Experimentation and Evaluation Plans for the 2010 Census: Interim Report

Lawrence D. Brown, Michael L. Cohen, and Daniel L. Cork, Editors, Panel on the Design of the 2010 Census Program of Evaluations and Experiments, National Research Council

ISBN: 0-309-11546-9, 108 pages, 6 x 9, (2008)

EXPERIMENTATION AND EVALUATION PLANS FOR THE 2010 CENSUS: INTERIM REPORT

Panel on the Design of the 2010 Census Program of Evaluations and Experiments

Lawrence D. Brown, Michael L. Cohen and Daniel L. Cork, 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. YA1323-06-CN-0031 between the National Academy of Sciences and the U.S. Census Bureau. Support of the work of the Committee on National Statistics is provided by a consortium of federal agencies through a grant from the National Science Foundation (Number SBR-0112521). Any opinion, 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-13: 978-0-309-11545-2
International Standard Book Number-10: 0-309-11545-0

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

Printed in the United States of America

Suggested citation: National Research Council. (2008). *Experimentation and Evaluation Plans for the 2010 Census: Interim Report.* Panel on the Design of the 2010 Census Program of Evaluations and Experiments, Lawrence D. Brown, Michael L. Cohen, and Daniel L. Cork, editors. 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. Ralph J. Cicerone 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. Charles M. Vest 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. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

Panel on the Design of the 2010 Census Program of Evaluations and Experiments

LAWRENCE D. BROWN (*Chair*), Department of Statistics, Wharton School of Business, University of Pennsylvania
RICHARD A. BERK, Department of Criminology, University of Pennsylvania
ERIC BRADLOW, Wharton School of Business, University of Pennsylvania
IVAN P. FELLEGI, Statistics Canada, Ottawa
LINDA GAGE, California Department of Finance, Sacramento
VIJAY NAIR, Department of Statistics, University of Michigan, Ann Arbor
JESSE H. POORE, JR., Department of Computer Science, University of Tennessee, Knoxville
NORA CATE SCHAEFFER, Department of Sociology, University of Wisconsin–Madison
ALLEN L. SCHIRM, Mathematica Policy Research, Inc., Washington, DC
JUDITH SELTZER, Department of Sociology, University of California, Los Angeles
STANLEY SMITH, Department of Economics, University of Florida, Gainesville
ROGER TOURANGEAU, Survey Research Center, University of Michigan, Ann Arbor, and Joint Program in Survey Methodology, University of Maryland, College Park
KIRK WOLTER, National Opinion Research Center, University of Chicago

MICHAEL L. COHEN, *Co-Study Director*
DANIEL L. CORK, *Co-Study Director*
AGNES E. GASKIN, *Senior Program Assistant*
MEYER ZITTER, *Consultant*

Committee on National Statistics
2007-2008

ACKNOWLEDGMENTS

The Panel on the Design of the 2010 Census Program of Evaluations and Experiments (CPEX) wishes to thank the many people who have contributed to our work. To start, the panel is indebted to Deputy Director Preston Jay Waite's support during the developmental stage of the study. The contracting officer for the study is Deborah Bolton, and she has been assisted throughout by Gary Chappell, Joyce Price, Jennifer Reichert, and Donna Souders. All of these Census Bureau staff have been extremely responsive to requests from the panel for help in the development of agendas, the collection of supporting materials, and assistance with other logistical details. They have been a pleasure to work with.

Also, a number of Census Bureau personnel provided extremely useful presentations and supporting materials during the two plenary meetings of the panel on April 30–May 1, 2007, and July 16, 2007. In this regard, we would like to thank Theresa Angueira, Larry Cahoon, Joan Hill, Elizabeth Martin, Mary Mulry, Manuel de la Puente, Jennifer Reichert, Annetta Clark Smith, and Frank Vitrano. Many of these along with others also contributed to three very productive small group meetings of the panel (held July 13 and July 16, 2007), and here we would like to mention interactions with Nancy Bates, William Bell, Sharon Boyer, Mary Destasio, Donna Kostanich, Lauri Schwede, Jennifer Tancreto, and Jim Treat. We would like to single out Jim Treat to thank him for his help in learning about a number of disparate areas of census planning, experimentation, and evaluation.

