1	51.7
1990	0.3	0.8	4.7	1.0	2.8	20.0	—	32.6
Hospitals and Schools for Handicapped								
2000	4.6	10.9	8.8	4.8	21.9	39.8	65.4	52.8
1990	0.6	1.2	9.4	2.1	4.6	25.8	—	30.7
College Dormitories								
2000	1.9	3.4	7.1	5.4	8.1	20.7	20.2	19.2
1990	0.2	1.3	3.4	1.4	1.2	3.2	—	2.8
Military Quarters								
2000	1.7	1.7	4.2	4.9	2.9	4.6	99.8	3.8
1990	0.5	0.9	7.1	1.4	1.5	6.4	—	5.4
Shelters								
2000	5.1	4.7	15.5	9.2	14.1	24.7	40.9	28.2
1990	2.1	3.3	20.3	5.6	13.3	23.1	—	20.6
Group Homes								
2000	4.1	2.9	6.7	3.5	17.6	30.1	49.8	42.2
1990	1.3	1.7	7.8	2.2	3.9	18.3	—	21.2
Other Group Quarters								
2000	4.4	3.7	6.6	4.5	19.4	32.0	47.6	42.9
1990	1.3	2.6	7.1	3.4	5.0	16.9	—	17.2

Table H.8 (continued)

Type of Group Quarters and Year	Speak Another Lang.	Language Spoken	English-Speaking Ability	Place of Birth	Citizen-ship	Year Entered U.S.	Where Lived 5 Yrs. Ago	State 5 Years Ago
Total								
2000	39.2	38.3	33.9	40.2	36.5	29.8	44.9	42.8
1990	18.2	27.7	22.1	19.2	14.0	30.1	18.1	19.3
Prisons								
2000	58.3	59.2	56.8	54.0	53.0	50.4	70.6	66.7
1990	29.8	35.7	29.8	31.7	24.7	38.0	33.5	32.8
Juvenile Institutions								
2000	44.4	38.0	34.1	46.2	42.1	37.9	49.9	51.5
1990	16.2	25.6	18.6	18.4	12.1	25.2	16.2	21.9
Nursing Homes								
2000	46.6	46.0	36.8	49.2	42.4	49.1	47.4	50.7
1990	25.5	34.8	26.8	25.5	18.8	46.2	23.9	29.2
Hospitals and Schools for Handicapped								
2000	50.7	47.5	40.2	54.3	47.4	41.1	53.7	59.7
1990	30.3	43.4	34.0	32.4	24.2	46.0	28.7	38.7
College Dormitories								
2000	20.6	19.0	16.3	22.2	19.9	11.3	23.7	23.9
1990	5.8	14.2	11.0	6.7	3.9	11.7	4.7	6.5
Military Quarters								
2000	4.1	7.7	4.1	5.4	3.8	10.6	12.8	6.3
1990	7.1	12.5	8.1	6.8	5.5	12.2	7.0	7.4
Shelters								
2000	28.0	31.4	24.2	35.2	32.9	27.1	45.0	34.1
1990	24.0	40.1	32.0	29.3	19.8	42.1	26.5	29.6
Group Homes								
2000	38.3	43.0	35.3	45.3	36.4	33.0	42.3	46.4
1990	20.4	31.3	22.1	24.2	15.7	35.1	21.0	24.4
Other Group Quarters								
2000	39.4	34.2	29.7	43.8	38.1	21.6	43.7	47.2
1990	18.3	22.9	20.5	18.7	14.5	26.2	18.9	23.3

Table H.8 (continued)

Type of Group Quarters and Year	Disability						Grandchildren	
	Senses	Physical Activity	Mental	Self-Care	Mobility	Work	In Home	Respon-sible for
Total								
2000	44.2	45.4	45.1	45.6	46.9	47.7	30.0	25.2
1990	—	—	—	17.3	16.7	18.1	—	—
Prisons								
2000	58.2	63.5	63.7	64.8	66.2	66.7	36.5	18.7
1990	—	—	—	33.1	31.5	34.2	—	—
Juvenile Institutions								
2000	43.3	45.1	45.9	46.2	47.9	48.4	—	—
1990	—	—	—	19.8	18.7	18.2	—	—
Nursing Homes								
2000	46.9	45.4	45.3	45.2	47.4	49.0	44.0	48.6
1990	—	—	—	21.6	21.2	23.0	—	—
Hospitals and Schools for Handicapped								
2000	56.0	55.3	54.6	55.1	56.2	57.3	47.4	25.3
1990	—	—	—	28.1	27.3	27.7	—	—
College Dormitories								
2000	21.8	22.1	21.9	22.1	22.3	22.7	0.5	15.7
1990	—	—	—	6.6	6.3	6.0	—	—
Military Quarters								
2000	—	—	—	—	—	—	83.3	14.8
1990	—	—	—	—	—	—	—	—
Shelters								
2000	30.5	32.1	29.7	30.8	32.2	33.3	39.3	14.9
1990	—	—	—	26.5	25.1	32.8	—	—
Group Homes								
2000	38.8	38.7	37.4	38.3	39.1	40.1	31.2	28.5
1990	—	—	—	20.0	19.2	25.8	—	—
Other Group Quarters								
2000	40.8	41.3	40.3	40.8	42.5	43.8	33.2	23.9
1990	—	—	—	18.2	17.4	21.2	—	—

Table H.8 (continued)

Type of Group Quarters and Year	Veteran Status	Years Served	Employment Recode	Place of Work (State)	How Get to Work	Use Carpool?	Time Leave for Work	When Last Worked
Total								
2000	39.6	35.7	5.5	11.4	12.0	14.4	20.0	45.8
1990	18.0	32.1	5.1	14.1	12.1	14.6	21.9	30.1
Prisons								
2000	57.5	54.5	—	—	—	—	—	70.2
1990	29.2	37.4	4.7	73.7	84.6	86.7	85.8	60.7
Juvenile Institutions								
2000	41.6	62.7	—	—	—	—	—	51.9
1990	16.0	59.3	0.2	14.3	14.2	50.9	41.1	44.5
Nursing Homes								
2000	48.6	52.7	—	—	—	—	—	50.1
1990	27.1	46.5	0.1	17.0	17.8	17.7	23.0	38.8
Hospitals and Schools for Handicapped								
2000	50.5	39.0	—	—	—	—	—	54.5
1990	27.9	43.6	—	21.7	28.0	71.4	44.6	47.8
College Dormitories								
2000	21.6	30.1	9.3	9.6	10.3	13.0	18.0	24.8
1990	5.7	20.8	6.6	10.2	7.7	8.3	17.8	9.7
Military Quarters								
2000	1.8	9.4	3.8	9.0	8.7	10.0	18.0	15.1
1990	2.5	26.3	—	10.0	8.4	11.3	18.1	4.1
Shelters								
2000	31.1	30.5	20.5	21.3	20.5	22.2	29.1	35.6
1990	25.8	32.9	27.6	36.0	30.9	33.6	44.2	35.1
Group Homes								
2000	38.3	43.6	10.8	12.5	13.8	15.0	20.5	43.4
1990	20.1	28.2	14.1	27.3	24.7	26.6	32.0	33.9
Other Group Quarters								
2000	40.2	49.7	22.1	24.9	28.5	33.9	36.3	46.2
1990	18.7	42.9	20.6	24.2	23.1	26.7	33.6	23.5

Table H.8 (continued)

Type of Group Quarters and Year	Industry	Occu-pation	Class of Worker	Work Last Year	Number Weeks Worked	Usual Hours Worked	Wage/ Salary Income	Self-Employ. Income
Total								
2000	46.2	46.9	54.4	47.7	42.8	41.3	50.1	42.7
1990	20.9	21.3	22.5	26.9	21.4	20.6	27.4	23.2
Prisons								
2000	75.2	75.4	78.3	70.6	72.5	71.2	74.3	67.4
1990	46.4	44.2	46.6	43.3	40.2	38.6	49.7	42.6
Juvenile Institutions								
2000	54.9	55.6	58.1	55.1	50.7	49.6	43.7	38.1
1990	25.2	25.1	25.8	26.0	23.2	23.2	21.6	18.5
Nursing Homes								
2000	78.6	78.2	80.4	52.6	69.9	69.5	48.2	47.7
1990	35.6	32.0	35.6	32.2	23.3	22.5	27.3	26.9
Hospitals and Schools for Handicapped								
2000	51.5	52.0	65.5	57.3	43.5	42.6	56.1	51.0
1990	36.9	35.9	37.1	37.8	33.9	32.5	36.6	32.5
College Dormitories								
2000	29.3	30.7	32.4	26.8	29.1	27.4	34.7	23.4
1990	9.9	11.1	11.5	11.2	13.2	12.7	15.0	10.1
Military Quarters								
2000	5.7	5.9	80.4	16.4	20.3	18.6	17.4	3.1
1990	1.1	5.7	7.8	21.3	17.5	17.2	15.8	9.6
Shelters								
2000	38.9	38.6	42.1	38.8	40.9	38.5	41.2	32.1
1990	38.2	36.4	37.2	36.6	39.0	38.1	37.1	33.3
Group Homes								
2000	42.8	43.8	44.9	46.1	46.1	44.6	50.5	39.9
1990	30.6	29.5	30.3	29.5	31.9	31.1	32.4	25.8
Other Group Quarters								
2000	43.3	43.5	46.3	49.0	51.3	50.4	49.8	41.4
1990	30.6	29.0	29.3	29.3	33.1	32.3	30.4	24.1

Table H.8 (continued)

Type of Group Quarters and Year	Interest Income	Social Security	SSI Income	Public Assist.	Retire-Ment	Other Income	All Income Imputed	Some/All Income Imputed
Total								
2000	54.5	57.4	55.5	55.2	55.5	54.5	59.9	63.1
1990	29.9	30.7	—	29.0	29.1	29.6	35.4	37.5
Prisons								
2000	74.1	74.1	74.1	74.2	74.2	73.8	78.3	79.7
1990	46.8	46.7	—	46.5	46.7	46.7	54.4	56.1
Juvenile Institutions								
2000	61.0	61.2	61.4	61.4	61.2	60.7	65.0	66.1
1990	29.1	29.1	—	29.0	28.9	28.9	32.4	33.4
Nursing Homes								
2000	68.1	75.9	68.7	68.3	69.4	67.1	76.9	79.0
1990	44.9	50.8	—	45.3	45.5	44.2	51.8	53.4
Hospitals and Schools for Handicapped								
2000	61.5	64.8	63.9	63.3	62.9	61.9	65.9	70.5
1990	38.8	40.0	—	39.2	39.0	38.8	41.3	46.6
College Dormitories								
2000	31.8	31.2	31.2	31.1	31.2	30.8	35.7	38.2
1990	11.1	10.8	—	10.6	10.6	10.7	14.9	16.5
Military Quarters								
2000	15.8	31.1	31.1	31.1	31.1	30.5	18.7	32.9
1990	10.3	0.6	—	0.6	0.6	10.2	15.3	17.2
Shelters								
2000	41.1	41.0	41.1	41.2	40.9	39.7	44.9	51.3
1990	37.5	37.7	—	37.1	37.5	37.3	40.9	44.0
Group Homes								
2000	51.6	55.1	54.9	53.0	52.6	51.1	58.3	66.0
1990	30.8	32.2	—	31.9	31.1	31.1	35.2	41.2
Other Group Quarters								
2000	56.6	59.8	57.4	55.9	57.0	55.3	63.1	67.0
1990	28.0	28.7	—	27.7	27.8	27.8	34.2	36.9

Table H.8 (continued)

NOTES: 2000 imputation rates include assignments. —, 100 percent assigned because of group quarters type (e.g., residents of nursing homes are assigned status as out of the labor force); alternatively, for 1990, question not asked (grade attending school, sensory disability, physical activity disability, mental disability, grandchildren in home, whether responsible for grandchildren, Supplemental Security Income [SSI]). Imputation rates for employment status recode include only cases for which none of the half-dozen relevant questions were answered. Some income imputed includes 100 percent and less than 100 percent imputed.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 Sample Census Edited File (SCEF) and the 1990 Sample Edited Data File (SEDF), provided to the panel summer 2003.

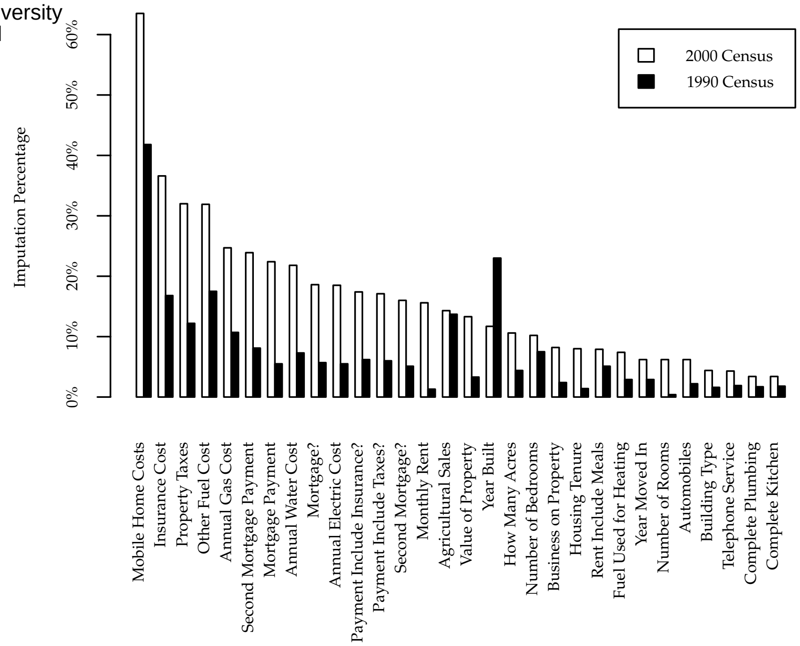

Figure H.1 Imputation/Assignment Rates for Housing Items, 2000 and 1990 Census, Persons Receiving the Long Form (weighted)

NOTES: 2000 rates include imputations and assignments (comparable to 1990), and are the “2000b” figures reported in Table H.7.

SOURCE: Tabulations by U.S. Census Bureau staff from the 2000 Sample Census Edited File (SCEF) and the 1990 Sample Edited Data File (SEDF), provided to the panel spring 2003.

Appendix I

Census 2000 Evaluations and Experiments

Every modern decennial census has included a program of formal evaluation studies intended to assess the quality of census operations and the resulting data. This appendix lists the evaluation studies planned in conjunction with the 2000 census and gives information on their status (notation if the studies were cancelled, citation information if they have been completed and released). It also briefly describes the experiments that were conducted in 2000.

I.1 CENSUS 2000 EVALUATIONS

The Census Bureau's original plan of evaluations for the 2000 census was very ambitious, including 149 studies. The Panel on Research on Future Census Methods reviewed the early evaluation plan in its first interim report (National Research Council, 2000a), and in a subsequent letter report (National Research Council, 2001c) urged the Census Bureau to "give high priority to evaluation studies and data analyses that are important to building an overall 2010 census framework."

Subsequently, in at least two major waves (early and late 2002), the evaluation program "was refined and priorities reassessed due to resource constraints at the Census Bureau" (U.S. Census Bureau, 2003a). As a result, dozens of studies were cancelled, reducing the total list of studies from 149 to 91. Eighteen planned studies were "cancelled" from the evaluation program because they were expedited and completed as part of the Executive Steering Committee for A.C.E. Policy (ESCAP) report series, as support for the Census Bureau's decisions on census adjustment in March and October 2001. The remaining studies were cancelled by the Census Bureau in an "[attempt] to obtain the best balance of resources" between "completing and releasing Census 2000 data products" and "conducting key Census 2000 evaluations" (U.S. Census Bureau, 2003a).

Response Rates and Behavior Analysis (Series A)

A.1.a: Telephone Questionnaire Assistance Operational Analysis (Chesnut, 2003b)

A.1.b: Telephone Questionnaire Assistance Customer Satisfaction Survey (Stevens, 2002)

A.2.a: Internet Questionnaire Assistance Operational Analysis: *Cancelled, early 2002*

A.2.b: Internet Data Collection Operational Analysis (Whitworth, 2002)

A.2.c: Census 2000 Internet Web Site and Questionnaire Customer Satisfaction Survey (Stapleton and Irwin, 2002)

A.3: Be Counted Campaign for Census 2000 (Carter, 2002)

A.4: Language Program—Use of Non-English Questionnaires and Guides (Smith and Jones, 2003)

A.5.a: Census 2000 Response Methods for Selected Language Groups (Lestina, 2003)

A.5.b: Awareness and Participation in the Census 2000 Language Assistance Programs Among Selected Language Groups [not yet released]

A.6.a: U.S. Postal Service Undeliverable Rates for Census 2000 Mailout Questionnaires (Kohn, 2003)

A.6.b: Detailed Reasons for Undeliverability of Census 2000 Mailout Questionnaires by the USPS (Chesnut, 2003a)

A.7.a: Census 2000 Mailback Response Rates (Stackhouse and Brady, 2003a)

A.7.b: Census 2000 Mail Return Rates (Stackhouse and Brady, 2003b)

A.8: Puerto Rico Focus Groups on Why Households Did Not Mail Back the Census 2000 Questionnaire (Berkowitz, 2002)

Content and Data Quality (Series B)

B.1.a: Analysis of the Imputation Process for 100 Percent Household Population Items (Zajac, 2003)

B.1.b: Analysis of Item Nonresponse Rates for 100 Percent Household Population Items (Norris, 2003)

B.2: Documentation of Characteristics and Data Quality by Response Type: *Cancelled, early 2002; some material incorporated into B.1*

B.3: Census Quality Survey to Evaluate Responses to the Census 2000 Question on Race (Bentley et al., 2003)

B.4: Match Study of Accuracy and Coverage Evaluation to Census to Compare Consistency of Race and Hispanic Origin Responses: *Cancelled; material subsumed by ESCAP I report B-10 (Farber, 2001a)*

B.5: Content Reinterview Survey: Accuracy of Data for Selected Population and Housing Characteristics as Measured by Reinterview (Singer and Ennis, 2002, 2003)

B.6: Master Trace Sample (Hill and Machowski, 2003)

B.7: Match Study of Current Population Survey to Census 2000 [not yet released]

B.8: Comparisons of Income, Poverty, and Unemployment Estimates Between Census 2000 and Three Census Demographic Surveys [not yet released]

B.9: Housing Measures Compared to the American Housing Survey: *Cancelled, early 2002*

B.10: Housing Measures Compared to the Residential Finance Survey: *Cancelled, early 2002*

B.11: American Community Survey Evaluation of Follow-Up, Edits, and Imputations: *Cancelled; "available data cannot answer the specified questions for this study"*

B.12: Puerto Rico Race and Ethnicity (Christenson, 2003)

B.13: Puerto Rico Focus Groups on the Race and Ethnicity Questions (Berkowitz, 2001)

Data Products (Series C)

C.1: Effects of Disclosure Limitation on Data Utility in Census 2000 (Steel and Zayatz, 2003)

C.2: Usability Evaluation of User Interface With American FactFinder: *Cancelled, early 2002; results of some focus group work summarized in separate document and presented at professional meetings*

C.3: Data Products Strategy: *Cancelled, early 2002, after some material had been shifted to C.2*

Partnership and Marketing (Series D)

D.1: Partnership and Marketing Program (Wolter et al., 2002)

D.2: Census in Schools/Teacher Customer Satisfaction Survey (Macro International, 2002)

D.3: Survey of Partners/Partnership Evaluation (Westat, 2002c)

Special Places and Group Quarters (Series E)

E.1.a: Special Place/Group Quarters Facility Questionnaire—Operational Analysis: *Cancelled, early 2002*

E.1.b: Facility Questionnaire—CATI and PV (Stevens, 2003)

E.2: Special Place Local Update of Census Addresses: *Cancelled, late 2002*

E.3: Assess the Inventory Development Process for Service-Based Enumerations: *Cancelled, mid-2002*

E.4: Decennial Frame of Group Quarters and Sources [not yet released]

E.5: Group Quarters Enumeration (Jonas, 2002, 2003b)

E.6: Service-Based Enumeration (McNally, 2002)

Address List Development (Series F)

F.1: Impact of the Delivery Sequence File Deliveries on the Master Address File Through Census 2000 Operations: *Cancelled, late 2002*

F.2: Address Listing Operation and Its Impact on the Master Address File (Ruhnke, 2002)

F.3: Local Update of Census Addresses 1998 (Owens, 2003)

F.4: Evaluation of the Census 2000 Master Address File Using Earlier Evaluation Data: *Cancelled, early 2002*

F.5: Block Canvassing Operation (Burcham, 2002)

F.6: Local Update of Census Addresses 1999 (Owens, 2002)

F.7: Criteria for the Initial Decennial Master Address File Delivery: *Cancelled, early 2002*

F.8: The Decennial Master Address File Update Rules: *Cancelled, early 2002*

F.9: New Construction Adds: *Cancelled, early 2002; some material shifted to I.4*

F.10: Update/Leave (Pennington, 2003)

F.11: Urban Update/Leave (Rosenthal, 2002b)

F.12: Update/Enumerate (Rosenthal, 2002a)

F.13: List/Enumerate (Zajac, 2002)

F.14: Overall Master Address File Building Process for Housing Units: *Cancelled, early 2002; material subsumed by Address List Development Topic Report (Vitrano et al., 2003b)*

F.15: Quality of the Geocodes Associated With Census Addresses (Ruhnke, 2003)

F.16: Block Splitting Operation for Tabulation Purposes (Green and Rothhaas, 2002)

Field Recruiting and Management (Series G)

G.1: Census 2000 Staffing Programs (Westat, 2002a,b)

G.2: Operation Control System: *Cancelled as report in this series, moved to R.2.a*

Field Operations (Series H)

H.1: Use of 1990 Data for Census 2000 Planning: *Cancelled, early 2002; some material presented at professional meetings*

H.2: Operational Analysis of Field Verification Operation for Non-ID Housing Units (Tenebaum, 2001)

H.3: Local Census Office Delivery of Census 2000 Mailout Questionnaires Returned by U.S. Postal Service With Undeliverable as Addressed Designation: *Cancelled, early 2002*

H.4: Questionnaire Assistance Centers for Census 2000 (Jones and Barrett, 2003)

H.5: Nonresponse Follow-Up for Census 2000 (Moul, 2002)

H.6: Operational Analysis of Non-Type of Enumeration Area Tool Kit Methods: *Cancelled; some material to be contained in separate research paper*

H.7: Nonresponse Follow-Up Enumerator Training (Burt and Mangaroo, 2003)

H.8: Operational Analysis of Enumeration of Puerto Rico (McNally, 2003)

H.9: Local Census Office Profile (Imel, 2003)

H.10: Date of Reference for Respondents of Census 2000 (Carter and Brady, 2002)

Coverage Improvement (Series I)

I.1: Coverage Edit Follow-Up for Census 2000 (Sheppard, 2003)

I.2: Nonresponse Follow-Up Whole Household Usual Home Elsewhere Probe (Viator and Alberti, 2003)

I.3: Nonresponse Follow-Up Mover Probe (Keathley, 2003)

I.4: Coverage Improvement Follow-Up (Moul, 2003)

I.5: Coverage Gain from Coverage Questions on Enumerator Completed Questionnaire (Nguyen and Zelenak, 2003)

I.6: Coverage, Rostering Methods and Household Composition: A Comparative Study of the Current Population Survey and Census 2000: *Cancelled, early 2002*

Ethnographic Studies (Series J)

J.1: Coverage, Rostering Methods and Household Composition: A Comparative Study of the Current Population Survey and Census 2000: *Cancelled, after having been reclassified under I.6*

J.2: Ethnographic Social Network Tracing (Brownrigg, 2003)

J.3: Comparative Ethnographic Research on Mobile Populations (Hunter et al., 2003)

J.4: Colonias on the U.S./Mexico Border: Barriers to Enumeration in Census 2000 (de la Puente and Stemper, 2003)

Data Capture (Series K)

K.1.a: Data Capture Audit Resolution Process (Rosenthal, 2003a)

K.1.b: Quality of the Data Capture System and the Impact of Questionnaire Capture and Processing on Data Quality (Conklin, 2003)

K.1.c: Analysis of Data Capture System 2000 Keying Operations: *Cancelled, early 2002*

K.1.d: Synthesis of Results from K.1.a, K.1.b, and K.1.c: *Cancelled, subsumed by topic report*

K.2: Analysis of the Interaction Between Aspects of Questionnaire Design, Printing, and Completeness With Data Capture: *Cancelled, early 2002; some material shifted to K.1.a and K.1.b*

K.3: Impact of Data Capture Errors on Autocoding, Clerical Coding and Autocoding Referrals in Industry and Occupation Coding [not yet released]

K.4: Performance of the Data Capture System 2000: *Cancelled, early 2002; some material contained in Titan Corporation (2003) and R.3.d*

Processing Systems (Series L)

L.1: Invalid Return Detection

L.2: Decennial Response File Stage 2 Linking and Setting of Expected Household Population (Rosenthal, 2003b)

L.3.a: Analysis of Primary Selection Algorithm Results (Operational Assessment) (Baumgardner, 2002)

L.3.b: Resolution of Multiple Census Returns Using Reinterview (Baumgardner, 2003)

L.4: Census Unedited File Creation (Jonas, 2003a)

L.5: Beta Site (Titan Systems Corporation, 2003)

Quality Assurance Evaluations (Series M)

M.1: Evaluation of the Census 2000 Quality Assurance Philosophy and Approach Used for the Address List Development and Enumeration Operations (Morganstein et al., 2003)

M.2: Effectiveness of Existing Variables in the Model Used to Detect Discrepancies During Reinterview, and the Identification of New Variables (Johanson, 2003)

Accuracy and Coverage Evaluation Survey Operations (Series N)

N.1: Contamination of Census Data Collected in A.C.E. Blocks (Bench, 2002)

N.2: Analysis of Listing Future Construction and Multi-Units in Special Places: *Cancelled, early 2002*

N.3: Analysis of Relisted Blocks: *Cancelled, early 2002*

N.4: Analysis of Blocks With No Housing Unit Matching: *Cancelled, early 2002*

N.5: Analysis of Blocks Sent Directly for Housing Unit Follow-Up: *Cancelled, early 2002*

N.6: Analysis of Person Interview With Unresolved Housing Unit Status: *Cancelled, early 2002*

N.7: Analysis on the Effects of Census Questionnaire Data Capture in A.C.E.: *Cancelled, early 2002*

N.8: Analysis of the Census Residence Questions Used in A.C.E.: *Cancelled; material subsumed by ESCAP report B-16 (Stiers, 2000)*

N.9: Analysis of the Person Interview Process: *Cancelled; material subsumed by ESCAP report B-5 (Feindt and Byrne, 2000)*

N.10: Discrepant Results in A.C.E. (Krejsa, 2003)

N.11: Extended Roster Analysis: *Cancelled, early 2002*

N.12: Matching Stages Analysis: *Cancelled; subsumed by ESCAP report B-6 (Childers et al., 2001)*

N.13: Analysis of Unresolved Codes in Person Matching: *Cancelled; some material subsumed by ESCAP report B-11 (Starsinic et al., 2001)*

N.14: Evaluation of Matching Error (Bean, 2002)

N.15: Outlier Analysis in the 2000 A.C.E.: *Cancelled, early 2002*

N.16: Impact of Targeted Extended Search: *Cancelled; some material shifted to N.17*

N.17: Targeted Extended Search Block Cluster Analysis (Wolfgang et al., 2002)

N.18: Effect of Late Census Data on Final Estimates: *Cancelled; some material subsumed by ESCAP II report 13 (Raglin, 2001)*

N.19: Field Operations and Instruments for A.C.E. (Green et al., 2003)

N.20: Group Quarters Analysis: *Cancelled, early 2002*

N.21: Analysis of Mobile Homes: *Cancelled, early 2002*

Coverage Evaluations of the Census and of the Accuracy and Coverage Evaluation Survey (Series O)

O.1: Type of Enumeration Area Summary: *Cancelled, early 2002*

O.2: Coverage of Housing Units in the Early Decennial Master Address File: *Cancelled, early 2002*

O.3: Census 2000 Housing Unit Coverage Study (Barrett et al., 2003)

O.4: Analysis of Conflicting Households (Liu et al., 2002)

O.5: Analysis of Proxy Data in the A.C.E. (Wolfgang et al., 2003)

O.6: P-Sample Nonmatches Analysis: *Cancelled; some material subsumed by Wolfgang et al. (2001)*

O.7: Analysis of Person Coverage in Puerto Rico: *Cancelled, early 2002*

O.8: Analysis of Housing Unit Coverage in Puerto Rico: *Cancelled, early 2002*

O.9: Geocoding Error Analysis: *Cancelled; some material subsumed by Feldpausch (2002) and Keathley et al. (2001)*

O.10: Housing Unit Duplication in Census 2000 (Jones, 2003a)

O.11: E-Sample Erroneous Enumeration Analysis: *Cancelled; some material subsumed by Feldpausch (2002)*

O.12: Analysis of Nonmatches and Erroneous Enumerations Using Logistic Regression: *Cancelled; some material subsumed by Beaghen et al. (2001)*

O.13: Analysis of Various Household Types and Long Form Variables: *Cancelled, late 2002*

O.14: Measurement Error Reinterview Analysis: *Cancelled; some material subsbumed by Adams and Krejsa (2001); Krejsa and Raglin (2001)*

O.15: Impact of Housing Unit Coverage on Person Coverage Analysis: *Cancelled; some material covered by Robinson and Wolfgang (2002)*

O.16: Person Duplication in Census 2000; revised to Person Duplication in the Search Area Measured by the 2000 Accuracy and Coverage Evaluation (Jones, 2003b)

O.17: Analysis of Households Removed Because Everyone in the Household Is Under 16 Years of Age: *Cancelled, early 2002*

O.18: Synthesis of What We Know About Missed Census People: *Cancelled, early 2002*

O.19: Analysis of Deleted and Added Housing Units in Census 2000 Measured by the Accuracy and Coverage Evaluation (Smith et al., 2003)

O.20: Consistency of Census Estimates with Demographic Benchmarks (Adlakha et al., 2003)

O.21: Implications of Net Census Undercount on Demographic Measures and Program Uses: *Cancelled, prior to February 2001*

O.22: Evaluation of Housing Units Coded as Erroneous Enumerations: *Cancelled, after having been added after February 2001; some material subsumed by Adams and Liu (2001)*

O.23: Analysis of Insufficient Information for Matching and Follow-Up: *Cancelled, after having been added after February 2001; some material subsumed by Feldpausch (2002)*

O.24: Evaluation of Lack of Balance and Geographic Errors Affecting Person Estimates: *Cancelled, after having been added after February 2001; some material subsumed by Adams and Liu (2001)*

O.25: Mover Analysis: *Cancelled, after having been added after February 2001; some material subsumed by Liu et al. (2001)*

O.26: Analysis of Balancing in the Targeted Extended Search: *Cancelled, after having been added after February 2001; some material subsumed by Hogan (2001a) and Adams and Liu (2001)*

Accuracy and Coverage Evaluation Survey Statistical Design and Estimation (Series P)

P.1: Measurement of Bias and Uncertainty Associated With Application of the Missing Data Procedures: *Cancelled; some material subsumed by Keathley et al. (2001)*

P.2: Synthetic Design Research/Correlation Bias: *Cancelled, early 2002*

P.3: Variance of Dual System Estimates and Adjustment Factors: *Cancelled; some material subsumed by Starsinic et al. (2001)*

P.4: Overall Measures of A.C.E. Quality: *Cancelled*

P.5: Total Error Analysis: *Cancelled, early 2002*

Organization, Budget, and Management Information System (Series Q)

Q.1: Management Processes and Systems of the 2000 Decennial Census (IBM Business Consulting Services, 2003)

Automation of Census Processes (Series R)

R.1.a: Telephone Questionnaire Assistance (Titan Systems Corporation, 2001d)
R.1.b: Coverage Edit Follow-Up (Titan Systems Corporation, 2001a)
R.1.c: Internet Questionnaire Assistance (Titan Systems Corporation, 2001c)
R.1.d: Internet Data Collection (Titan Systems Corporation, 2001b)
R.2.a: Operations Control System 2000 (Titan Systems Corporation, 2002g)
R.2.b: Laptop Computers for Accuracy and Coverage Evaluation (Titan Systems Corporation, 2002d)
R.2.c: Accuracy and Coverage Evaluation 2000 Control System (Titan Systems Corporation, 2002a)
R.2.d: Matching and Review Coding System for the Accuracy and Coverage Evaluation (Titan Systems Corporation, 2002f)
R.3.a: Pre-Appointment Management System/Automated Decennial Administrative Management System (Titan Systems Corporation, 2002h)
R.3.b: American FactFinder (Titan Systems Corporation, 2002b)
R.3.c: Management Information System 2000 (Titan Systems Corporation, 2002e)
R.3.d: Census 2000 Data Capture (Titan Systems Corporation, 2002c)

I.2 CENSUS 2000 EVALUATION TOPIC REPORTS

The Census Bureau issued a series of "topic reports," which were intended to synthesize the results of multiple individual-topic evaluations. The evaluation studies referenced in those topic reports were to be made public at the same time that the topic report was made public. Topic reports began to be issued in summer 2003.

The topic reports are:

- Address List Development (Vitrano et al., 2003b,a)
- Automation of Census 2000 Processes (Dawson and Stoudt, 2003)
- Content and Data Quality (Schneider, 2003, 2004)
- Coverage Improvement (Clark and Moul, 2003)
- Coverage Measurement (Petroni and Childers, 2003)
- Data Capture (Titan Corporation, 2003)

- Data Collection (Hough and Borsa, 2003)
- Data Processing (Alberti, 2003)
- Ethnographic Studies (de la Puente, 2004)
- Partnership and Marketing Program (Edwards and Wilson, 2003)
- Privacy Research in Census 2000 (Singer, 2003)
- Puerto Rico (Hovland and Buckley-Ess, 2003)
- Race and Ethnicity (del Pinal, 2003)
- Response Rates and Behavior Analysis (Treat, 2003)
- Special Places and Group Quarters (Abramson, 2003)

I.3 CENSUS 2000 EXPERIMENTS

Census 2000 Alternative Questionnaire Experiment (AQE2000)

This experiment used additional mailings and reinterview studies to manipulate three questionnaire design components:

- *Presentation of residence rules on the short form:* Does providing a brief, reformatted version of the rules improve data quality?

- *Comparing the 1990 and 2000 census presentation of race and Hispanic origin questions:* The two censuses presented the questions in slightly different ways (see Figures 8.1 and 8.2), not only in wording but also in format and design.

- *Design of "skip to" and "go to" instructions in the census long form:* Determine whether respondents were able to navigate through the paper question correctly and efficiently.

The results are reported in Gerber et al. (2002); Martin (2002); Redline et al. (2002); Martin et al. (2003).

Administrative Records Census 2000 Experiment (AREX 2000)

Initial planning for the 2000 census included experimentation with an administrative records census as a possible way to save costs; the idea had been raised but not endorsed by National Research Council (1995b:Ch.4) and National Research Council (1999b:Ch.5). The AREX 2000 experiment assembled national-level administrative records (unduplicated using Social Security Numbers) and assigned block-level geographic codes. Records for the five selected test sites were then extracted and tallied at the census block level. A separate branch of the experiment sought to reconcile administrative records with the Master Address File to generate block-level population and housing unit counts.

The results of the experiment are reported in Bauder and Judson (2003); Berning and Cook (2003); Berning (2003); Heimovitz (2003); Judson and Bye (2003).

Social Security Number, Privacy Attitudes, and Notification Experiment (SPAN)

Related to the administrative records research, the SPAN experiment probed for behavioral and attitudinal data on public response to queries for their Social Security Numbers (SSNs) on census questionnaires. The experiment also tested public response to variations in wording in notices about Census Bureau use of administrative records, as well as surveying public concerns about privacy and confidentiality raised by use of administrative records.

The results of the experiment are reported in Brudvig (2003); Guarino et al. (2001); Trentham and Larwood (2003).

Response Mode and Incentive Experiment (RMIE)

This experiment studied the effectiveness of three electronic modes of data collection:

- *Operator telephone interview:* Also known as reverse computer-assisted telephone interview (Reverse CATI); respondents were encouraged to call a toll-free telephone number, at which time a telephone interviewer administered the questionnaire.

- *Computer telephone interview:* Also known as the Automated Spoken Questionnaire; respondents were asked to call a toll-free telephone number, at which time the short-form questionnaire was administered by interactive voice response (an automated system).

- *Internet:* Respondents were encouraged to answer the questionnaire using a Web address provided in a cover letter.

The experiment also tested the impact on response of offering an incentive for completing the questionnaire (specifically, a telephone calling card valid for 30 minutes of free long-distance calling).

The results were reported in Caspar (2003); Guarino (2001); Schneider et al. (2002).

Census 2000 Supplementary Survey

The Census 2000 Supplementary Survey extended pilot work on the American Community Survey (ACS), the Census Bureau's planned continuous measurement survey to replace the census long form in 2010. In addition to data collection in 36 ongoing ACS test sites, the Supplementary Survey extended data collection to 1,203 additional counties. The Census 2000 Supplementary Survey was conducted as an experiment, with the intent of determining whether it is feasible to collect long-form-census data at the same time, but in a separate process from, the decennial census data collection. The Census Bureau concluded that this simultaneous collection is feasible and that ACS work is feasible for a full national sample; the results are reported in Griffin and Obenski (2001).