We would like to thank Connie Citro for helping, along with Jay Waite, in the development of the study. In addition, Connie participated actively in the panel meetings, providing her usual insightful comments and suggestions. We would also like to thank Meyer Zitter, consultant to the panel, for providing important advice on the potential use of administrative records in census taking. We express our gratitude to Agnes Gaskin, who has seen to it that the meetings of the panel run as smoothly as possible, facilitating the travel and other needs of the panel members, and formatting this report. The panel is also indebted to Christine McShane, who provided expert technical editing of the draft report.

Finally, it has been a pleasure interacting with a very talented, energetic, and collaborative panel. We are excited at the prospect of working together in the future to provide additional assistance to the Census Bureau as it makes final preparations for the 2010 decennial census and the associated experiments and evaluations.

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. 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: C.A. (Al) Irvine, Consultant, San Diego, California; Benjamin King, Statistical Consultant, Durham, North Carolina; J. Michael Oakes, Division of Epidemiology, University of Minnesota, Minneapolis;

Joseph Salvo, Population Division, New York City Department of City Planning, New York, New York; Robert Scardamalia, Center for Research and Information Analysis, Department of Economic Development, Albany, New York; Frederick J. Scheuren, Consultant, Alexandria, Virginia; and Judith M. Tanur, Department of Sociology and Behavior, State University of New York, Stony Brook.

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 this report was overseen by William F. Eddy, Department of Statistics, Carnegie Mellon University. Appointed by the National Research Council, he was responsible for making certain that an independent examination of this 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 committee and the institution.

Lawrence Brown, *Chair*
Michael Cohen, *Co-Study Director*
Daniel Cork, *Co-Study Director*

Contents

Appendixes

EXECUTIVE SUMMARY

In connection with every recent decennial census, the U.S. Census Bureau has carried out experiments and evaluations. A census "experiment" usually involves field data collection during the census in which alternatives to current census processes are assessed for a subset of the population. An "evaluation" is usually a post hoc analysis of data collected as part of the decennial census processing to determine whether individual steps in the census operated as expected. The Census Bureau program for evaluations and experiments for the 2010 decennial census is referred to as the 2010 CPEX Program.

CPEX, like its predecessor programs, has enormous potential to help improve the next census, which is the federal government's single most important, and most costly, data collection activity. A well-planned and well-executed CPEX is a sound investment to ensure that the 2020 census is as cost-effective as possible.

The Census Bureau is now determining the topics for experiments during the 2010 census. The specific designs of the experiments have to be final by summer 2008 to meet the planning needs for the census. Because the data needed to support census evaluations are typically output files from the census itself, various post-censal data collections, and possibly extracts from administrative records, the exact structure of individual evaluations is not yet as time-sensitive as the experiments. However, some early planning for evaluations is crucial so that the necessary data extracts can be prepared and retained. This is especially true because much of the data collection in 2010 will be carried out by contractors, and so data retention requirements need to be arranged with contractors as early as possible.

The Panel on the Design of the 2010 Census Program of Evaluations and Experiments has been broadly charged to review proposed topics for evaluations and experiments and recommend priorities for them for the 2010 census, to consider what can be learned from the 2010 testing cycle to better plan for the 2020 census, and to assess the Census Bureau's overall continuing research program for the nation's decennial censuses.

The primary purpose of this interim report is to help reduce the possible subjects for census experimentation from an initial list of 52 research topics compiled by the Census Bureau to perhaps 6, which is consistent with the size of the experimentation program in 2000. This interim report also offers broad advice on plans for evaluations of the 2010 census. The panel expects to provide fuller details of individual experiments and evaluations in its subsequent reports.

CENSUS EXPERIMENTS

The panel identified three priority experiments for inclusion in the 2010 census to assist 2020 census planning (in one instance, there might be several related experiments): an experiment on

the use of the Internet for data collection; an experiment on the use of administrative records for various census purposes; and an experiment (or set of experiments) on features of the census questionnaire.