Privacy Schemas and Data Collection: An Ethnographic Account

The goal of the experiment was to collect qualitative and attitudinal data on survey participation and response, including further probing of privacy concerns and elaborating reasons for choosing to participate in survey data collections. The results of the study are in Gerber (2003).

Complex Households and Relationships in the Decennial Census and Demographic Surveys (Ethnographic Studies)

This ethnographic research project assembled six teams to study how well census methods, questions, and categories matched the diversity and experience of modern households. The six teams targeted particular ethnic or race groups: African Americans, Hispanics, Inupiaq Eskimos, Koreans, Navajos, and whites. The results are reported in Schwede (2003).

Generation X Speaks Out on Censuses, Surveys, and Civic Engagement: An Ethnographic Approach (Ethnographic Studies)

This ethnographic study was intended to probe the civic engagement and attitude toward censuses (and surveys in general) among the Generation X population, those born between 1968 and 1979. Within this age cohort, differences by other factors—socioeconomic background, ethnicity, immigrant status, and so forth—were also considered. Members of the subsequent Millenial generation (14–18 years of age) were also interviewed for comparison. The results are reported in Crowley (2003).

Glossary and Abbreviations

Accuracy and Coverage Evaluation (A.C.E.): A *coverage evaluation* program conducted by the Census Bureau following the 2000 census; it produces estimates of *undercount* and *overcount* in the census and forms the basis for statistical adjustment of census counts through *dual-systems estimation*. In the A.C.E., a sample survey is conducted in a sample of census *block clusters* after the *nonresponse follow-up* phase of the census is complete. The resulting sample of individuals found by the survey in the selected block clusters—called the *P-sample*—is matched to the set of census enumerations from the sample block clusters (the *E-sample*).

A.C.E.: *Accuracy and Coverage Evaluation*

A.C.E. Revision II: The estimates and evaluation research produced between October 2001 and December 2002 as a further revision to the original *Accuracy and Coverage Evaluation*. The Revision II work informed the Census Bureau's March 2003 decision not to use statistically adjusted census figures in calculating intercensal population estimates. See *Accuracy and Coverage Evaluation*.

ACF: *Address Control File*

ACS: *American Community Survey*

Address Control File (ACF): The 1990 census analogue to the *Master Address File* used in 2000. The ACF was the residential address list used to label questionnaires, control the mail response check-in operation, and determine the *nonresponse follow-up* workload.

Address List Improvement Act of 1994 (P.L. 103-430): The law that enabled two innovations in the construction of the *Master Ad-*

dress File for the 2000 census: the *Local Update of Census Addresses* Program (allowing local governments to receive and review the address list) and address list updates from the U.S. Postal Service's *Delivery Sequence File.*

Administrative records: Records that are collected as part of the operation of federal, state, and local programs, typically fund allocation and tax programs, such as Internal Revenue Service and Food Stamp Program records.

American Community Survey (ACS): A continuous survey program under development by the Census Bureau to collect the detailed socioeconomic and other data currently asked of the census *long-form sample*. ACS estimates would be based on monthly surveys of respondents and released annually; for smaller population groups, estimates would be based on 3 or 5 years of data. The Census Bureau hopes that implementation of the American Community Survey will allow the switch to a short-form-only census in 2010. Pilot ACS data collection began in 1996, and a larger prototype (the *Census 2000 Supplementary Survey*) was fielded in 2000. Data collection continued at the 2000 level from 2001–2003.

American FactFinder: The Internet site established and hosted by the Census Bureau as a primary means of disseminating data from the 2000 census, the 1990 census, the *Census 2000 Supplementary Survey*, and the *American Community Survey*. It is accessible at http://factfinder.census.gov [1/10/04].

Apportionment: The reallocation of the 435 seats in the U.S. House of Representatives, on the basis of population size, using data from a new decennial census. Under current law, apportionment is conducted using the method of *equal proportions*.

Balancing error: Type of error cited by the *Executive Steering Committee for A.C.E. Policy* in its March 2001 recommendation not to adjust census counts for congressional redistricting. Balancing error occurs when cases in the *P-sample* and *E-sample* are not treated identically (e.g., when the search area used to identify P-sample matches and E-sample correct enumerations is defined differently).

BAS: *Boundary and Annexation Survey*

Basic data; basic data item: See *complete count*.

Basic street address: The portion of an address consisting of a house number with a street or road number (but not designation for apartments, units, or subdivisions within structures).

Be Counted: A program in the 2000 census that made census questionnaires available in public places, so that residents who believed that they had been missed in the regular census enumeration could file a questionnaire.

Block: See *census block*.

Block canvass: A field operation implemented shortly before the 2000 census to improve the completeness of the *Master Address File*. Census Bureau field staff were assigned to verify information for every *mailout/mailback* address and to add addresses for housing units and living quarters when possible. In the 2000 census process, the Census Bureau decided to perform a complete block canvass after concluding that the *Delivery Sequence File* and updates from the *Local Update of Census Addresses* Program were insufficient as primary sources of address updates.

Block cluster: Group of one or more *census blocks* expected to contain about 30 housing units, defined for use in the *Accuracy and Coverage Evaluation* Program.

Blue line: A Census Bureau-defined boundary that is meant to identify areas that have predominantly *city-style addresses* and those that do not. Areas "inside the blue line" are said to have city-style addresses; those "outside the blue line" typically have non-city-style addresses. Named after the color originally used to draw the boundary on early maps, the blue line was used to identify local and tribal government eligibility for the 1998 or the 1999 wave of the *Local Update of Census Addresses* Program.

Boundary and Annexation Survey: An annual survey by the Census Bureau of all counties (or equivalents), minor civil divisions, incorporated places, and American Indian reservations to determine their legal geographic boundaries.

BSA: *Basic street address*

C2SS: *Census 2000 Supplementary Survey*

CAI: *Computer-Assisted Interviewing*

CAPI: *Computer-assisted personal interviewing*

Casing check: A program in which postal workers determine addresses for which they did not receive a questionnaire and notify the Census Bureau.

CATI: *Computer-assisted telephone interviewing*

CAUS: *Community Address Updating System*

CCF: *Coverage correction factor*

CEFU: *Coverage edit follow-up*

Census 2000 Supplementary Survey (C2SS): Pilot program for the *American Community Survey*; a survey to collect data items from the census long form that was conducted in monthly samples totaling 700,000 households in 2000.

Census Act: See *Title 13*.

Census block: The smallest entity for which the Census Bureau collects and tabulates decennial census information; bounded on all sides by visible and nonvisible features shown on Census Bureau maps. Occasionally, especially in rural areas, drainage ditches or power lines may be used to define blocks. Because most blocks have small population and housing unit counts, only *complete-count* data are tabulated for them.

Census Day: The target date of a decennial census. Census Day is the date for which census respondents are supposed to describe their household population, and for which the results of a decennial census are supposed to be an accurate representation of the nation's population. Since 1930, Census Day has been April 1 of years ending in zero.

Census Monitoring Board: An entity established by law (P.L. 105-119) during preparations for the 2000 census that was charged with observing and monitoring all aspects of census implementation. Members of the board appointed by House and Senate Republican leaders and by the president (in consultation with House and Senate Democratic leaders) functioned independently, with separate staffs and reports. Under the terms of the enacting legislation, the Census Monitoring Board ceased to exist on September 30, 2001.

Census Plus: A method considered as an alternative to *dual-systems estimation* in 2000, in which enumerators would revisit a sample of households to obtain a roster of household members and immediately reconcile that roster with the roster from the census. When weighted, the union of the two rosters (after making any deletions or additions) would provide an estimate of the population. Plans for Census Plus were set aside following the results of a 1995 census test.

Census tract: A census-defined geographic area of roughly 2,500 households. Census tracts are aggregations of *census blocks* (roughly 150 blocks, dependent on the population of the area). Tracts are intended to be relatively stable entities over time, though their definitions do shift with each census.

CIFU: *Coverage improvement follow-up*

Closeout: The last stage of *nonresponse follow-up* when enumerators are instructed to make a last attempt to obtain at least minimal information, from a *proxy* if necessary. *Imputation* is used to fill in any missing information.

CNSTAT: Committee on National Statistics

Coefficient of variation (CV): An assessment of the variability of an estimate, calculated as the ratio of the standard error of an estimate to the value of the estimate. This is expressed as a percentage of the size of the quantity being measured.

Community Address Updating System (CAUS): A program to be initiated under the Census Bureau's *MAF/TIGER Enhancements Program* leading up to the 2010 census to improve geographic coverage, particularly in rural areas. The CAUS Program uses *American Community Survey* enumerators to collect geographic updates as they perform their duties, using laptop computer systems equipped with *GPS* receivers.

Complete count; complete-count items: The basic data items asked of all census respondents, whether they received the census *short form* or the *long form* (the long form asks additional *sample items* of an approximate one-sixth sample of the population). In 2000, the complete-count items were name, age, sex, race, Hispanic origin, date of birth, relationship to census respondent (reference person), and housing tenure (own or rent).

Computer-assisted interviewing (CAI): A group of methods for using computers to assist with data collection. CAI surveys can be either interviewer-administered (conducted in person using a laptop computer or by telephone using a shared computer) or self-administered (conducted by surveys disseminated to respondents by telephone, by the Internet, or on a computer disk). See *computer-assisted personal interviewing; computer-assisted telephone interviewing*.

Computer-assisted personal interviewing (CAPI): The use of a computer to assist an interviewer in carrying out an interview.

Advantages include avoiding errors in skip patterns, providing immediate edit checks, and expediting electronic data capture.

Computer-assisted telephone interviewing (CATI): Method of data collection featuring interviewers administering survey interviews via telephone, reading questions presented on a computer screen, and recording responses on the computer.

Continuation form: A census questionnaire, asking items for up to six additional household members, used in follow-up interviewing in 2000 if there were six or more people in a household (the standard enumerator form allowed for only five household members in 2000). Multiple continuation forms could be completed for a household, as necessary.

Count Question Resolution (CQR): A process established by the Census Bureau to enable state, local, and tribal governments to challenge their 2000 census counts if they believed that inappropriate tabulation boundaries had been used or that specific living quarters had been miscounted. Upon review, the local governments could be issued a letter attesting to a revised count for use in such purposes as fund allocation; the enabling regulation stipulated that counts revised under CQR would not affect a state's *apportionment* count. The CQR program began in 2001 and ended in 2003.

Correlation bias: A (technical) bias in *dual-systems estimation* by which the estimated counts would be, on the average, either too low or too high, caused by heterogeneity in enumeration probabilities for both the census and the *postenumeration survey*. The heterogeneities of the probabilities for these two attempted enumerations are typically positively related, which causes the estimated counts to be on the average too low. The *A.C.E. Revision II* estimates included an attempted correction for correlation bias.

COTS: Commercial off-the-shelf

Coverage correction factor (CCF): In *dual-systems estimation*, the dual-systems estimate for a *poststratum* divided by the census count (including whole person imputations and late additions). The CCF is interpreted as the multiplier that can be applied to the population count for a poststratum in a particular area to generate an adjusted count.

Coverage edit follow-up (CEFU): A 2000 census follow-up operation of mail-return households whose census responses showed population count discrepancies (e.g., the number of people for whom census information was included on the form did not match the number of residents reported elsewhere on the questionnaire). In particular, CEFU concentrated on large households (those with seven or more members for which there was room to report characteristics for only the first six members on the mail questionnaire).

Coverage evaluation: Statistical studies conducted to evaluate the level and sources of coverage error in censuses and surveys.

Coverage improvement follow-up (CIFU): The second-stage follow-up operation used in the 2000 census (performed between June and August, 2000), verifying findings from the initial *nonresponse follow-up*.

Coverage improvement programs: Often (but not always) nationally applied methods and programs that attempt to collect information from individuals and households that might be missed using *mailout/mailback* or *nonresponse follow-up*. Before the 2000 census cycle this term referred to such programs as the parolee and probationer program (used in 1990), in which lists of these individuals were checked to see whether they were enumerated, and the non-household sources program, in which several *administrative record* lists were matched to census records to try to identify people missed in the census for purposes of field follow-up (used in 1980). For the 2000 census, coverage improvement refers more to efforts to complete the address list, use of *multiple response modes*, and *service-based enumeration*.

CPS: *Current Population Survey*

Curbstoning: The practice by which a census enumerator fabricates a questionnaire for a residence without actually visiting it.

Current Population Survey (CPS): Monthly sample survey of the U.S. population that provides employment and unemployment figures as well as current data about other social and economic characteristics of the population. The CPS is collected for the Bureau of Labor Statistics by the Census Bureau. The sample size for the CPS is about 50,000 households per month.

CV: *coefficient of variation*

Data capture: The process by which survey responses are transferred from written questionnaires to an electronic format for tabulation. In the 2000 census, data capture was done by *optical character* and *optical mark recognition*; from 1890 to 1950, punch cards were used for data capture, and the *FOSDIC* process of *optical mark recognition* was used from 1960 to 1990.

Data Capture System 2000 (DCS 2000): The system used in the 2000 census to digitally extract and capture information from paper census forms. A replacement of the *FOSDIC* system used in earlier census, the Data Capture System 2000 made use of *optical character recognition*.

Data-defined: In census data processing, an assessment of the completeness of a census record. For the *complete count* (basic data items), a household is data-defined if at least one member has reported values for at least two complete-count items (including name). For the *long-form sample*, a record for a household is said to be data-defined if at least one member of the household has at least two sample data items reported.

Data Preparation Division: Until 1998, the name of the Census Bureau's permanent processing center in Jeffersonville, Indiana (now known as the *National Processing Center*).

DCS 2000: *Data Capture System 2000*

Decennial Master Address File (DMAF): See *Master Address File*.

Delivery Sequence File (DSF): The master list of deliverable mail addresses maintained by the U.S. Postal Service, organized by carrier route. The Delivery Sequence File was first used as a source of updates to the *Master Address File* for the 2000 census, following enactment of the *Address List Improvement Act of 1994*.

Demographic analysis: A method that uses various *administrative records* (especially birth and death records, information on immigration and emigration, and Medicare records) and information from previous censuses to estimate the total number of people in various demographic groups resident in the United States on a specific date, and therefore their census undercoverage.

DMAF: Decennial Master Address File; see *Master Address File*.

Domain: A classification based on race and ethnicity (Hispanic origin) used in the definition of *poststrata* in the *Accuracy and Coverage Evaluation*.

Dress rehearsal: The largest census test, typically 2 years before the decennial census, in which the methods and procedures of the upcoming decennial census are given their final test to identify any operational problems.

DSE: *Dual-systems estimation*

DSF: *Delivery Sequence File*

Dual-systems estimation (DSE): An estimation methodology that uses two independent attempts to collect information from households to estimate the total population, including the number of people missed by both attempts.

Enumerator: A census field operations employee who collects information from respondents through interviews.

Equal proportions: Since 1941, the method used in *apportionment* of the U.S. House of Representatives. Under the method, priority values for seats are generated by multiplying a state's apportionment population by the reciprocal of the geometric mean $\left(1/\sqrt{n(n-1)}\right)$; for example, a state's priority value for its tenth seat in the House equals its population multiplied by $1/\sqrt{90}$. Priority values for all states are ranked and seats are assigned beginning with the 51st seat (each state's second seat; the Constitution provides each state with a minimum of one seat in the House.)

Erroneous enumeration: The inclusion of someone in the census in error. Such inclusions may be people born after *Census Day* or deceased before Census Day, people in the United States temporarily, and people in the wrong location. They also include people counted more than once, i.e., duplicates.

Error: The difference between an estimate and the true value.

E-sample: The set of census enumerations for a sample of census *block clusters*; part of the *Accuracy and Coverage Evaluation*, it is used to calculate the correct enumeration rate in the *dual-systems estimation* formula.

ESCAP: *Executive Steering Committee for A.C.E. Policy*

Executive Steering Committee for A.C.E. Policy (ESCAP): The committee of senior Census Bureau staff charged with analyzing information from the 2000 census and *Accuracy and Coverage Evaluation* in order to decide whether census counts should be adjusted for estimated net *undercount*. The ESCAP reported to the director of the Census Bureau, who in turn submitted a formal recommendation to the U.S. secretary of commerce.

Follow-up: A secondary census or survey operation, predominantly in data collection, carried out to successfully complete an initial operation. It is most often a telephone or personal visit interview to obtain missing data or clarify original responses.

FOSDIC (Film Optical Sensing Device for Input to Computers): From 1960 to 1990, census questionnaires were microfilmed. The answers were read from the microfilmed questionnaires using FOSDIC and converted to electronic codes on computer tape.

FSPD: *Further Study of Person Duplication*

Further Study of Person Duplication (FSPD): A research study completed as part of *A.C.E. Revision II* that made further refinements to the *Person Duplication Studies* conducted in summer 2001. The FSPD included improved statistical matching techniques and assessed probabilities of duplicate links.

FWA: Final weighting area

Geocoding: The assignment of a geographic location code to an address or a map spot (longitude/latitude location). For example, geocoding may identify the *census block* in which a street address is located; with that knowledge, information from that street address can be associated with the correct block (and, thus, the correct census tract, place, and higher-level geographic aggregates). Geocoding is a major function of the Census Bureau's *TIGER* database.

GIS: Geographic information system

GPS: Global positioning system

GQ: *Group quarters*

Gross error: The sum of erroneous enumerations and omissions in the census. See also *Erroneous enumeration*; *Omission*; *Overcount*; *Undercount*.

Group quarters (GQ): A place where people live that is not a housing unit. There are two types of group quarters: institutional (for example, nursing homes, mental hospitals, and correctional institutions) and noninstitutional (for example, college dormitories, ships, hotels, group homes, and shelters). A structure that houses group quarters may also include one or more housing units (e.g., the apartment for a resident faculty member in a dormitory).

Hot-deck imputation: The technique used by the Census Bureau to impute missing responses on census questionnaires. Imputations are made based on a continually updated distribution of responses from other, filled-in questionnaires that match characteristics that are known from an incomplete questionnaire.

Household: All the persons who occupy a *housing unit* as their *usual place of residence*.

Housing unit (HU): A house, an apartment, etc., that is occupied (or, if vacant, is intended for occupancy) as separate living quarters, which are those in which the occupants live and eat separately from any other persons in the building. See also *household*.

HU: *housing unit*

Hundred percent data: See *complete count*. Tabulations from the Hundred Percent Census Edited File—the hundred percent data, after consistency edits and imputation has been performed—are used to generate apportionment counts.

ICM: *Integrated Coverage Measurement*

***II*s:** Census respondents whose records contain insufficient information for matching, such as would be necessary to obtain adjusted counts through *dual-systems estimation*. For the 2000 census in particular, the set of *II*s contained both persons with substantially incomplete questionnaires and people who were reinstated in the census count at a late stage of processing but excluded from the *Accuracy and Coverage Evaluation*.

Imputation: A method for filling in missing information. Sequential *hot-deck imputation* fills in information from a previously processed respondent (and therefore geographically close) with other similar characteristics.

Individual census report: A special version of the census questionnaire used during *group quarters* enumeration and other special operations that asks questions regarding only one person (rather than up to six persons in a household, as on the usual census short and long forms).

Inmover: A person who moved into a *housing unit* after *Census Day* but before the reference date for a *postenumeration survey*, as in the *Accuracy and Coverage Evaluation*.

Integrated coverage measurement (ICM): The use of a *postenumeration survey* and some type of estimation method, e.g., *dual-systems estimation*, to produce adjusted census counts in time for apportionment and therefore all uses of census data. ICM was a key part of initial Census Bureau plans for the 2000 census, but was abandoned after the Supreme Court's 1999 decision ruling out the use of sampling in generating apportionment counts.

Interactive voice response (IVR): A technology for survey data collection in which surveys are conducted using an automated telephone system, using voice recognition to allow respondents to select from possible question responses.

IRS: Internal Revenue Service

IVR: *Interactive voice response*

IWA: Initial weighting area

Key from image (KFI): A *data capture* operation in which the data from questionnaires whose entries could not be parsed by *optical character recognition* with sufficient confidence were entered by hand by census staff.

KFI: key from image

Last resort: Term used in the 1990 census to describe the collection of data from neighbors, apartment managers, post office employees, etc., when a response from a resident could not be obtained.

LCO: *local census office*

List/enumerate: A method of enumeration in which enumerators canvass a geographic area, list each residential address, and collect a questionnaire from or enumerate a household.

List/leave: A method of enumeration in which the enumerators list each residential address and at the same time deliver the census form for return by mail.

Local census office (LCO): A temporary office established for decennial census operations, including coordination of address listing and nonresponse follow-up operations.

Local review: Census Bureau program in the 1980 and 1990 censuses in which local officials were given the opportunity to review housing unit counts in census blocks.

Local Update of Census Addresses (LUCA): A Census Bureau program in which local officials were given the opportunity to review individual addresses on the *Master Address File* and make corrections, additions, and deletions to that list, and to make corrections to census maps to match any changes that may be needed. The LUCA 98 Program covered only local governments in *mailout/mailback* enumeration areas; LUCA 99 covered governments in *update/leave* enumeration areas.

Long-form data-defined: See *data-defined*.

Long-form sample: The approximate one-sixth sample of the population that is asked an additional set of socioeconomic and housing characteristics questions (the *sample items*) in addition to the *complete-count* items asked of everyone. See also *complete count*.

LUCA: *Local Update of Census Addresses*

MAF: *Master Address File*

MAF/TIGER Enhancements Program: A Census Bureau program in anticipation of the 2010 census to make improvements to the Census Bureau's *Master Address File* and *TIGER* geographic database. Major objectives include the realignment of the TIGER database to be consistent with *GPS* readings, modernization of the TIGER database system, and implementation of the *Community Address Updating System* to collect address information as part of *American Community Survey* operations.

Mail Response Rate: Measure of respondent cooperation in the census, defined as the number of households returning a questionnaire by mail divided by the total number of questionnaires sent out in mailback areas. See also *mail return rate*.

Mail Return Rate: Measure of respondent cooperation in the census, defined as the number of households returning a questionnaire by mail divided by the total number of occupied households that were sent questionnaires in mailback areas (excluding vacant households and nonresidential units). The mail return rate is considered a more refined measure of cooperation than the earlier-available mail response rate.

Mailout/mailback: A method of census enumeration used primarily in urban areas in which questionnaires are mailed to each address and the residents are asked to mail back the completed questionnaires.

Master Address File (MAF): The master list of addresses of residential and nonresidential addresses maintained by the Census Bureau. An extract of the MAF, called the *Decennial Master Address File (DMAF)*, is the list of residential addresses on which the 2000 census enumeration was based. It is derived from the 1990 census address list (the *Address Control File*) in *mailout/mailback* areas or an address listing by census field staff, and is updated using a variety of sources, including information from the U.S. Postal Service, local officials, and a *block canvass* in mailout/mailback areas. See also *Topologically Integrated Geographic Encoding and Referencing (TIGER) System*.

Master Trace Sample: A sample of census records (possibly by selecting all records in a sample of decennial census blocks) for which all information relevant to census data collection is retained to assist in analyzing and comparing methodologies suggested for use in the subsequent census.

Matching: The process through which it is determined how many persons are included in both the *postenumeration survey* and the census (in PES *block clusters*) and how many persons are only included in one or the other attempted enumeration.

MCD: minor civil division. The Census Bureau has also used this abbreviation for the "mobile computing devices" it intends to use for follow-up data collection in 2010.

Monitoring Board: See *Census Monitoring Board*.

MSA: metropolitan statistical area

Multiple response modes: Generally speaking, alternative methods for being enumerated, not including *mailout/mailback* and enumeration as part of usual *nonresponse follow-up*. In 2000 these methods included obtaining and returning questionnaires available in public places (*"Be Counted"* forms), the use of the telephone and the Internet to obtain or provide census information, and the enumeration of persons at places that offer services to the homeless.

Multiunit structure: A building that contains more than one *housing unit*.

National Processing Center: The Census Bureau's permanent processing facility in Jeffersonville, Indiana.

Noncity-style address: See *city-style address*.

Nonmover: A person who lived in the same *housing unit* on both *Census Day* and the reference date for a *postenumeration survey*, as in the *Accuracy and Coverage Evaluation*.

Nonresponse: The failure to obtain all or part of the information requested on a questionnaire.

Nonresponse follow-up (NRFU): The field operation whereby census enumerators attempt to obtain completed questionnaires from interviewing members of households for which no questionnaire was returned in the mail. For the 2000 census, NRFU was performed between April and June, 2000. NRFU was conducted on a 100 percent basis in accordance with the Supreme Court's decision on sampling for apportionment; the Census Bureau's initial plans for the 2000 census called for sampling in this follow-up phase, which was then sampling for nonresponse follow-up (SNRFU).

Nonsampling error: Any *error* that occurs during the measuring or data collection process that is not directly related to the mechanics of *sampling* from the population. Nonsampling error includes *undercount* and *overcount*, resulting from people being missed or erroneously counted by enumeration processes.

NORC: National Opinion Research Center

NRC: National Research Council

NRFU: *nonresponse follow-up*

OCR: *optical character recognition*

OCS 2000: Operations Control System 2000

OMB: Office of Management and Budget (U.S.)

Omission: A person missed in the census. See also *Erroneous enumeration*; *Overcount*; *Undercount*.

OMR: *optical mark recognition*

One-number census: The goal of the original plan for the 2000 census involving *Integrated Coverage Measurement* and *sampling for nonresponse follow-up*; a census involving production of a single set of estimates. That set of estimates would be based on a combination of traditional counting methods and statistical estimation. The idea of a one-number census in 2000 was abandoned after enactment of legislation that required reporting of both statistically adjusted and unadjusted counts (if the decision to adjust census results was made) and after the U.S. Supreme Court ruled sampling unlawful for purposes of deriving *apportionment* counts.

Optical character recognition (OCR): Data capture technology that uses a scanner that "reads" and interprets human handwriting and converts it into electronic form.

Optical mark recognition (OMR): Data capture technology that scans a page, looking for marks in prespecified locations (e.g., circles keyed to the possible responses to a question; the respondent fills in one of those circles). Responses are coded based on the location of detected marks.

Outmover: A person who lived in a particular *housing unit* on *Census Day* but no longer lived there as of the reference date for a *postenumeration survey*, as in the *Accuracy and Coverage Evaluation*.

Overcount: The total number of people counted more than once or otherwise enumerated erroneously in the census. See also *Erroneous enumeration*; *Omission*; *Undercount*.

P-sample: Sample survey of respondents in a sample of *block clusters* conducted after (and independent of) the census enumeration; part of the *Accuracy and Coverage Evaluation*, it is used to calculate the match rate in the *dual-systems estimation* formula.

P.L. 94-171: The public law that requires the Census Bureau to provide the decennial census data required for congressional redistricting to the states by April of the year following the year of the census enumeration. The data file containing population counts by basic characteristics (age, race, Hispanic origin, and sex) down to the block level is commonly known as the "P.L. 94-171 data" or "redistricting data."

PALS: *Program for Address List Supplementation*

PAMS/ADAMS: *Pre-Appointment Management System/Automated Decennial Administrative Management System*

Person Duplication Studies: An evaluation of person duplication in the census conducted during summer 2001 (in preparation for the second decision on statistical adjustment of the 2000 census in October 2001). The Person Duplication Studies made use of the capability to match records by name and date of birth.

Person record: A record for an individual created from data captured from a census form.

PES: *Postenumeration survey; Post-Enumeration Survey (1990)*

PES-A: A variant form of a *postenumeration survey* that involves interviewers asking current residents in the survey sample about people who moved from those residences after *Census Day*.

PES-B: A variant form of a *postenumeration survey* used in 1990 that involves interviewers collecting information from survey respondents about their current residence as well as asking them where they lived (at the current *P-sample* address or another address) on *Census Day*. PES-B was ruled out as a strategy for the original 2000 census plan including *Integrated Coverage*

Measurement and, later, for the *Accuracy and Coverage Evaluation*; it was replaced with *PES-C*.

PES-C: A variant form of a *postenumeration survey* used in the 2000 *Accuracy and Coverage Evaluation* that involves interviewers (1) finding out who lived at each P-sample address on *Census Day*, including people who had left that address after Census Day, and (2) finding out who lived at each P-sample address as of the A.C.E. interview day. The information thus generated on *outmovers* (as well as *inmovers* and *nomovers*) made A.C.E. estimation more complex.

Postenumeration survey (PES): The independent follow-up survey conducted in some coverage evaluation programs using *dual-systems estimation*. The specific postenumeration survey conducted following the 1990 census is known as the *Post-Enumeration Survey* (*P-sample* component); in 2000, the postenumeration survey produced the P-sample of the *Accuracy and Coverage Evaluation*.

Post-Enumeration Survey (1990; PES): The 1990 census analogue of the *Accuracy and Coverage Evaluation* in 2000, including a *P-sample* and *E-sample*.

Poststratification: Separating a data set collected through use of sampling into strata on the basis of information gathered during data collection, and then treating each stratum separately in estimation.

Poststratum: A collection (of individuals in the census context) that share some characteristics (e.g., race, age, sex, region, owner/renter) obtained during data collection and that are separately treated in estimation.

Primary selection algorithm (PSA): Algorithm developed by the Census Bureau to consolidate multiple responses from the same address into a single return; given concerns about opening a loophole for duplicates, the details of the PSA have not been made public.

Program for Address List Supplementation (PALS): A short-lived Census Bureau program prior to the 2000 census that invited local governments to send address files to the Census Bureau for matching to and updating the MAF. PALS was later replaced by the *Local Update of Census Addresses* Program.

Proxy: A census interview in which the respondent is not a member of the household being enumerated. The respondent might be a neighbor, a landlord, or some other knowledgeable person.

PSA: *Primary selection algorithm*

Public use microdata sample: Computerized files containing a small sample of individual *long-form sample* records, subject to confidentiality protection, giving detailed population and housing characteristics for the people on those forms.

PUMS: *Public use microdata sample*

Race and Ethnicity Targeted Test (RAETT): A 1996 test conducted in selected areas of the United States to evaluate alternative formats and sequencing of the race, Hispanic, and ancestry questions for the 2000 census questionnaire.

RAETT: *Race and Ethnicity Targeted Test*

Reapportionment: See *apportionment*.

Redistricting: The revision of political boundaries—most commonly congressional districts, but also state legislative districts and municipal election wards and districts—based on the results of a new decennial census. Standards for redrawing the boundaries vary by location, but districts are generally intended to be as equal as possible in population.

Response rate: See *mail response rate*.

Return rate: See *mail return rate*.

Revision II: See *A.C.E. Revision II*.

Sample data; sample item: See *long-form sample*.

Sample data-defined: See *data-defined*

Sampling error: An error that occurs because only part of the population is contacted directly.

SBE: *Service-based enumeration*

Service-based enumeration (SBE): Enumeration of typically homeless people at food kitchens and shelters.

SF; SF1; SF2; SF3; SF4: *Summary File*

Short form: The census questionnaire that is mailed to about five-sixths of all *households* that asks only the basic *complete-count* data items.

Short-form data-defined: See *data-defined*.

SNRFU: Sampling for nonresponse follow-up; see *nonresponse follow-up*.

Special place: One or more *group quarters* where people live or stay that is different from the usual private house, apartment, or mobile home and that requires different decennial census procedures. Such places (e.g., a university or military installation) are administrative units; the individual group quarters (e.g., dormitories) are where people sleep.

Statistical Policy Directive 15: Common name for the 1977 version of standards issued by the U.S. Office of Management and Budget for the collection of data on race and ethnicity (Hispanic origin). The most recent revision (1997) of these guidelines is not labeled or numbered as a directive or circular. The guidelines make it possible to identify with more than one race, rather than choosing from only one of a few mutually exclusive racial categories.

Summary File (SF): The detailed data files of tabulations that are the primary data products of the census. Box 2.1 describes the content of each SF variant.

Synthetic error: Type of error cited by the *Executive Steering Committee for A.C.E. Policy* in its recommendation not to adjust census counts for congressional redistricting. The match rate, correct enumeration rate, and other rates involved in the *dual-systems estimation* for a population *poststratum* are supposed to apply at all lower geographic levels; synthetic error is produced in adjusted counts when this synthetic assumption is not satisfied.

Targeted extended search: An *Accuracy and Coverage Evaluation (A.C.E.)* operation that extended the search for matches between *E-* and *P-sample* records to the ring of *blocks* surrounding the A.C.E. sample *block cluster*.

TEA: *Type of enumeration area*

Tenure: The status of an occupied *housing unit* as either owner-occupied or renter-occupied.

TIGER: *Topologically Integrated Geographic Encoding and Referencing System*

Title 13: The portion of the U.S. Code under which the Census Bureau operates; it is also known as the *Census Act*. Originally enacted in 1929, Title 13 was enacted into positive law in 1954. In addition to outlining the authority to conduct the census, Title 13 also protects the confidentiality of census information and establishes penalties for disclosing this information.

Topologically Integrated Geographic Encoding and Referencing (TIGER) System: The framework for identifying the exact geographic location of residential addresses (as well as other physical features).

Tract: See *census tract*.

Type of enumeration area (TEA): A classification identifying how the Census Bureau conducted the decennial census in particular census blocks. There were nine TEAs in 2000. See *mailout/mailback*; *update/leave*; *urban update/leave*; *list/enumerate*.

Under(over)coverage; under(over)count: A nonspecific term representing either the rate or number of individuals missed (erroneously included) in the decennial census. More specifically, gross undercoverage and gross undercount are the rate or number of those missed for a demographic group or geographic area (similarly for gross overcoverage and gross overcount); net undercoverage and net undercount are the difference between the rate or number of those missed for a demographic group or geographic area and the rate or number of those erroneously included; differential (net) undercoverage and differential (net) undercount are the difference between the rate or number of net undercoverage between two demographic groups or between two geographic areas.

Unduplication: The process by which individuals reported on more than one census questionnaire are identified and counted once at only one geographic location.

Update/leave: (also known as **update/leave/mailback**). A method of census enumeration used in areas lacking city-style addresses in which the census questionnaire is delivered to an address by a census enumerator. The *Master Address File* is corrected at the time of delivery (if necessary). Residents at the address are asked to fill out the questionnaire and mail it back.

Urban update/leave: A new *type of enumeration area* used by the Census Bureau in the 2000 census. This data collection mode targeted urban areas in which mail delivery was thought to be problematic (e.g., large apartment complexes with common mailbox sites, at which a mail carrier might leave questionnaires in a common area rather than deliver to specific addresses). These areas were counted using *update/leave* methods rather than *mailout/mailback*.

USPS: U.S. Postal Service

Usual home elsewhere: A housing unit that is temporarily occupied by one or more people who have a usual residence elsewhere.

Usual residence: Under Census Bureau definition, the living quarters where a person spends more nights during a year than any other place.

Bibliography

Abramson, F. H. (2003, September 17). *Special Place/Group Quarters Enumeration*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 5. Washington, DC: U.S. Census Bureau.

Adams, T. S. and E. A. Krejsa (2001). *ESCAP II: Results of the Person Followup and Evaluation Followup Forms Review*. Executive Steering Committee for A.C.E. Policy II Supporting Report 24. Washington, DC: U.S. Census Bureau.

Adams, T. S. and X. Liu (2001). *ESCAP II: Evaluation of Lack of Balance and Geographic Errors Affecting Person Estimates*. Executive Steering Committee for A.C.E. Policy II Supporting Report 2. Washington, DC: U.S. Census Bureau.

Adlakha, A. L., J. G. Robinson, K. K. West, and A. Bruce (2003, August 18). *Assessment of Consistency of Census Data with Demographic Benchmarks at the Subnational Level*. Census 2000 Evaluation O.20. Washington, DC: U.S. Census Bureau.

Alberti, N. (2003, September 25). *Data Processing in Census 2000*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 7. Washington, DC: U.S. Census Bureau.

Anderson, M. J. (1988). *The American Census: A Social History*. New Haven, Conn.: Yale University Press.

Anderson, M. J. (2000). Advertising and the census. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 13–14. Washington, DC: CQ Press.

Anderson, M. J. and S. E. Fienberg (1999). *Who Counts? The Politics of Census-Taking in Contemporary America*. New York: Russell Sage Foundation.

Anderson, M. J. and S. E. Fienberg (2001). Counting and estimation: Methodology for improving the quality of censuses—the U.S. 2000 census adjustment decision. Technical report, Department of History, University of Wisconsin–Milwaukee, and Department of Statistics and Center for Automated Learning and Discovery, Carnegie Mellon University. Paper presented at the International Conference on Quality in Official Statistics, Stockholm.

Angueira, T. (2003). 2010 census planning, development, and testing: An update on the 2003 and 2004 census tests. Presentation to the Panel on Research on Future Census Methods, September 11, 2003, Washington, DC.

Baker, G. E. (1986). Whatever happened to the reapportionment revolution in the United States? In B. Grofman and A. Lijphart (Eds.), *Electoral Laws and Their Political Consequences*, pp. 257–276. New York: Agathon Press, Inc.

Bankier, M. (1999). Experience with the new imputation methodology used in the 1996 Canadian census with extensions for future censuses. Working Paper 24, Conference of European Statisticians, Session on Statistical Data Editing, Rome, Italy, June 2–4, 1999.