One important opportunity for improving census quality and possibly reducing census costs in 2020 is the use of the Internet as a means of enumeration. Although Internet response was permitted (but not advertised) in the 2000 census, the Census Bureau has elected not to allow online response in 2010. The panel does not second-guess that decision, but we think that it is essential to have a full and rigorous test of Internet methodologies in the 2010 CPEX. Internet response provides important advantages for data collection, including alternate ways of presenting residence rules and concepts, increased facility for the presentation of questionnaires in foreign languages, and real-time editing. It also has the feature of immediate transmission of data, which has important benefits regarding minimizing the overlap of census data collection operations.

RECOMMENDATION 1: The Census Bureau should include, in the 2010 census, a test of Internet data collection as an alternative means of enumeration. Such a test should investigate means of facilitating Internet response and should measure the impact on data quality, the expeditiousness of response, and the impact on the use of foreign language forms.

Another important opportunity for reducing costs and improving data quality is the use of administrative records. These are data collected as a by-product of the management of federal, state, and local governmental programs, such as birth and death records, building permit records, and welfare program records. In 2000, administrative records were the subject of an experiment intended to study their use as a complementary type of enumeration (that is, whether person counts for some geographic areas derived from records were consistent with census returns). However, administrative records could be used more broadly to assist a number of census tasks, including such uses as (1) to improve the Master Address File, (2) as an alternative to last-resort proxy response, (3) as an alternative to item and unit imputation, (4) to resolve duplicate search, (5) to validate edit protocols, (6) for coverage measurement and coverage evaluation, (7) for coverage improvement, and (8) to help target households for various purposes. It is important for the Census Bureau to determine, starting now, which of these various potential uses of administrative records would or would not be effective for use in 2020.

RECOMMENDATION 2: The Census Bureau should develop an experiment (or evaluation) that assesses the utility of administrative records for assistance in specific census component processes—for example, for improvement of the Master Address File, for nonresponse follow-up, for assessment of duplicate status, and for coverage improvement. In addition, either as an experiment or through evaluations, the Census Bureau should collect sufficient data to support assessment of the degree to which targeting various census processes, using administrative records, could reduce census costs or improve census quality.

Finally, given the crucial importance of the census questionnaire as a driver of census data quality, especially with regard to the nation's data on race and ethnicity, and to correctly locate

each person at the proper census residence, the Census Bureau should conduct either a large experiment or several smaller experiments on the content and method of presentation of the census questionnaire.

RECOMMENDATION 3: The Census Bureau should include one or more alternate questionnaire experiments during the 2010 census to examine:

- **the representation of questions on race and ethnicity on the census questionnaire, particularly asking about race and Hispanic origin as a single question;**
- **the representation of residence rules and concepts on the census questionnaire; and**
- **the usefulness of including new or improved questions or other information on the questionnaire with regard to (1) coverage probes, (2) the motivation of census questions, (3) the request of information on usual home elsewhere on group quarters questionnaires, and (4) deadline messaging and mailing dates for questionnaires.**

In such experiments, both the 2000 and the 2010 census questionnaires should be included in the assessments to serve as controls. The Census Bureau should explore the possibility of joining the recommended experiments listed above into a single experiment, through use of fractional factorial experimental designs.

CENSUS EVALUATIONS

It is important that sufficient data be retained to enable postcensus evaluations of the processes used to update the Master Address File from census to census. The success of a mailout-mailback census is most dependent on the quality of its address list, and therefore understanding the contribution of the various processes used to update the address list, especially Local Update of Census Addresses (LUCA) and address canvassing, is crucially important. In addition, given the expense of address canvassing in all blocks, it is important to be able to ascertain the extent to which canvassing can be targeted to blocks that are likely to have changes. Both administrative records, especially building permit data, and commercial mailing lists may have value in assisting in the targeting of blocks for canvassing.

RECOMMENDATION 4: The Census Bureau should design its Master Address File so that the complete operational history—when list-building operations have added, deleted, modified, or simply replicated a particular address record—can be reconstructed. This information will support a comprehensive evaluation of the Local Update of Census Addresses and address canvassing. In addition, sufficient information should be retained, including relevant information from administrative records and the American Community Survey, to support evaluations of methods for targeting blocks that may not benefit from block canvassing. Finally, efforts should be made to obtain addresses from commercial mailing lists to determine whether they also might be able to reduce the need for block canvassing.

More broadly, a master trace sample database could be used to address a substantial number of questions about the functioning of the 2010 census. Such a database would necessitate the retention of the entire census processing history (including the coverage measurement processes) of all addresses for a selected sample of areas, structured in a way to facilitate analysis. For example, such a database would help determine what percentage of census omissions are in partially enumerated households, or it could assess the benefits of the coverage follow-up interview. The panel therefore recommends that the process for creating such a database be initiated.

RECOMMENDATION 5: The Census Bureau should initiate efforts now for planning the general design of a master trace sample database and should plan for retention of the necessary information to support its creation.

Also, evaluations should be carried out on the feasibility of coverage measurement through use of a reverse record check based on the American Community Survey. The reverse record check is an alternative method for estimating the completeness of census coverage of the population, which may have advantages over the methods of dual-systems estimation and demographic analysis that have been used for this purpose to date.

RECOMMENDATION 6: The Census Bureau, through the use of an evaluation of the 2010 census (or an experiment in the 2010 census) should determine the extent to which the American Community Survey could be used as a means for evaluating the coverage of the decennial census through use of a reverse record check.

Finally, the Census Bureau has no program for assessing the rate of omissions of residents of group quarters in the 2010 census, nor can it assess the rate of placement of group quarters in the wrong census geography. The Census Bureau should therefore take the first steps toward remedying this by collecting sufficient information in 2010 to evaluate ideas on how to include this capability in the 2020 census coverage measurement program.

RECOMMENDATION 7: The Census Bureau should collect sufficient data in 2010 to support the evaluation of potential methods for assessing the omission rate of group quarters residents and the rate of locating group quarters in the wrong census geography. This is a step toward the goal of improving the accuracy of group quarters data.

OVERALL CENSUS RESEARCH PROGRAM

It appears that basic census research is not receiving the priority and support needed to best guide census redesign. For example, tests on some topics have been unnecessarily repeated, and previous research has sometimes been ignored in designing newer tests. Also, some topics, by their nature, require a relatively long time to understand and therefore need to be separated from the decennial census operational cycle. The lack of priority of research can also be seen in that the results of the 2006 test census tests were not all completed in time for the design of the 2008 census dress rehearsal. Research continuity is important not only to reduce redundancy and to ensure that findings are known and utilized, but also because there are a number of issues that

come up repeatedly over many censuses that are inherently complex and therefore benefit from testing in a variety of circumstances in an organized way, as unaffected as possible by the census cycle. Finally, given the fielding of the American Community Survey, there is now a real opportunity for research on census and survey methodology to be more continuous.

RECOMMENDATION 8: The Census Bureau should support a dedicated research program in census and survey methodology, whose work is relatively unaffected by the cycle of the decennial census. In that way, a body of research findings can be generated that will be relevant to more than one census and to other household surveys.

THE 2010 CENSUS DESIGN

In carrying out our charge to advise on the development of plans for experimentation and evaluation for the 2010 census, and more generally to review the full program of research and testing for improving census methodology, three issues arose that relate to the 2010 census design itself and, consequently, its evaluation. While the panel is aware that most aspects of the 2010 census design have already been decided and cannot be easily changed given time constraints, there remains the possibility that some of the following recommendations may still be able to be acted on prior to 2010.

The first issue is the possibility of the introduction of errors into the data collection transmissions by the handheld computing devices that will be used to follow up households that do not return a mail questionnaire. The second issue is the possibility of interoperability problems in the various software systems constituting the management information system for the 2010 census. The third issue is the role of telephone questionnaire assistance in 2010.

RECOMMENDATION 9: The Census Bureau should use dual-recording systems, quantitative validation metrics, dedicated processing systems, periodic system checkpoints, strict control over handheld devices, and related techniques to ensure and then verify the accuracy of the data collected from handheld computing devices.

RECOMMENDATION 10: The Census Bureau should provide for a check to ensure that the subsystems of the management information system used in 2010 have no interoperability problems.

RECOMMENDATION 11: The Census Bureau should strongly consider, for the 2010 census, explicit encouragement of the collection of all data on the census questionnaire for people using Telephone Questionnaire Assistance. In addition, the Census Bureau should collect sufficient information to estimate the percentage of callers to Telephone Questionnaire Assistance who did not ultimately send back their census questionnaires. This would provide an estimate of the additional costs of nonresponse follow-up due to the failure to collect the entire census questionnaire for those cases. The Census Bureau should also consider carrying out an experiment whereby a sample of callers to Telephone Questionnaire Assistance are asked whether they would mind providing their full

information to better estimate the additional resources required as a result of expanding Telephone Questionnaire Assistance in this way.