Bankier, M., P. Mason, and P. Poirier (2002). Imputation of demographic variables from the 2001 Canadian census of population. Working Paper 25, Conference of European Statisticians, Session on Statistical Data Editing, Helsinki, Finland, May 27–29, 2002.

Barrett, D. F., M. Beaghen, D. Smith, and J. Burcham (2001). *ESCAP II: Census 2000 Housing Unit Coverage Study*. Executive Steering Committee for A.C.E. Policy II Supporting Report 17. Washington, DC: U.S. Census Bureau.

Barrett, D. F., M. Beaghen, D. Smith, and J. Burcham (2003, February 21). *Census 2000 Housing Unit Coverage Study*. Census 2000 Evaluation O.3. Washington, DC: U.S. Census Bureau.

Bateman, D. V. (1991). *Post Enumeration Survey Evaluation Project P11: Balancing Error Evaluation*. 1990 Coverage Studies and Evaluation Memorandum Series M-2. Washington, DC: U.S. Census Bureau.

Bauder, M. and D. Judson (2003). *Administrative Records Experiment in 2000 (AREX 2000): Household Level Analysis*. Census 2000 Experiment Report. Washington, DC: U.S. Census Bureau.

Baumgardner, S. (2002, November 26). *Analysis of the Primary Selection Algorithm*. Census 2000 Evaluation L.3.a (Executive Summary Only). Washington, DC: U.S. Census Bureau.

Baumgardner, S. (2003, September 10). *Resolution of Multiple Census Returns Using a Re-interview*. Census 2000 Evaluation L.3.b (Executive Summary Only). Washington, DC: U.S. Census Bureau.

Baumgardner, S. K., D. A. Moul, R. A. Pennington, R. I. Piegari, H. F. Stackhouse, K. J. Zajac, N. S. Alberti, J. W. Reichert, and J. B. Treat (2001). *Quality of 2000 Census Processes*. DSSD Census 2000 Procedures and Operations Memorandum Series B-3. Washington, DC: U.S. Census Bureau.

Beaghen, M., R. Feldpausch, and R. Byrne (2001). *ESCAP II: Analysis of Non-Matches and Erroneous Enumerations Using Logistic Regression*. Executive Steering Committee for A.C.E. Policy II Supporting Report 19. Washington, DC: U.S. Census Bureau.

Beaghen, M. and R. Sands (2002, December 31). *Results from the Imputation of Unresolved Enumeration, Residency, and Match Status*. DSSD Revised A.C.E. Estimates Memorandum Series PP-57. Washington, DC: U.S. Census Bureau.

Bean, S. L. (2001). *ESCAP II: Accuracy and Coverage Evaluation Matching Error*. Executive Steering Committee for A.C.E. Policy II Supporting Report 7. Washington, DC: U.S. Census Bureau.

Bean, S. L. (2002, June 20). *Evaluation of Matching Error*. Census 2000 Evaluation N.14. Washington, DC: U.S. Census Bureau.

Bean, S. L. and D. M. Bauder (2002, December 31). *A.C.E. Revision II Report: Census and Administrative Records Duplication Study*. DSSD Revised A.C.E. Estimates Memorandum Series PP-44. Washington, DC: U.S. Census Bureau.

Bell, W. R. (1993). Using information from demographic analysis in post-enumeration survey estimation. *Journal of the American Statistical Association 88*(423), 1106–1118.

Bench, K. (2002, August 22). *Contamination of Census 2000 Data Collected in Accuracy and Coverage Evaluation Block Clusters*. Census 2000 Evaluation N.1. Washington, DC: U.S. Census Bureau.

Bentley, M., T. Mattingly, C. Hough, and C. Bennett (2003, April 3). *Census Quality Survey to Evaluate Responses to the Census 2000 Question on Race: An Introduction to the Data*. Census 2000 Evaluation B.3. Washington, DC: U.S. Census Bureau.

Berkowitz, S. (2001, July 17). *Puerto Rico Focus Groups on the Census 2000 Race and Ethnicity Questions*. Census 2000 Evaluation B.13. Washington, DC: U.S. Census Bureau.

Berkowitz, S. (2002, July 17). *Puerto Rico Focus Groups on Why Households Did Not Mail Back the Census 2000 Questionnaire*. Census 2000 Evaluation A.8. Washington, DC: U.S. Census Bureau.

Berning, M. A. (2003). *Administrative Records Experiment in 2000 (AREX 2000): Request for Physical Address Evaluation*. Census 2000 Experiment Report. Washington, DC: U.S. Census Bureau.

Berning, M. A. and R. H. Cook (2003). *Administrative Records Experiment in 2000 (AREX 2000): Process Evaluation*. Census 2000 Experiment Report. Washington, DC: U.S. Census Bureau.

Binder, D. A. (1996). Comment. *Journal of the American Statistical Association 91*(434), 510–512.

Brinson, A. and C. Fowler (2003, February 19). *Assessment Report for Data Capture of Paper Questionnaires*. Washington, DC: U.S. Census Bureau, U.S. Department of Commerce.

Brownrigg, L. A. (2003, October 16). *Ethnographic Social Network Tracing of Highly Mobile People*. Census 2000 Evaluation J.2. Washington, DC: U.S. Census Bureau.

Brudvig, L. (2003). *Analysis of the Social Security Number Validation Component of the Social Security Number, Privacy Attitudes, and Notification Experiment*. Census 2000 Testing, Experimentation, and Evaluation Program Report. Washington, DC: U.S. Census Bureau.

Bryant, B. E. (2000). Decennial census: 1990 census. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 158–161. Washington, DC: CQ Press.

Burcham, J. L. (2002, April 5). *Block Canvassing Operation*. Census 2000 Evaluation F.5. Washington, DC: U.S. Census Bureau.

Bureau of the Census (1964). *Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960: Accuracy of Data on Population Characteristics as Measured by CPS-Census Match*. ER 60–5. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1970). *Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960: Record Check Study of Accuracy of Income Reporting*. ER 60–8. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1975a). *1970 Census of Population and Housing Evaluation and Research Program: Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews*. PHC(E)-10. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1975b). *1970 Census of Population and Housing Evaluation and Research Program: Accuracy of Data for Selected Population Characteristics as Measured by the 1970 CPS-Census Match*. PHC(E)-11. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1976). *Language Minority, Illiteracy, and Voting Data Used in Making Determinations for the Voting Rights Act Amendments of 1975 (Public Law 94-73)*. Current Population Reports, Population Estimates and Projections. Series P-25, Number 627. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1982a). *The Meaning of Enumeration*. 1990 Planning Conference Series, Number 1. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1982b). *User's Guide—1980 Census of Population and Housing*. Part A, Text. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1983a). *Census of Population and Housing 1980: Public Use Microdata Samples Technical Documentation*. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1983b). *Introduction and Overview of the 1980 Census*, Chapter 1. Draft. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1984). 1990 census planning. Background paper for the Census Advisory Committee on Population Statistics, Washington, DC.

Bureau of the Census (1992). *Assessment of Accuracy of Adjusted Versus Unadjusted 1990 Census Base for Use in Intercensal Estimates: Recommendation. Report of the Committee on Adjustment of Postcensal Estimates*. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1993). *1990 Census of Population and Housing: History—Part A*. 1990 CPH-R-2A. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1994, February). *Nonresponse Followup Reinterview*. 1990 Census of Population and Housing Evaluation and Research Reports: Effectiveness of Quality Assurance. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1995a, October). *1990 Census of Population and Housing: History—Part B*. 1990 CPH-R-2B. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1995b, October). *1990 Census of Population and Housing: History—Part C*. 1990 CPH-R-2C. Washington, DC: U.S. Department of Commerce.

Bureau of the Census (1997). *Report to Congress—The Plan for Census 2000*. Originally issued July 1997; revised and reissued August 1997. Washington, DC: U.S. Department of Commerce.

Burt, G. and R. Mangaroo (2003, March 28). *Nonresponse Followup (NRFU) Enumerator Training*. Census 2000 Evaluation H.7. Washington, DC: U.S. Census Bureau.

Butcher, L. and N. Dunton (1999). *Uses of Small-Area Poverty Estimates, Final Report*. Kansas City, Missouri: Midwest Research Institute.

Byrne, R., M. Beaghen, and M. H. Mulry (2002, December 31). *Clerical Review of Census Duplicates (CRCD)*. DSSD Revised A.C.E. Estimates Memorandum Series PP-43. Washington, DC: U.S. Census Bureau.

Cantwell, P. J., D. McGrath, N. Nguyen, and M. F. Zelenak (2001). *Accuracy and Coverage Evaluation: Missing Data Results*. DSSD Census 2000 Procedures and Operations Memorandum Series B-7. Washington, DC: U.S. Census Bureau.

Carter, N. (2002, September 25). *Be Counted Campaign for Census 2000*. Census 2000 Evaluation A.3. Washington, DC: U.S. Census Bureau.

Carter, N. and S. Brady (2002, November 14). *Date of Reference for Age and Birth Date Used by Respondents of Census 2000*. Census 2000 Evaluation H.10. Washington, DC: U.S. Census Bureau.

Caspar, R. A. (2003). *Synthesis of Results from the Response Mode and Incentive Experiment*. Census 2000 Testing, Experimentation, and Evaluation Program Report. Washington, DC: U.S. Census Bureau.

Chesnut, J. (2003a, September 30). *Study of the U.S. Postal Service Reasons for Undeliverability of Census 2000 Mailout Questionnaires*. Census 2000 Evaluation A.6.b. Washington, DC: U.S. Census Bureau.

Chesnut, J. (2003b, March 20). *Telephone Questionnaire Assistance*. Census 2000 Evaluation A.1.a. Washington, DC: U.S. Census Bureau.

Childers, D. (1993). *The Impact of Housing Unit Coverage on Person Coverage*. Housing Unit Coverage Study Results Memorandum Number 2. Distributed with cover memo from Ruth Ann Killion to Thomas C. Walsh (June 24, 1993). Washington, DC: Bureau of the Census, U.S. Department of Commerce.

Childers, D. R. (2000, January 11). *Accuracy and Coverage Evaluation: The Design Document*. DSSD Census 2000 Procedures and Operations Memorandum Series S-DT-1. Washington, DC: U.S. Census Bureau.

Childers, D. R., R. L. Byrne, T. S. Adams, and R. Feldpausch (2001). *Accuracy and Coverage Evaluation: Person Matching and Follow-up Results*. DSSD Census 2000 Procedures and Operations Memorandum Series B-6. Washington, DC: U.S. Census Bureau.

Childers, D. R. and D. A. Fenstermaker (2000). *Accuracy and Coverage Evaluation: Overview of Design*. DSSD Census 2000 Procedures and Operations Memorandum Series S-DT-2. Washington, DC: U.S. Census Bureau.

Chiu, C. W. F., R. M. Yucel, E. Zanutto, and A. M. Zaslavsky (2001). Using matched substitutes to improve imputations for geographically linked databases. Paper prepared for the August 2001 Joint Statistical Meetings, Indianapolis, IN, and printed in *Proceedings of the Survey Research Methods Section of the American Statistical Association*, Alexandria, VA.

Choldin, H. (1994). *Looking for the Last Percent: The Controversy over Census Undercount*. New Brunswick, N.J.: Rutgers University Press.

Christenson, M. (2003, July 14). *Puerto Rico Census 2000 Responses to the Race and Ethnicity Questions*. Census 2000 Evaluation B.12. Washington, DC: U.S. Census Bureau.

Citro, C. F. (2000a). Advisory committees. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 14–18. Washington, DC: CQ Press.

Citro, C. F. (2000b). Editing and imputation. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 195–197. Washington, DC: CQ Press.

Citro, C. F. (2000c). Enumeration: Special populations. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 201–206. Washington, DC: CQ Press.

Citro, C. F. (2000d). Long form. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 273–277. Washington, DC: CQ Press.

Citro, C. F. (2000e). Population estimates and projections. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 300–303. Washington, DC: CQ Press. Assisted by Meyer Zitter.

Clark, J. R. and D. Moul (2003, September 29). *Coverage Improvement in Census 2000 Enumeration*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 10. Washington, DC: U.S. Census Bureau.

Clark, S. L., J. Iceland, T. Palumbo, K. Posey, and M. Weismantle (2003, September). *Comparing Employment, Income, and Poverty: Census 2000 and the Current Population Survey*. Housing and Household Economic Statistics Division Report. Washington, DC: U.S. Census Bureau.

Commerce Secretary's 2000 Census Advisory Committee (1995, March). Report to Ronald Brown, U.S. Secretary of Commerce.

Commerce Secretary's 2000 Census Advisory Committee (1999, January 22). Final Report to William Daley, U.S. Secretary of Commerce.

Conklin, J. (2003, March 12). *Evaluation of the Quality of the Data Capture System and the Impact of the Data Capture Mode on the Data Quality*. Census 2000 Evaluation K.1.b. Washington, DC: U.S. Census Bureau.

Cresce, A. (2003). *Overstatement of More Than One Race in Census 2000.* Memorandum dated June 10. Washington, DC: U.S. Census Bureau.

Crowley, M. (2003). *Generation X Speaks Out on Civic Engagement and the Decennial Census: An Ethnographic Approach*. Census 2000 Ethnographic Study. Washington, DC: U.S. Census Bureau.

Davis, M. and P. Biemer (1991a, July 11). *Estimates of P-sample Clerical Matching Error from a Rematching Evaluation*. 1990 Coverage Studies and Evaluation Memorandum Series H-2. Washington, DC: Bureau of the Census, U.S. Department of Commerce.

Davis, M. and P. Biemer (1991b, July 11). *Measurement of the Census Erroneous Enumerations—Clerical Error Made in the Assignment of Enumeration Status*. 1990 Coverage Studies and Evaluation Memorandum Series L-2. Washington, DC: Bureau of the Census, U.S. Department of Commerce.

Davis, P. P. (2001). *Accuracy and Coverage Evaluation: Dual System Estimation Results*. DSSD Census 2000 Procedures and Operations Memorandum Series B-9. Washington, DC: U.S. Census Bureau.

Dawson, J. A. and D. W. Stoudt (2003, July 15). *Automation of Census 2000 Processes*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 2. Washington, DC: U.S. Census Bureau.

de la Puente, M. (2004, February 5). *Census 2000 Ethnographic Studies*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 15. Washington, DC: U.S. Census Bureau.

de la Puente, M. and D. Stemper (2003, September 22). *The Enumeration of Colonias in Census 2000: Perspectives of Ethnographers and Census Enumerators*. Census 2000 Evaluation J.4. Washington, DC: U.S. Census Bureau.

del Pinal, J. (2003, September 26). *Race and Ethnicity in Census 2000*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 9. Washington, DC: U.S. Census Bureau.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B 39*, 1–38.

Dillman, D., J. R. Clark, and M. D. Sinclair (1993). The 1992 simplified questionnaire test: Effects of questionnaire length, respondent-friendly design and request for Social Security numbers on completion rates. In *Proceedings of the 1993 Bureau of the Census Annual Research Conference*, pp. 8–17. Washington, DC: U.S. Department of Commerce.

Dillman, D. A., J. R. Clark, and J. B. Treat (1994). Influence of 13 design factors on completion rates to decennial census questionnaires. Paper presented at the Annual Research Conference of the U.S. Bureau of the Census, Arlington, Va. (March).

Dimitri, C. R. (1999, June). *Mail Implementation Strategy*. Census 2000 Dress Rehearsal Evaluation Memorandum A1a. Washington, DC: U.S. Census Bureau.

Durbin, T. M. and L. P. Whitaker (1991). *Congressional and State Reapportionment and Redistricting: A Legal Analysis*. Congressional Research Service Report for Congress. 91-292-A. Washington, DC: U.S. Government Printing Office.

Edwards, W. S. and M. J. Wilson (2003, September 24). *Evaluations of the Census 2000 Partnership and Marketing Program*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 6. Washington, DC: U.S. Census Bureau.

Ehrenhalt, A. (1983). Reapportionment and redistricting. In T. E. Mann and N. J. Ornstein (Eds.), *The American Elections of 1982*, pp. 44–71. Washington, DC: American Enterprise Institute for Public Policy Research.

Eltinge, J. L. (1996). Comment. *Journal of the American Statistical Association 91*(434), 513–515.

Ericksen, E. P., L. F. Estrada, J. W. Tukey, and K. M. Wolter (1991). *Report on the 1990 Decennial Census and the Post-Enumeration Survey*. Washington, DC: U.S. Department of Commerce.

Executive Steering Committee for A.C.E. Policy (2001a). *Analysis Plan for Further ESCAP Deliberations Regarding the Adjustment of Census 2000 Data for Future Uses*. Washington, DC: U.S. Census Bureau.

Executive Steering Committee for A.C.E. Policy (2001b). *Report of the Executive Steering Committee for Accuracy and Coverage Evaluation Policy on Adjustment for Non-Redistricting Uses*. Washington, DC: U.S. Census Bureau.

Executive Steering Committee for A.C.E. Policy (2001c). *Report: Recommendation Concerning the Methodology to be Used in Producing Tabulations of Population Reported to States and Localities Pursuant to 13 U.S.C. 141(c)*. Washington, DC: U.S. Census Bureau.

Fairchild, L. (2001). *Long Form Direct Variance Estimation Specifications for Census 2000*. Working Draft, DSSD Census 2000 Procedures Memorandum Series LL (unnumbered; dated December 21). Washington, DC: U.S. Census Bureau.

Farber, J. (2001a). *Accuracy and Coverage Evaluation: Consistency of Post-Stratification Variables*. DSSD Census 2000 Procedures and Operations Memorandum Series B-10. Washington, DC: U.S. Census Bureau.

Farber, J. (2001b). *Quality Indicators of Census 2000 and the Accuracy and Coverage Evaluation*. DSSD Census 2000 Procedures and Operations Memorandum Series B-2. Washington, DC: U.S. Census Bureau.

Farley, R. (2001). *Identifying with Multiple Races: A Social Movement that Succeeded but Failed?* PSC Research Report 01-491. Ann Arbor: Population Studies Center, Institute for Social Research, University of Michigan.

Fay, R. E. (1992). When are inferences from multiple imputation valid? Paper prepared for the August 1992 Joint Statistical Meetings, Boston, MA, and printed in *Proceedings of the Survey Research Methods Section of the American Statistical Association*, Alexandria, VA.

Fay, R. E. (1996). Alternative paradigms for the analysis of imputed survey data. *Journal of the American Statistical Association 91*(434), 490–498.

Fay, R. E. (2001, October 26). *ESCAP II: Evidence of Additional Erroneous Enumerations from the Person Duplication Study*. Executive Steering Committee for A.C.E. Policy II Supporting Report 9. Washington, DC: U.S. Census Bureau.

Feindt, P. and R. Byrne (2000, September 21). *Accuracy and Coverage Evaluation Survey: Person Interviewing Results (Prototype)*. DSSD Census 2000 Procedures and Operations Memorandum Series B-5. Washington, DC: U.S. Census Bureau.

Feldpausch, R. (2002). *ESCAP II: E-Sample Erroneous Enumerations*. Executive Steering Committee for A.C.E. Policy II Supporting Report 5 (revised). Washington, DC: U.S. Census Bureau.

Fellegi, I. P. and D. Holt (1976). A systematic approach to automatic edit and imputation. *Journal of the American Statistical Association 71*(353), 17–35.

Fenstermaker, D. (2002, December 31). *A.C.E. Revision II: Summary of Estimated Net Coverage*. DSSD Revised A.C.E. Estimates Memorandum Series PP-54. Washington, DC: U.S. Census Bureau.

Fienberg, S. E. (2000). Capture-recapture methods. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 49–55. Washington, DC: CQ Press.

Gaines, L. M., L. Gage, and J. J. Salvo (2000). State and local governments: use of census data. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 337–340. Washington, DC: CQ Press.

Gerber, E. (2003). *Privacy Schemas and Data Collection: An Ethnographic Account*. Census 2000 Ethnographic Study. Washington, DC: U.S. Census Bureau.

Gerber, E., A. Dajani, and M. A. Scaggs (2002). *An Experiment to Improve Coverage Through Revised Roster Instructions*. Census 2000 Alternative Questionnaire Experiment Report. Washington, DC: U.S. Census Bureau.

Green, A., R. Watson, D. Smith, D. Barrett, R. Byrne, and S. Spratt (2003). *Evaluation of Housing Unit Field Operations and Instruments for the Accuracy and Coverage Evaluation*. Census 2000 Evaluation N.19. Washington, DC: U.S. Census Bureau.

Green, S. and C. Rothhaas (2002, December 10). *Evaluation of the Block Splitting Operation for Tabulation Purposes*. Census 2000 Evaluation F.16. Washington, DC: U.S. Census Bureau.

Griffin, D. H. and S. M. Obenski (2001, September 28). *A Demonstration of the Operational Feasibility of the American Community Survey*. Census 2000 Testing, Experimentation, and Evaluation Program Report. Washington, DC: U.S. Census Bureau.

Griffin, R. (2001). *Census 2000: Missing Housing Unit Status and Population Data*. DSSD Census 2000 Procedures and Operations Memorandum Series B-17. Washington, DC: U.S. Census Bureau.

Griffin, R. and D. Haines (2000). *Accuracy and Coverage Evaluation Survey: Post-Stratification for Dual System Estimation*. DSSD Census 2000 Procedures and Operations Memorandum Series Q-21. Washington, DC: U.S. Census Bureau.

Griffin, R. and D. Malec (2001). *ESCAP II: Sensitivity Analysis for the Assessment of the Synthetic Assumption*. Executive Steering Committee for A.C.E. Policy II Supporting Report 23. Washington, DC: U.S. Census Bureau.

Griffin, R. A. (2002, December 31). *Analysis of Synthetic Assumption*. DSSD Revised A.C.E. Estimates Memorandum Series PP-49. Washington, DC: U.S. Census Bureau.

Groves, R. M. and M. P. Couper (2002). Designing surveys acknowledging nonresponse. In M. Ver Ploeg, R. A. Moffitt, and C. F. Citro (Eds.), *Studies of Welfare Populations: Data Collection and Research Issues*, pp. 13–54. Panel on Data and Methods for Measuring the Effects of Changes in Social Welfare Programs, Committee on National Statistics. Washington, DC: National Academy Press.

Guarino, J. (2001). *Assessing the Impact of Differential Incentives and Alternative Data Collection Modes on Census Response*. Census 2000 Testing, Experimentation, and Evaluation Program Report. Washington, DC: U.S. Census Bureau.

Guarino, J. A., J. M. Hill, and H. F. Woltman (2001). *Analysis of the Social Security Number Notification Component of the Social Security Number, Privacy Attitudes, and Notification Experiment*. Census 2000 Testing, Experimentation, and Evaluation Program Report. Washington, DC: U.S. Census Bureau.

Haines, D. (1999a, September 22). *Accuracy and Coverage Evaluation Survey: Logistic Regression Modeling for Poststratification Variable Selection*. DSSD Census 2000 Procedures and Operations Memorandum Series Q-6. Washington, DC: U.S. Census Bureau.

Haines, D. (1999b, September 22). *Accuracy and Coverage Evaluation Survey: Poststratification Research Variables and Evaluation Statistics*. DSSD Census 2000 Procedures and Operations Memorandum Series Q-9. Washington, DC: U.S. Census Bureau.

Haines, D. (2000). *Accuracy and Coverage Evaluation Survey: Final Post-Stratification Plan for Dual-System Estimation*. DSSD Census 2000 Procedures and Operations Memorandum Series Q-24. Washington, DC: U.S. Census Bureau.

Haines, D. (2002, December 30). *Computer Specifications for Person Dual System Estimation Output Files*. DSSD Revised A.C.E. Estimates Memorandum Series PP-29. Washington, DC: U.S. Census Bureau.

Harris, D. (2003). Racial classification and the 2000 census. Commissioned paper, Panel to Review the 2000 Census, Committee on National Statistics. University of Michigan, Ann Arbor.

Hefter, S. P. (2000). *Long Form Weighting Specifications for Census 2000*. DSSD Census 2000 Procedures Memorandum Series LL (unnumbered; no date given, 2000 assumed). Washington, DC: U.S. Census Bureau.

Heimovitz, H. K. (2003). *Administrative Records Experiment in 2000 (AREX 2000): Outcomes Evaluation*. Census 2000 Experiment Report. Washington, DC: U.S. Census Bureau.

Hill, J. M. and J. D. Machowski (2003, September 29). *Master Trace Sample*. Census 2000 Evaluation B.6. Washington, DC: U.S. Census Bureau.

Hirschman, C., R. Alba, and R. Farley (2000). The meaning and measurement of race in the U.S. census: Glimpses into the future. *Demography 37*(2), 381–393.

Hogan, H. (1992). The 1990 Post-Enumeration Survey: An overview. *American Statistician 46*, 261–269.

Hogan, H. (1993). The 1990 Post-Enumeration Survey: Operations and results. *Journal of the American Statistical Association 88*(423), 1047–1060.

Hogan, H. (2000a). Accuracy and Coverage Evaluation. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 3–5. Washington, DC: CQ Press.

Hogan, H. (2000b). Accuracy and Coverage Evaluation: Theory and application. Paper prepared for the Second Workshop of the Panel to Review the 2000 Census, Committee on National Statistics. U.S. Census Bureau, Washington, DC.

Hogan, H. (2000c, September 27). *Data and Analysis to Inform the ESCAP Recommendation*. DSSD Census 2000 Procedures and Operations Memorandum Series B-1 (prototype). Washington, DC: U.S. Census Bureau.

Hogan, H. (2001a). *Accuracy and Coverage Evaluation: Data and Analysis to Inform the ESCAP Report*. DSSD Census 2000 Procedures and Operations Memorandum Series B-1. Washington, DC: U.S. Census Bureau.

Hogan, H. (2001b). *Accuracy and Coverage Evaluation Survey: Effect of Excluding "Late Census Adds"*. DSSD Census 2000 Procedures and Operations Memorandum Series Q-43. Washington, DC: U.S. Census Bureau.

Hough, C. L. and F. R. Borsa (2003, September 30). *Data Collection in Census 2000*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 13. Washington, DC: U.S. Census Bureau.

Hovland, I. and J. Buckley-Ess (2003, September 30). *Puerto Rico*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 14. Washington, DC: U.S. Census Bureau.

Hunter, J., M. de la Puente, and M. Solo (2003, September 11). *Comparative Ethnographic Research on Mobile Populations*. Census 2000 Evaluation J.3. Washington, DC: U.S. Census Bureau.

IBM Business Consulting Services (2003, October 8). *Management Evaluation of Census 2000*. Census 2000 Evaluation Q.1. Washington, DC: U.S. Census Bureau.

Imel, J. D. (2003, September 30). *Local Census Office Profile for Census 2000*. Census 2000 Evaluation H.9. Washington, DC: U.S. Census Bureau.

Johanson, C. (2003, September 9). *Effectiveness of Variables Used in the Model to Detect Discrepant Results During Reinterview and the Identification of New Variables*. Census 2000 Evaluation M.2. Washington, DC: U.S. Census Bureau.

Jonas, K. (2002, November 7). *Group Quarters Enumeration*. Census 2000 Evaluation E.5. Washington, DC: U.S. Census Bureau.

Jonas, K. (2003a, July 31). *Census Unedited File Creation*. Census 2000 Evaluation L.4. Washington, DC: U.S. Census Bureau.

Jonas, K. (2003b, August 6). *Group Quarters Enumeration*. Census 2000 Evaluation E.5, Revision 1. Washington, DC: U.S. Census Bureau.

Jones, J. (2003a, April 8). *Housing Unit Duplication in Census 2000*. Census 2000 Evaluation O.10. Washington, DC: U.S. Census Bureau.

Jones, J. (2003b, January 21). *Person Duplication in the Search Area Measured by the 2000 Accuracy and Coverage Evaluation*. Census 2000 Evaluation O.16. Washington, DC: U.S. Census Bureau.

Jones, J. and D. F. Barrett (2003, June 25). *Questionnaire Assistance Centers for Census 2000*. Census 2000 Evaluation H.4. Washington, DC: U.S. Census Bureau.

Judkins, D. R. (1996). Comment. *Journal of the American Statistical Association 91*(434), 507–510.

Judson, D. H. and B. Bye (2003, October 21). *Synthesis of Results from the Administrative Records Experiment in 2000 (AREX 2000)*. Washington, DC: U.S. Census Bureau.

Kearney, A. T. (2002, December 31). *A.C.E. Revision II: Missing Data Evaluation Final Report*. DSSD Revised A.C.E. Estimates Memorandum Series PP-48. Washington, DC: U.S. Census Bureau.

Keathley, D. (2002, December 31). *Report on the Error Due to Estimating Outmovers Using Inmovers in the PES-C*. DSSD Revised A.C.E. Estimates Memorandum Series PP-47. Washington, DC: U.S. Census Bureau.

Keathley, D. (2003, September 29). *Evaluation of Nonresponse Followup—Mover Probe*. Census 2000 Evaluation I.3. Washington, DC: U.S. Census Bureau.

Keathley, D., A. Kearney, and W. Bell (2001). *ESCAP II: Analysis of Missing Data Alternatives for the Accuracy and Coverage Evaluation*. Executive Steering Committee for A.C.E. Policy II Supporting Report 12. Washington, DC: U.S. Census Bureau.

Keeley, C. (1993). *Could the Census Bureau Reduce the Undercount by Not Using a "Long Form?"*. Washington, DC: Bureau of the Census, U.S. Department of Commerce.

Kilmer, A. (2002, November 5). *Census 2000 Count Imputation—Results*. DSSD Census 2000 Procedures and Operations Memorandum Series Q-81. Washington, DC: U.S. Census Bureau.

Kim, J.-k. and W. A. Fuller (2002). Variance estimation for nearest neighbor imputation. Unpublished manuscript.

Kincannon, C. L. (2003, March 12). *Statement of Census Bureau director C. Louis Kincannon on Accuracy and Coverage Evaluation Revision II*. Washington, DC: Public Information Office, U.S. Census Bureau.

Kohn, F. (2003, April 10). *The United States Postal Service Undeliverable Rates for Census 2000 Mailout Questionnaires*. Census 2000 Evaluation A.6.a. Washington, DC: U.S. Census Bureau.

Kominski, R. (1985). *Final Report—Documentation of Voting Rights Act Determinations*. Washington, DC: Bureau of the Census, U.S. Department of Commerce.

Kominski, R. (1992). *1992 Voting Rights Act Bilingual Ballots Determinations*. Washington, DC: Bureau of the Census, U.S. Department of Commerce.

Kostanich, D. (Ed.) (2003a, March 11). *A.C.E. Revision II: Design and Methodology*. DSSD Revised A.C.E. Estimates Memorandum Series PP-30. Washington, DC: U.S. Census Bureau.

Kostanich, D. (2003b, January 2). *A.C.E. Revision II: Summary of Methodology*. DSSD Revised A.C.E. Estimates Memorandum Series PP-35. Washington, DC: U.S. Census Bureau.

Kostanich, D., R. Griffin, and D. Fenstermaker (1999). Accuracy and Coverage Evaluation survey: Plans for Census 2000. Unpublished paper prepared for the March 19, 1999, meeting of the Panel to Review the 2000 Census, Committee on National Statistics. U.S. Bureau of the Census, Department of Commerce, Washington, D.C.

Krejsa, E. A. (2003, March 25). *Discrepant Results in the Accuracy and Coverage Evaluation*. Census 2000 Evaluation N.10. Washington, DC: U.S. Census Bureau.

Krejsa, E. A. and T. S. Adams (2002, December 31). *Results of the A.C.E. Revision II Measurement Coding*. DSSD Revised A.C.E. Estimates Memorandum Series PP-55. Washington, DC: U.S. Census Bureau.

Krejsa, E. A. and D. A. Raglin (2001). *ESCAP II: Evaluation Results for Changes in A.C.E. Enumeration Status*. Executive Steering Committee for A.C.E. Policy II Supporting Report 3. Washington, DC: U.S. Census Bureau.

Kulka, R. A., N. A. Holt, W. Carter, and K. L. Dowd (1991). Self-reports of time pressures, concerns for privacy, and participation in the 1990 mail census. In *Proceedings of the Bureau of the Census 1991 Annual Research Conference*, pp. 33–54. Washington, DC: U.S. Department of Commerce.

Laney, G. P. (1992). *The Voting Rights Act of 1965, As Amended: Its History and Current Issues*. Congressional Research Service Report for Congress. 92-578-GOV. Washington, DC: U.S. Government Printing Office.

Lestina, Frederic Allen, J. (2003, September 30). *Analysis of the Linguistically Isolated Population in Census 2000*. Census 2000 Evaluation A.5.a. Washington, DC: U.S. Census Bureau.

Little, R. J. A. and D. B. Rubin (1987). *Statistical Analysis With Missing Data*. New York: John Wiley & Sons.

Liu, X. J., R. L. Byrne, and L. M. Imel (2001). *ESCAP II: Analysis of Movers*. Executive Steering Committee for A.C.E. Policy II Supporting Report 15. Washington, DC: U.S. Census Bureau.

Liu, X. J., R. Feldpausch, and D. Smith (2002, October 15). *Analysis of Conflicting Households*. Census 2000 Evaluation O.4. Washington, DC: U.S. Census Bureau.

Love, S. and D. Dalzell (2001, March 27). Comparison of record substitution and 100% data allocation. Memorandum for the Record. U.S. Census Bureau, Washington, DC.

Macro International (2002, July 10). *Evaluation of the Census in Schools Program: Materials and Distribution*. Census 2000 Evaluation D.2. Washington, DC: U.S. Census Bureau.

Magnuson, D. L. (2000a). Decennial censuses: 1910 census. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 135–136. Washington, DC: CQ Press.

Magnuson, D. L. (2000b). Decennial censuses: 1930 census. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 139–140. Washington, DC: CQ Press.

Marks, E. S. (1978). The role of dual system estimation in census, evaluation. In K. J. Krotki (Ed.), *Developments in Dual Systems Estimation of Population Size and Growth*, Chapter 20. Edmonton: University of Alberta Press.

Martin, E. (1999). Who knows who lives here? Within-household disagreements as a source of survey coverage error. *Public Opinion Quarterly 63*, 220–236.

Martin, E. (2001). Privacy concerns and the census long form: Some evidence from census 2000. Paper prepared for the Joint Statistical Meetings, Atlanta, GA. U.S. Census Bureau, Washington, DC.

Martin, E. (2002). *Questionnaire Effects on Reporting of Race and Hispanic Origin: Results of a Replication of the 1990 Mail Short Form in Census 2000*. Census 2000 Alternative Questionnaire Experiment Report. Washington, DC: U.S. Census Bureau.

Martin, E., E. Gerber, and C. Redline (2003, August 28). *Synthesis Report: Census 2000 Alternative Questionnaire Experiment*. Census 2000 Alternative Questionnaire Experiment Report. Washington, DC: U.S. Census Bureau.

Martin, E., D. Sheppard, M. Bentley, and C. Bennett (2003). *Results of 2003 National Census Test of Race and Hispanic Questions*. Unpublished report, dated October 1. Washington, DC: U.S. Census Bureau.

McMillen, D. (2000). Americans overseas. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 31–33. Washington, DC: CQ Press.

McMillen, D. B. (1998). The census under attack. *CHANCE 11*(2), 50–55.

McNally, T. (2002, November 6). *Service-Based Enumeration*. Census 2000 Evaluation E.6. Washington, DC: U.S. Census Bureau.

McNally, T. (2003, May 15). *Operational Analysis of Enumeration of Puerto Rico*. Census 2000 Evaluation H.8. Washington, DC: U.S. Census Bureau.

Meng, X.-L. (1994). Multiple-imputation inferences with uncongenial sources of input. *Statistical Science 9*(4), 538–573.

Miskura, S. M. (1993). *Definition, Clarification, and Issues: "One-Number Census"*. Memorandum dated April 14 for R.D. Tortora. Washington, DC: Bureau of the Census, U.S. Department of Commerce.

Miskura, S. M. (2000a). *Addresses in the Master Address File (MAF) Excluded from the April 7, 2000 MAF Extract*. Census 2000 Informational Memorandum 67. Washington, DC: U.S. Census Bureau.

Miskura, S. M. (2000b). *Results of Reinstatement Rules for the Housing Unit Duplication Operations*. Census 2000 Informational Memorandum 82. Washington, DC: U.S. Census Bureau.

Morganstein, D., D. Marker, and D. Levine (2003, September 25). *Evaluation of the Census 2000 Quality Assurance Philosophy and Approach Used in the Address List Development and Enumeration Operations*. Census 2000 Evaluation M.1. Washington, DC: U.S. Census Bureau.

Moul, D. A. (2002, July 25). *Nonresponse Followup for Census 2000*. Census 2000 Evaluation H.5. Washington, DC: U.S. Census Bureau.

Moul, D. A. (2003, May 9). *Coverage Improvement Followup*. Census 2000 Evaluation I.4. Washington, DC: U.S. Census Bureau.

Mule, T. (2001). *ESCAP II: Person Duplication in Census 2000*. Executive Steering Committee for A.C.E. Policy II Supporting Report 20. Washington, DC: U.S. Census Bureau.