1

Introduction

The Census of Population and Housing is carried out in the United States every 10 years, and the next census is scheduled to begin its mailout-mailback operations in March 2010. For at least the past 50 years, each decennial census has been accompanied by a research program of evaluation or experimentation. The Census Bureau typically refers to a census "experiment" as a study involving field data collection—typically carried out simultaneously with the decennial census itself—in which alternatives to census processes currently in use are assessed for a subset of the population. By comparison, census "evaluations" are usually post hoc analyses of data collected as part of the decennial census process to determine whether individual steps in the census operated as expected. Collectively, census experiments and evaluations are designed to inform the Census Bureau as to the quality of the processes and results of the census, as well as to help plan for modifications and innovations that will improve the (cost) efficiency and accuracy of the next census. The Census Bureau is currently developing a set of evaluations and experiments to accompany the 2010 census, which the Bureau refers to as the 2010 Census Program for Evaluations and Experiments or CPEX.

These two activities of the more general census research program are concentrated during the conduct of the census itself, but census-related research activities continue throughout the decade. Traditionally, the Census Bureau's intercensal research has been focused on a series of census tests, some of which are better described as "test censuses" because they are conducted in specific geographic areas and can include fieldwork (e.g., in-person follow-up for nonresponse) as well as contact through the mail or other means. The sequence of tests usually culminates in a dress rehearsal two years prior to the decennial census. In addition to the test censuses, the Census Bureau has also conducted some smaller scale experimental data collections during the intercensal period.

CHARGE TO THE PANEL

As it began to design its CPEX program for 2010, the Census Bureau requested that the Committee on National Statistics of the National Academies convene the Panel on the Design of the 2010 Census Program of Evaluations and Experiments. The panel's charge is to:

> . . . consider priorities for evaluation and experimentation in the 2010 census. [The panel] will also consider the design and documentation of the Master Address File and operational databases to facilitate research and evaluation, the design of experiments to embed in the 2010 census, the design of evaluations of the 2010 census processes, and what can be learned from the pre-2010 testing that was conducted in 2003-2006 to enhance the

> testing to be conducted in 2012-2016 to support census planning for 2020. Topic areas for research, evaluation, and testing that would come within the panel's scope include questionnaire design, address updating, nonresponse follow-up, coverage follow-up, unduplication of housing units and residents, editing and imputation procedures, and other census operations. Evaluations of data quality would also be within scope. . . .

More succinctly, the Census Bureau requests that the panel:

- Review proposed topics for evaluations and experiments;
- Assess the completeness and relevance of the proposed topics for evaluation and experimentation;
- Suggest additional research topics and questions;
- Recommend priorities;
- Review and comment on methods for conducting evaluations and experiments; and
- Consider what can be learned from the 2010 testing cycle to better plan research for 2020.

The panel is charged with evaluating the 2010 census research program, primarily in setting the stage for the 2020 census. As the first task, the panel was asked to review an initial list of research topics compiled by the Census Bureau, with an eye toward identifying priorities for specific experiments and evaluations in 2010. This first interim report by the panel uses the Bureau's initial suggestions for consideration as a basis for commentary on the overall shape of the research program surrounding the 2010 census and leading up to the 2020 census. It is specifically the goal of this report to suggest priorities for the experiments to be conducted in line with the 2010 census because they are the most time-sensitive. To some observers, a two-year time span between now and the fielding of the 2010 census may seem like a long time; in the context of planning an effort as complex as the decennial census, however, it is actually quite fleeting. Experimental treatments must be specified, questionnaires must be tested and approved, and systems must be developed and integrated with standard census processes—all at the same time that the Bureau is engaged in an extensive dress rehearsal and final preparations for what has long been the federal government's largest and most complex non-military operation. Accordingly, the Census Bureau plans to identify topics for census experiments to be finalized by winter 2007 and to have more detailed plans in place in summer 2008; this report is an early step in that effort.