Mule, T. (2002a, December 31). *A.C.E. Revision II Results: Further Study of Person Duplication*. DSSD Revised A.C.E. Estimates Memorandum Series PP-51. Washington, DC: U.S. Census Bureau.

Mule, T. (2002b, April 4). *Revised Preliminary Estimates of Net Undercounts for Seven Race/Ethnicity Groupings*. DSSD Revised A.C.E. Estimates Memorandum Series PP-2. Washington, DC: U.S. Census Bureau.

Mule, T. (2003, March 4). *A.C.E. Revision II Results: Change in Estimated Net Undercount*. DSSD A.C.E. Revision II Memorandum Series PP-58. Washington, DC: U.S. Census Bureau.

Mulry, M. H. and R. S. ZuWallack (2002, December 31). *Confidence Intervals and Loss Function Analyses*. DSSD Revised A.C.E. Estimates Memorandum Series PP-42. Washington, DC: U.S. Census Bureau.

Nash, F. F. (2000). *Overview of the Duplicate Housing Unit Operations*. Washington, DC: U.S. Census Bureau.

National Conference of State Legislatures (1992). Customer feedback for the 1990 census redistricting data program—a preliminary survey. Prepared for discussion at the National Conference of State Legislatures Annual Meeting, Cincinnati, Ohio.

National Research Council (1978). *Counting the People in 1980: An Appraisal of Census Plans*. Panel on Decennial Census Plans, Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1985). *The Bicentennial Census: New Directions for Methodology in 1990*. Panel on Decennial Census Methodology, Constance F. Citro and Michael L. Cohen, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1987). Letter from Benjamin F. King to John Keane, director, U.S. Bureau of the Census. Panel on Decennial Census Methodology, Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1988). *Priorities for the 1990 Census: Research, Evaluation, and Experimentation (REX) Program*. Panel on Decennial Census Methodology, Constance F. Citro and Michael L. Cohen, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1994). *Counting People in the Information Age*. Panel to Evaluate Alternative Census Methods, Duane L. Steffey and Norman M. Bradburn, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1995a). *Measuring Poverty: A New Approach*. Panel on Poverty and Family Assistance: Concepts, Information Needs, and Measurement Methods, Constance F. Citro and Robert T. Michael, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1995b). *Modernizing the U.S. Census*. Panel on Census Requirements in the Year 2000 and Beyond, Barry Edmonston and Charles Schultze, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1996). *Spotlight on Heterogeneity—the Federal Standards for Racial and Ethnic Classification: Summary of a Workshop*. Barry Edmonston, Joshua Goldstein, and Juanita Tamayo Lott, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1997). *Preparing for the 2000 Census: Interim Report II*. Panel to Evaluate Alternative Census Methodologies, Andrew A. White and Keith F. Rust, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1999a, May). Letter from Janet L. Norwood to Kenneth Prewitt, director, U.S. Census Bureau. Panel to Review the 2000 Census, Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (1999b). *Measuring a Changing Nation: Modern Methods for the 2000 Census*. Panel on Alternative Census Methodologies, Michael L. Cohen, Andrew A. White, and Keith F. Rust, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2000a). *Designing the 2010 Census: First Interim Report*. Panel on Research on Future Census Methods, Michael L. Cohen and Benjamin F. King, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2000b, November). Letter from Janet L. Norwood to Kenneth Prewitt, director, U.S. Census Bureau. Panel to Review the 2000 Census, Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2001a). *The 2000 Census: Interim Assessment*. Panel to Review the 2000 Census, Constance F. Citro, Daniel L. Cork, and Janet L. Norwood, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2001b). *Choosing the Right Formula: Interim Report*. Panel on Formula Allocations, Thomas B. Jabine, Thomas A. Louis, and Allen L. Schirm, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2001c). Letter from Benjamin F. King to William G. Barron, acting director, U.S. Census Bureau. Dated February 15. Panel on Research on Future Census Methods, Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2001d, November). Letter from Janet L. Norwood to William Barron, acting director, U.S. Census Bureau. Panel to Review the 2000 Census, Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2001e). *Proceedings, First Workshop of Panel to Review the 2000 Census (October 6, 1999)*. Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2001f). *Proceedings, Second Workshop of Panel to Review the 2000 Census (February 2-3, 2000)*. Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2001g). *Proceedings, Third Workshop of Panel to Review the 2000 Census (October 2, 2000)*. Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2003a). *Planning the 2010 Census: Second Interim Report*. Panel on Research on Future Census Methods, Daniel L. Cork, Michael L. Cohen, and Benjamin F. King, eds., Committee on National Statistics. Washington, DC: The National Academies Press.

National Research Council (2003b). *Statistical Issues in Allocating Funds by Formula*. Panel on Formula Allocations, Thomas A. Louis, Thomas B. Jabine, and Marisa A. Gerstein, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2004). *Reengineering the 2010 Census: Risks and Challenges*. Panel on Research on Future Census Methods, Daniel L. Cork, Michael L. Cohen, and Benjamin F. King, eds., Committee on National Statistics. Washington, DC: The National Academies Press.

Navarro, A. and D. Olson (2001). *Accuracy and Coverage Evaluation: Effect of Targeted Extended Search*. DSSD Census 2000 Procedures and Operations Memorandum Series B-18. Washington, DC: U.S. Census Bureau.

Naymark, J. and K. Hodges (2000). Uses of census data by the private sector. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 356–358. Washington, DC: CQ Press.

Nguyen, N. T. and M. F. E. Zelenak (2003, June 24). *Coverage Gains From Coverage Questions C1 and C2 on Enumerator-Completed Questionnaires for Census 2000*. Census 2000 Evaluation I.5. Washington, DC: U.S. Census Bureau.

Norris, S. (2003, September 23). *Analysis of Item Nonresponse Rates for the 100 Percent Housing and Population Items from Census 2000*. Census 2000 Evaluation B.1.b. Washington, DC: U.S. Census Bureau.

O'Rourke, T. G. (1980). *The Impact of Reapportionment*. New Brunswick, NJ: Transaction Books.

Owens, K. L. (2000). *Census 2000 Address List Review*. Washington, DC: U.S. Census Bureau.

Owens, K. L. (2002, May 14). *Evaluation of the Local Update of Census Addresses 99 (LUCA 99)*. Census 2000 Evaluation F.6. Washington, DC: U.S. Census Bureau.

Owens, K. L. (2003, April 16). *Evaluation of the Local Update of Census Addresses 98 (LUCA 98)*. Census 2000 Evaluation F.3. Washington, DC: U.S. Census Bureau.

Parker, F. R. (1989). Changing standards in voting rights law. In W. P. O'Hare (Ed.), *Redistricting in the 1990s: A Guide for Minority Groups*, pp. 55–66. Washington, DC: Population Reference Bureau, Inc.

Passel, J. S. (2000). Immigration. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 253–260. Washington, DC: CQ Press.

Pennington, R. A. (2003, June 6). *Evaluation of the Update/Leave Operation*. Census 2000 Evaluation F.10. Washington, DC: U.S. Census Bureau.

Petroni, R. and D. R. Childers (2003, August 28). *Coverage Measurement from the Perspective of March 2001 A.C.E.* Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 4. Washington, DC: U.S. Census Bureau.

Philipp, D. (2001, July 16). *2000 Decennial Census Documentation: Sample Census Edited File (SCEF)*. Washington, DC: Decennial Systems and Contracts Management, U.S. Census Bureau.

Prewitt, K. (2000, June). *Accuracy and Coverage Evaluation: Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of Census 2000*. Washington, DC: U.S. Census Bureau.

Raghunathan, T. E., J. M. Lepkowski, J. Van Hoewyk, and P. Solenberger (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology 27*(1), 85–95.

Raglin, D. A. (2001). *ESCAP II: Effect of Excluding Reinstated Census People from the A.C.E. Person Process*. Executive Steering Committee for A.C.E. Policy II Supporting Report 13. Washington, DC: U.S. Census Bureau.

Raglin, D. A. and E. A. Krejsa (2001). *ESCAP II: Evaluation for Results for Changes in Mover and Residence Status in the A.C.E.* Executive Steering Committee for A.C.E. Policy II Supporting Report 16. Washington, DC: U.S. Census Bureau.

Rao, J. N. K. and J. Shao (1992). Jackknife variance estimation with survey data under hot deck imputation. *Biometrika 79*(4), 811–822.

Redline, C., D. Dillman, A. Dajani, and M. A. Scaggs (2002). *The Effects of Altering the Design of Branching Instructions on Navigational Performance in Census 2000*. Census 2000 Alternative Questionnaire Experiment Report. Washington, DC: U.S. Census Bureau.

Robinson, J. G. (2000, September 21). *Accuracy and Coverage Evaluation Survey: Demographic Analysis Results (Prototype)*. DSSD Census 2000 Procedures and Operations Memorandum Series B-4. Washington, DC: U.S. Census Bureau.

Robinson, J. G. (2001a). *Accuracy and Coverage Evaluation: Demographic Analysis Results*. DSSD Census 2000 Procedures and Operations Memorandum Series B-4. Washington, DC: U.S. Census Bureau.

Robinson, J. G. (2001b). *ESCAP II: Demographic Analysis Results*. Executive Steering Committee for A.C.E. Policy II Supporting Report 1. Washington, DC: U.S. Census Bureau.

Robinson, J. G. and A. Adlakha (2002, December 31). *Comparison of A.C.E. Revision II Results with Demographic Analysis*. DSSD Revised A.C.E. Estimates Memorandum Series PP-41. Washington, DC: U.S. Census Bureau.

Robinson, J. G., B. Ahmed, P. Das Gupta, and K. A. Woodrow (1993). Estimation of population coverage in the 1990 United States census based on demographic analysis. *Journal of the American Statistical Association 88*(423), 1061–1071.

Robinson, J. G. and G. S. Wolfgang (2002, December 31). *Comparison of A.C.E. Revision II Population Coverage Results with HUCS Housing Unit Coverage Results*. DSSD Revised A.C.E. Estimates Memorandum Series PP-50. Washington, DC: U.S. Census Bureau.

Rochester Institute of Technology Research Corporation (2002, September 20). *DCS 2000 Data Quality, v.2.1 Final*. Rochester, NY: Rochester Institute of Technology Research Corporation. Prepared for the U.S. Census Bureau, Washington, DC.

Rosenthal, M. (2002a, December 10). *Update/Enumerate*. Census 2000 Evaluation F.12. Washington, DC: U.S. Census Bureau.

Rosenthal, M. (2002b, October 3). *Urban Update/Leave*. Census 2000 Evaluation F.11. Washington, DC: U.S. Census Bureau.

Rosenthal, M. (2003a, October 24). *DCS 2000 Data Capture Audit Resolution Process*. Census 2000 Evaluation K.1.a. Washington, DC: U.S. Census Bureau.

Rosenthal, M. (2003b, June 30). *Operational Analysis of the Decennial Response File Linking and Setting of Housing Unit Status and Expected Household Size*. Census 2000 Evaluation L.2. Washington, DC: U.S. Census Bureau.

Rubin, D. B. (1996). Multiple imputation after 18+ years (with discussion). *Journal of the American Statistical Association 91*(434), 473–489.

Ruggles, S. (2000). IPUMS. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 264–267. Washington, DC: CQ Press.

Ruhnke, M. C. (2002, January 30). *The Address Listing Operation and Its Impact on the Master Address File*. Census 2000 Evaluation F.2. Washington, DC: U.S. Census Bureau.

Ruhnke, M. C. (2003, August 19). *An Assessment of Addresses on the Master Address File "Missing" in the Census or Geocoded to the Wrong Collection Block*. Census 2000 Evaluation F.15. Washington, DC: U.S. Census Bureau.

Salvo, J. J. (2000). Data capture. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 105–108. Washington, DC: CQ Press.

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. New York: Chapman and Hall.

Schindler, E. (2000, January 12). *Accuracy and Coverage Evaluation Survey: Post-stratification Preliminary Research Results*. DSSD Census 2000 Procedures and Operations Memorandum Series Q-23. Washington, DC: U.S. Census Bureau.

Schindler, E. (2001). *Accuracy and Coverage Evaluation (A.C.E.) Survey—Census Imputations by Post-stratum*. DSSD Census 2000 Procedures and Operations Memorandum Series Q-64. Washington, DC: U.S. Census Bureau.

Schindler, E. (2003, April 9). *Adjusted Data for States, Counties, and Places*. DSSD Revised A.C.E. Estimates Memorandum Series PP-60. Washington, DC: U.S. Census Bureau.

Schneider, P. (2003, September 29). *Content and Data Quality in Census 2000*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 12. Washington, DC: U.S. Census Bureau.

Schneider, P. (2004, January 22). *Content and Data Quality in Census 2000*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 12 (revised). Washington, DC: U.S. Census Bureau.

Schneider, S., D. Cantor, P. Segel, C. Arieira, and L. Nguyen (2002). *Response Mode and Incentive Experiment for Census 2000*. Census 2000 Testing, Experimentation, and Evaluation Program Report. Washington, DC: U.S. Census Bureau.

Schwede, L. (2003, August 27). *Complex Households and Relationships in the Decennial Census and in Ethnographic Studies of Six Race/Ethnic Groups*. Census 2000 Testing, Experimentation, and Evaluation Program Report. Washington, DC: U.S. Census Bureau.

Sheppard, D. (2003, July 29). *Coverage Edit Followup*. Census 2000 Evaluation I.1. Washington, DC: U.S. Census Bureau.

Siegel, P. M. (1993). *The Impact of Content on Census Coverage*. Washington, DC: Bureau of the Census, U.S. Department of Commerce.

Singer, E. (2003, June 20). *Privacy Research in Census 2000*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 1. Washington, DC: U.S. Census Bureau.

Singer, E., J. Van Hoewyk, R. Tourangeau, D. Steiger, M. Montgomery, and R. Montgomery (2001). *Final Report on the 1999–2000 Surveys of Privacy Attitudes*. Census 2000 Testing, Experimentation, and Evaluation Program Report. Washington, DC: U.S. Census Bureau.

Singer, P. and S. R. Ennis (2002, December 13). *Census 2000 Content Reinterview Survey: Accuracy of Data for Selected Population and Housing Characteristics as Measured by Reinterview*. Census 2000 Evaluation B.5. Washington, DC: U.S. Census Bureau.

Singer, P. and S. R. Ennis (2003, September 24). *Census 2000 Content Reinterview Survey: Accuracy of Data for Selected Population and Housing Characteristics as Measured by Reinterview*. Census 2000 Evaluation B.5 (revised and final). Washington, DC: U.S. Census Bureau.

Smith, D., D. F. Barrett, and M. Beaghen (2003, October 17). *Analysis of Deleted and Added Housing Units in Census 2000 Measured by the Accuracy and Coverage Evaluation*. Census 2000 Evaluation O.19. Washington, DC: U.S. Census Bureau.

Smith, D. and D. Whitford (2003). Census unduplication research plan for 2010. Paper prepared for the August 2003 Joint Statistical Meetings, San Francisco, CA. U.S. Census Bureau, Washington, DC.

Smith, D. R. and J. Jones (2003, September 22). *Use of Non-English Questionnaires and Guides in the Census 2000 Language Program*. Census 2000 Evaluation A.4. Washington, DC: U.S. Census Bureau.

Spar, E. J. (2000). Private sector. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 309–311. Washington, DC: CQ Press.

Stackhouse, H. F. and S. Brady (2003a, January 30). *Census 2000 Mail Response Rates*. Census 2000 Evaluation A.7.a. Washington, DC: U.S. Census Bureau.

Stackhouse, H. F. and S. Brady (2003b, January 30). *Census 2000 Mail Return Rates*. Census 2000 Evaluation A.7.b. Washington, DC: U.S. Census Bureau.

Stapleton, C. and J. Irwin (2002, April 15). *Census 2000 Internet Web Site and Questionnaire Customer Satisfaction Surveys*. Census 2000 Evaluation A.2.c. Washington, DC: U.S. Census Bureau.

Starsinic, M. D., C. D. Stissel, and M. E. Asiala (2001). *Accuracy and Coverage Evaluation: Variance Estimates by Size of Geographic Area*. DSSD Census 2000 Procedures and Operations Memorandum Series B-11. Washington, DC: U.S. Census Bureau.

Steel, P. and L. Zayatz (2003, April 15). *The Effects of the Disclosure Limitation Procedure on Census 2000 Tabular Data Products*. Census 2000 Evaluation C.1. (Abridged). Washington, DC: U.S. Census Bureau.

Stevens, N. (2002, January 30). *Telephone Questionnaire Assistance Customer Satisfaction Survey*. Census 2000 Evaluation A.1.b. Washington, DC: U.S. Census Bureau.

Stevens, N. (2003, May 27). *Evaluation of the Facility Questionnaire (Computer Assisted Telephone Interviewing and Personal Visit)*. Census 2000 Evaluation E.1.b. Washington, DC: U.S. Census Bureau.

Stiers, G. (2000, September 26). *Demographic Full Count Report*. DSSD Census 2000 Procedures and Operations Memorandum Series B-16. Washington, DC: U.S. Census Bureau.

Stiller, J. and D. R. Dalzell (2003). Hot-deck imputation with SAS™ arrays and macros for large surveys. Paper prepared for the Joint Statistical Meetings, San Francisco, CA. U.S. Census Bureau, Washington, DC.

Tenebaum, M. (2001, July 24). *Assessment of Field Verification*. Census 2000 Evaluation H.2. Washington, DC: U.S. Census Bureau.

Thernstrom, A. M. (1979). The odd evolution of the Voting Rights Act. *The Public Interest 55*(Spring), 49–76.

Thibaudeau, Y. (1998). *Model Explicit Item Imputation for Demographic Categories for Census 2000*. Statistical Research Division Research Report RR-99-02. Washington, DC: U.S. Census Bureau.

Thompson, J. H. (1992). *CAPE Processing Results*. Washington, DC: Bureau of the Census, U.S. Department of Commerce.

Thompson, J. H. (2000). Organization and administration of the census. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 295–299. Washington, DC: CQ Press.

Thompson, J. H., P. J. Waite, and R. E. Fay (2001). *Basis of "Revised Early Approximation" of Undercounts Released Oct. 17, 2001*. Executive Steering Committee for A.C.E. Policy II Supporting Report 9a. Washington, DC: U.S. Census Bureau.