Although this report is primarily about priorities for experiments, we also discuss the evaluation component of the CPEX. This is because even the basic possibilities for specific evaluations depend critically on the data that are collected during the conduct of the census itself. Hence, we offer comments about the need to finalize plans for 2010 data collection—whether in house by the Census Bureau or through its technical contractors—in order to facilitate a rich and useful evaluation program.

We will continue to study the CPEX program over the next few years, and we expect to issue at least one more report; these subsequent reports will respond to the Bureau's evolving development of the CPEX plan as well as provide more detailed guidance for the conduct of specific evaluations and experiments.

BACKGROUND: EXPERIMENTS AND EVALUATIONS IN THE 2000 CENSUS

As context for the discussion that follows and to get a sense of the scope of CPEX, it is useful to briefly review the experiments and evaluations of the previous census. The results of the full Census 2000 Testing, Experimentation, and Evaluation Program are summarized by Abramson (2004).

Experiments

The Census Bureau carried out five experiments in conjunction with the 2000 census. Several ethnographic studies were also conducted during the 2000 census; about half of these were considered to be part of the formal evaluation program, whereas the others were designated as a sixth experiment.

Census 2000 Alternative Questionnaire Experiment (AQE2000): AQE2000 comprised three experiments for households in the mailout-mailback universe of the 2000 census. The *skip instruction experiment* examined the effectiveness of different methods for guiding respondents through an alternative long-form questionnaire with skip patterns. The *residence instructions experiment* tested various methods (format, presentation, and wording of instructions) for representing the decennial census residence rules on the questionnaire. The hope was to improve within-household coverage by modifying the roster instructions. Finally, the *race and Hispanic origin experiment* compared the 1990 race and Hispanic origin questions with the questions on the Census 2000 short form, specifically assessing the effect of permitting the reporting of more than one race and reversing the sequence of the race and Hispanic origin items. This experiment is summarized by Martin et al. (2004).

Administrative Records Census 2000 Experiment (AREX 2000): AREX 2000 was designed to assess the value of administrative records data in conducting an administrative records census. As a by-product, it also provided useful information as to the value of administrative records in carrying out or assisting in various applications in support of conventional decennial census processes. AREX 2000 used administrative records to provide information on household counts, date of birth, race, Hispanic origin, and sex, linked to a corresponding block code.

The test was carried out in five counties in two sites, Baltimore City and Baltimore County, Maryland, and Douglas, El Paso, and Jefferson counties in Colorado, with approximately 1 million housing units and a population of approximately 2 million. The population coverage for the more thorough of the schemes tested was between 96 and 102 percent relative to the Census 2000 counts for the five test site counties. However, the AREX 2000 and the census counted the same number of people in a housing unit only

51.1 percent of the time. They differed by at most one person only 79.4 percent of the time. The differences between the administrative records–based counts and the census counts were primarily attributed to errors in address linkage and typical deficiencies in administrative records (missed children, lack of representation of special populations, and deficiencies resulting from the time gap between the administrative records extracts and Census Day). Another important finding was that administrative records are not currently a good source of data for race and Hispanic origin, and the models used to impute race and Hispanic origin were not sufficient to correct the deficiencies in the data. The experiment is summarized by Bye and Judson (2004).

Social Security Number, Privacy Attitudes, and Notification Experiment (SPAN): This experiment assessed the public's attitudes regarding the census and its uses, trust and privacy issues, the Census Bureau's confidentiality practices, possible data sharing across federal agencies, and the willingness of individuals to provide their Social Security number on the decennial census questionnaire. In addition, the public's attitude toward the use of administrative records in taking the census was also assessed. The experiment is described in detail by Larwood and Trentham (2004).