Titan Corporation (2003, July 22). *Census 2000 Data Capture*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 3. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2001a, December 28). *Coverage Edit Followup System Requirements Study*. Census 2000 Evaluation R.1.b. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2001b, December 28). *Internet Data Collection System Requirements Study*. Census 2000 Evaluation R.1.d. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2001c, December 28). *Internet Questionnaire Assistance System Requirements Study*. Census 2000 Evaluation R.1.c. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2001d, December 28). *Telephone Questionnaire Assistance System Requirements Study*. Census 2000 Evaluation R.1.a. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2002a, May 10). *Accuracy and Coverage Evaluation 2000 System Requirements Study*. Census 2000 Evaluation R.2.c. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2002b, June 6). *American FactFinder System Requirements Study*. Census 2000 Evaluation R.3.b. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2002c, August 23). *Census 2000 Data Capture System Requirements Study*. Census 2000 Evaluation R.3.d. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2002d, June 6). *Laptop Computers for Accuracy and Coverage Evaluation System Requirements Study*. Census 2000 Evaluation R.2.b. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2002e, July 22). *Management Information System 2000 System Requirements Study*. Census 2000 Evaluation R.3.c. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2002f, April 17). *Matching Review and Coding System for Accuracy and Coverage Evaluation (Housing Unit, Person and Final Housing Unit) System Requirements Study*. Census 2000 Evaluation R.2.d. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2002g, February 28). *Operations Control System 2000 System Requirements Study*. Census 2000 Evaluation R.2.a. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2002h, June 6). *Pre-Appointment Management System/Automated Decennial Administrative Management System System Requirements Study*. Census 2000 Evaluation R.3.a. Washington, DC: U.S. Census Bureau.

Titan Systems Corporation (2003, January 14). *Operational Requirements Study: The Beta Site Systems Testing & Management Facility*. Census 2000 Evaluation L.5. Washington, DC: U.S. Census Bureau.

Treat, J. B. (1993). *1993 National Census Test Appeals and Long-Form Experiment, Long-Form Component, Final Report*. Washington, DC: Bureau of the Census, U.S. Department of Commerce.

Treat, J. B. (2003, September 30). *Response Rates and Behavior Analysis*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 11. Washington, DC: U.S. Census Bureau.

Trentham, S. and L. Larwood (2003). *Synthesis of Results from the Social Security Number, Privacy Attitudes, and Notification Experiment*. Census 2000 Testing, Experimentation, and Evaluation Program Report. Washington, DC: U.S. Census Bureau.

Tucker, C. and B. Kojetin (1996, September). Testing racial and ethnic origin questions in the CPS supplement. *Monthly Labor Review*, 3–7.

U.S. Census Bureau (1998, November). *Census 2000 Operational Plan*. Revised. Washington, DC: U.S. Department of Commerce.

U.S. Census Bureau (1999a, December). *1990 Data for Census 2000 Planning [Data CD]*. U.S. Department of Commerce.

U.S. Census Bureau (1999b, January). *Census 2000 Operational Plan Using Traditional Census-Taking Methods*. Washington, DC: U.S. Department of Commerce.

U.S. Census Bureau (2001a). *2000 Census Household Education Edit*. Unnumbered specifications document shared with panel (dated May). Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2001b, February 16). *P-Sample and E-Sample Dual-System Estimation Output Files*. DSSD Census 2000 Procedures and Operations Memorandum Series Q-38. Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2002a). *Specifications for Edit and Allocation of Housing Tenure*. Unnumbered specifications document shared with panel. Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2002b). *Specifications for Joint Economic Edit, 2000 Census Long Forms*. Unnumbered specifications document shared with panel. Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2003a, February). *Census 2000 Testing, Experimentation, and Evaluation Program Summary Documentation: Program Modifications Since May 2002*. Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2003b, March 12). *Decision on Intercensal Population Estimates*. Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2003c, March 12). *Technical Assessment of A.C.E. Revision II*. Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2003d). *Technical Documentation, Summary File 3, 2000 Census of Population and Housing*. Washington, DC: U.S. Census Bureau.

U.S. Census Monitoring Board Congressional Members (1999, September 30). *Unkept Promise: Statistical Adjustment Fails to Eliminate Undercounts, as Revealed by Evaluation of Severely Undercounted Blocks from the 1990 Census Plan*. Suitland, MD: U.S. Census Monitoring Board.

U.S. Census Monitoring Board Congressional Members (2001, May 23). *A Guide to Statistical Adjustment: How It Really Works*. Washington, DC: U.S. Census Monitoring Board.

U.S. Census Monitoring Board Presidential Members (2001a, September). *Final Report to Congress*. Washington, DC: U.S. Census Monitoring Board.

U.S. Census Monitoring Board Presidential Members (2001b, April). *Report to Congress*. Washington, DC: U.S. Census Monitoring Board.

U.S. Department of Commerce, Office of Inspector General (1997). *Bureau of the Census: Headquarters Information Processing Systems for the 2000 Decennial Census Require Technical and Management Plans and Procedures*. OSE-10034-8-0001. Washington, DC: U.S. Department of Commerce.

U.S. General Accounting Office (1992). *Decennial Census: 1990 Results Show Need for Fundamental Reform*. GAO/GGD-92-94. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (1997, February). *High Risk Series: Quick Reference Guide*. GAO/HR-97-2. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (1999a, September). *2000 Census: Analysis of Fiscal Year 2000 Amended Budget Request*. Report GAO/AIMD/GGD-99-291. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (1999b). *Formula Grants: Effects of the Adjusted Population Counts on Federal Funding to States*. GAO/HEHS-99-69. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (2000a, October). *2000 Census: Headquarters Processing System Status and Risks*. GAO-01-1. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (2000b, February). *2000 Census: New Data Capture System Progress and Risks*. GAO/AIMD-00-61. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (2000c, May). *2000 Census: Status of Nonresponse Follow-up and Key Operations*. GAO/T-GGD/AIMD-00-164. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (2000d, September). *2000 Census: Update on Data Capture Operations and System*. GAO/AIMD-00-324R. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (2001a, August). *2000 Census: Review of Partnership Program Highlights Best Practices for Future Operations*. Report GAO-01-579. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (2001b, December). *2000 Census: Significant Increase in Cost Per Housing Unit Compared to 1990 Census*. Report GAO-02-31. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (2002a, February). *2000 Census: Best Practices and Lessons Learned for More Cost-Effective Nonresponse Follow-up*. GAO-02-196. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (2002b). *2000 Census: Complete Costs of Coverage Evaluation Programs Are Not Available*. GAO-03-41. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (2002c). *2000 Census: Lessons Learned for Planning a More Cost-Effective 2010 Census*. GAO-03-40. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (2003a). *2000 Census: Coverage Measurement Programs' Results, Costs, and Lessons Learned*. GAO-03-287. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (2003b). *Decennial Census: Methods for Collecting and Reporting Data on the Homeless and Others without Conventional Housing Need Refinement*. GAO-03-227. Washington, DC: U.S. Government Printing Office.

Viator, M. A. and N. Alberti (2003, February 20). *Evaluation of Nonresponse Followup—Whole Household Usual Home Elsewhere Probe*. Census 2000 Evaluation I.2. Washington, DC: U.S. Census Bureau.

Vitrano, F., J. Treat, and R. Pennington (2003a, November 4). *Address List Development in Census 2000*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 8 (revised). Washington, DC: U.S. Census Bureau.

Vitrano, F., J. Treat, and R. Pennington (2003b, September 26). *Address List Development in Census 2000*. Census 2000 Testing, Experimentation, and Evaluation Program, Topic Report 8. Washington, DC: U.S. Census Bureau.

Waite, P. J., S. M. Obenski, and L. E. Buckley (2001, August). 2010 census planning: The strategy. Paper prepared for the Joint Statistical Meetings, Atlanta, GA. U.S. Census Bureau, Washington, DC.

Waters, M. C. (1990). *Ethnic Options: Choosing Identities in America*. Berkeley: University of California Press.

West, K. (1991, July). *1990 Post-Enumeration Survey Evaluation Project P9a: Accurate Measurement of Census Erroneous Enumerations—Evaluation Followup*. 1990 Coverage Studies and Evaluation Memorandum Series K-2. Washington, DC: Bureau of the Census, U.S. Department of Commerce.

West, K. K. and J. G. Robinson (2001, May). *The Use of Demographic Benchmarks to Ensure Census Data Quality*. Washington, DC: U.S. Census Bureau.

Westat (2002a, June 7). *Census 2000 Staffing Programs, Pay Component*. Census 2000 Evaluation G.1. Washington, DC: U.S. Census Bureau.

Westat (2002b, June 7). *Census 2000 Staffing Programs, Recruiting Component*. Census 2000 Evaluation G.1. Washington, DC: U.S. Census Bureau.

Westat (2002c, May 21). *Report of Survey of Partners*. Census 2000 Evaluation D.3. Washington, DC: U.S. Census Bureau.

Whitworth, E. (2002, August 14). *Internet Data Collection*. Census 2000 Evaluation A.2.b. Washington, DC: U.S. Census Bureau.

Wolfgang, G., T. S. Adams, P. Davis, X. Liu, and P. Stallone (2001). *ESCAP II: P-Sample Nonmatch Analysis*. Executive Steering Committee for A.C.E. Policy II Supporting Report 18 (revised). Washington, DC: U.S. Census Bureau.

Wolfgang, G., R. Byrne, and S. Spratt (2003, March 19). *Analysis of Proxy Data in the Accuracy and Coverage Evaluation*. Census 2000 Evaluation O.5. Washington, DC: U.S. Census Bureau.

Wolfgang, G., P. Stallone, and T. S. Adams (2002, September 5). *Targeted Extended Search Analysis*. Census 2000 Evaluation N.17. Washington, DC: U.S. Census Bureau.

Wolter, K., B. Calder, E. Malthouse, S. Murphy, S. Pedlow, and J. Porras (2002, July 17). *Partnership and Marketing Program Evaluation*. Census 2000 Evaluation D.1. Washington, DC: U.S. Census Bureau.

Working Group on LUCA (2001). *Assessment of the 2000 Census LUCA Program*. Washington, DC: Committee on National Statistics.

Wright, T. (2000). Sampling for follow-up of nonresponding households. In M. J. Anderson (Ed.), *Encyclopedia of the U.S. Census*, pp. 325–327. Washington, DC: CQ Press.

Zajac, K. J. (2002, May 23). *List/Enumerate*. Census 2000 Evaluation F.13. Washington, DC: U.S. Census Bureau.

Zajac, K. J. (2003, September 25). *Analysis of Item Imputation Rates for the 100 Percent Person and Housing Unit Data Items from Census 2000*. Census 2000 Evaluation B.1.a. Washington, DC: U.S. Census Bureau.

Zhao, Z. (2003). *Analysis of the Dual System Estimate in the 2000 Census*. Ph. D. thesis, Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia.

Biographical Sketches of Panel Members and Staff

Janet L. Norwood (*Chair*) is a counselor and senior fellow at the Conference Board, where she chairs the Advisory Committee on the Leading Indicators. She served as U.S. Commissioner of Labor Statistics from 1979 to 1992 and then was a senior fellow at the Urban Institute until 1999. She is a past member of the Committee on National Statistics and the Division of Engineering and Physical Sciences of the National Research Council. She is a member of advisory committees at the National Science Foundation, at several statistical agencies, and at universities, and she chairs the advisory committee for the Bureau of Transportation Statistics. She has received honorary LL.D. degrees from Carnegie Mellon, Florida International, Harvard, and Rutgers Universities. She is a fellow and past president of the American Statistical Association, a member and past vice president of the International Statistical Institute, an honorary fellow of the Royal Statistical Society, and a fellow of the National Academy of Public Administration and the National Association of Business Economists. She has a B.A. degree from Rutgers University and M.A. and Ph.D. degrees from the Fletcher School of Law and Diplomacy of Tufts University.

Robert M. Bell is a researcher at the AT&T Laboratories. He is a member of the National Research Council's Committee on National Statistics and previously served on its Panel on Alternative Census

Methodologies. Earlier, he worked at the RAND Corporation, where he examined and coauthored a number of studies on social policy issues. His work with RAND's Child Policy Project included research on such topics as prenatal substance abuse, adolescent drug use and the importance of social bonding among different ethnic groups, and the effects of the availability of condoms in a high school in Los Angeles. Dr. Bell earned a B.S. in mathematics from Harvey Mudd College and an M.S. from the University of Chicago and a Ph.D. from Stanford University, both in statistics.

Norman M. Bradburn is the assistant director of the National Science Foundation for Social, Behavioral and Economic Sciences. Formerly, he held positions as a senior vice president for research of the National Opinion Research Center (NORC) at the University of Chicago, director of NORC, and provost of the University. In his work as a social psychologist he has focused on the areas of survey and questionnaire design and related fields. He is also a past member of the Commission on Behavioral and Social Sciences and Education and the Committee on National Statistics and chaired the Panel to Evaluate Alternative Census Methodologies. He is a fellow of the American Statistical Association, the International Statistical Institute, and the American Association for the Advancement of Science, and a member of the American Association for Public Opinion Research and the World Association for Public Opinion Research. He received B.A. degrees from the University of Chicago and Oxford University and M.A. and Ph.D. degrees in clinical and social psychology from Harvard University.

Lawrence D. Brown is the Miers Bush professor of the Department of Statistics at the Wharton School at the University of Pennsylvania. He is a member of the National Academy of Sciences. He is also a member of the National Research Council's Committee on National Statistics and a former member of the National Research Council's Commission on Physical Sciences, Mathematics, and Applications and its Board on Mathematical Sciences. He was a critic of the Census Bureau's plans to incorporate sampling in the census. He is a fellow of the American Statistical Association and is a fellow and past president of the Institute of Mathematical Statistics. He received a B.S. from the California Institute of Technology and a Ph.D. from Cornell University.

Constance F. Citro is a senior program officer for the Committee on National Statistics. She is a former vice president and deputy director of Mathematica Policy Research, Inc., and was an American Statistical Association/National Science Foundation research fellow at the U.S. Census Bureau. For the committee, she has served as study director for numerous projects, including the Panel on Estimates of Poverty for Small Geographic Areas, the Panel on Poverty and Family Assistance, the Panel to Evaluate the Survey of Income and Program Participation, the Panel to Evaluate Microsimulation Models for Social Welfare Programs, and the Panel on Decennial Census Methodology. Her research has focused on the quality and accessibility of large, complex microdata files, as well as analysis related to income and poverty measurement. She is a fellow of the American Statistical Association. She received a B.A. degree from the University of Rochester and M.A. and Ph.D. degrees in political science from Yale University.

Michael L. Cohen is a senior program officer for the Committee on National Statistics, currently serving as co-study director for the Panel on Research on Future Census Methods and staff to the Panel to Review the 2000 Census. He previously assisted the Panel on Estimates of Poverty for Small Geographic Areas. He also directed the Panel on Statistical Methods for Testing and Evaluating Defense Systems. Formerly, he was a mathematical statistician at the Energy Information Administration, an assistant professor in the School of Public Affairs at the University of Maryland, and a visiting lecturer in statistics at Princeton University. His general area of research is the use of statistics in public policy, with particular interest in census undercount, model validation, and robust estimation. He is a fellow of the American Statistical Association. He received a B.S. degree in mathematics from the University of Michigan and M.S. and Ph.D. degrees in statistics from Stanford University.

Daniel L. Cork is a program officer for the Committee on National Statistics, currently serving as co-study director of the Panel on Research on Future Census Methods and assisting the Panel to Review the 2000 Census. He has also served as study director for the Workshop on Survey Automation. His research interests include quantitative criminology (particularly space-time dynamics in homicide), Bayesian statistics, and statistics in sports. He holds a B.S. degree in statistics from George Washington University and an M.S. in statistics and a joint Ph.D. in statistics and public policy from Carnegie Mellon University.

William F. Eddy is professor of statistics at Carnegie Mellon University. His research concentrates on the computational and graphical aspects of statistics. He is particularly interested in dynamic graphics for the analysis and presentation of data, especially those dynamic graphical displays that can be rendered interactively. He is a former member and incoming chair of the Committee on National Statistics; a fellow of the American Association for the Advancement of Science, the American Statistical Association, the Institute of Mathematical Statistics, and the Royal Statistical Society; and an elected member of the International Statistical Institute. He was the founding coeditor of *CHANCE* magazine and is the founding editor of the *Journal of Computational and Graphical Statistics*. He has an A.B. degree from Princeton University, and M.A., M.Phil., and Ph.D. degrees from Yale University.

Robert M. Hauser is Vilas research professor of sociology at the University of Wisconsin-Madison, where he has directed the Center for Demography and Ecology and the Institute for Research on Poverty. He currently directs the Center for Demography of Health and Aging. He is a member of the National Academy of Sciences and a fellow of the National Academy of Education, the American Association for the Advancement of Science, the American Statistical Association, the Center for Advanced Study in the Behavioral Sciences, and the American Academy of Arts and Sciences. He has served on the National Research Council's Committee on National Statistics, Commission on Behavioral and Social Sciences and Education, and Board on Testing and Assessment. His current research interests include trends in educational progression and social mobility in the United States among racial and ethnic groups, the uses of educational assessment as a policy tool, the effects of families on social and economic inequality, and changes in socioeconomic standing, health, and well-being across the life course. He received a B.A. degree from the University of Chicago and a Ph.D. degree from the University of Michigan.

Ingram Olkin is professor of statistics and education at Stanford University. Before moving to Stanford, he was on the faculties of Michigan State University and the University of Minnesota, where he served as chair of the Department of Statistics. He has cowritten a number of books and been an editor of the *Annals of Statistics* and an associate editor of *Psychometrika*, the *Journal of Educational Statistics*, and the *Journal of the American Statistical Association*. He has also served as chair of the National Research Council's Committee on Applied and Theoretical Statistics and as president of the Institute of Mathematical Statistics. He received an honorary D.Sci from De-Montfort University; a Lifetime Contribution Award from the American Psychological Association, Division 5; and a Wilks Medal from the American Statistical Association. His recent interest has been methodology for research synthesis, and he has published both theoretical and applied papers in medicine. He is currently a fellow in the Primary Care Outcomes Research Program. He received a Ph.D. in mathematical statistics from the University of North Carolina.

D. Bruce Petrie is vice-president and chief operating officer at the Canadian Institute for Health Information. From 1984 to 1999, he was the assistant chief statistician of the Social, Institutions and Labour Statistics Field at Statistics Canada, where he was responsible for conducting the census in Canada. Petrie held several positions at Statistics Canada, including director general of the Household Surveys Branch from 1978 to 1984 and director of the Labour Force Survey (the Canadian equivalent of the Current Population Survey) from 1973 to 1978. He previously served as a member of the Committee on National Statistics' Panel to Evaluate Alternative Census Methodologies. He has a bachelor of commerce degree from Dalhousie University and an M.B.A. from the University of Western Ontario.

Meyer Zitter is an independent demographic consultant. Formerly, he was chief of the Census Bureau's Population Division and also served as assistant director for international programs. He is a fellow of the American Statistical Association and a member of the International Statistical Institute and the International Union for the Scientific Study of Population. He has a B.B.A. degree from City College of New York.

Index to Chapters 1–10

A

B

C

D

E

H

K

L

M

Q

R

S

U

V

W