Response Mode and Incentive Experiment (RMIE): The RMIE investigated the impact of three computer-assisted data collection techniques: computer-assisted telephone interviewing (CATI), the Internet, and interactive voice response, on the response rate and quality of the data collected. The households in six panels were given the choice of providing their data via the usual paper forms or by one of these alternate modes. Half of the panels were offered an incentive—a telephone calling card good for 30 minutes of calls—for using the alternate response mode. In addition, the experiment included a nonresponse component designed to assess the effects on response of an incentive to use alternative response mode options among a sample of census households that failed to return their census forms by April 26, 2000. This was to test the effect of these factors on a group representing those who would be difficult to enumerate. A final component of this experiment involved interviewing households assigned to the Internet mode who opted to complete the traditional paper census form to determine why these households did not use the Internet. One of the findings was that the Internet provided relatively high-quality data. However, among respondents who were aware of the Internet option, 35 percent reported that they believed the paper census form would be easier to complete. Other reasons for not using the Internet include no access to a computer, concerns about privacy, "forgot the Internet was an option," and insufficient knowledge of the Internet. The incentive did not increase response but instead redirected response to the alternate modes. The CATI option seemed to be preferred over the other two alternate modes. Caspar (2004) summarizes the experiment's results.

Census 2000 Supplementary Survey (C2SS): By 1999, the basic notion that the new American Community Survey (ACS) would take the role of the traditional census long-form sample had been established (this is discussed in more detail in the next section). ACS testing had grown to include fielding in about 30 test sites (counties), with full-scale implementation planned for the 2000-2010 intercensal period. Hence, the Census Bureau was interested in some assessment of the operational feasibility of conducting a large-

scale ACS at the same time as a decennial census. Formally an experiment in the 2000 census program, the C2SS escalated ACS data collection to include more than one-third of all counties in the United States; this step-up in collection—while well short of full-scale implementation—offered a chance to compare ACS estimates with those from the 2000 census. Operational feasibility was defined as the C2SS tasks being executed on time and within budget with the collected data meeting basic quality standards. No concerns about the operational feasibility of taking the ACS in 2010 were found. Griffin and Obenski (2001) wrote a summary report on the operational feasibility of the ACS, based on the C2SS.

Ethnographic Studies: Three studies were included in this experiment. One study examined the representation of and responses from complex households in the decennial census through ethnographic studies of six race/ethnic groups (Schwede, 2003). A second study examined shared attitudes among those individuals following the "baby boomers, i.e., those born between 1965 and 1975, about civic engagement and community involvement, government in general, and decennial census participation in particular (Crowley, 2003). A third study examined factors that respondents considered when they were asked to provide information about themselves in a variety of modes (Gerber, 2003). This research suggested that the following factors may contribute to decennial noncompliance and undercoverage errors: (1) noncitizenship status or unstable immigration status, (2) respondents not knowing about the decennial census, and (3) increased levels of distrust among respondents toward the government.

Evaluations

The Census Bureau initially planned to conduct 149 evaluation studies to assess the quality of 2000 census operations. Due to various resource constraints, as well as the overlap of some of the studies with assessments needed to evaluate the quality of the 2000 estimates of net undercoverage, 91 studies were completed. These evaluations were summarized in various topic reports, the subjects of which are listed in Table 1-1.

POST HOC ASSESSMENT OF THE 2000 EXPERIMENTS AND EVALUATIONS

We have described six experiments that were embedded in the 2000 census. We can now look back at these experiments to see the extent to which they were able to play a role in impacting the design of the 2010 census. In doing that we hope to learn how to improve the selection of experiments in the 2010 census, looking toward the design of the 2020 census. Before continuing, it is important to note that the very basic design of the 2010 census was determined before these 2000 census experiments had been carried out. Therefore, at a fundamental level, the 2000 census experiments were always limited in their impact on key aspects of the basic design of the next census.

On the one hand, with the benefit of hindsight, the choice of the general subject matter for these six experiments can be viewed as relatively successful, since many of the basic issues identified for experimentation were relevant to the design of the 2010 census. The

utility of information from administrative records for census purposes, the advantages and disadvantages of Internet data collection, various aspects of census questionnaire design, and the operational feasibility of the American Community Survey being carried out during a decennial census were issues for which additional information was needed to finalize the 2010 design.

On the other hand, the details of these studies also indicate that they could have played a more integral role in the specification of the design for the 2010 census if they had been modified in relatively modest ways. For example, as a test of residence instructions, AQE2000 varied many factors simultaneously so that individual design effects were difficult to separate out. Also, the test of long-form routing instructions was largely irrelevant to a short-form-only census. AREX 2000 focused on the use of administrative records to serve in place of the current census enumeration, whereas examination of the use of administrative records to help with specific operations, such as for targeted improvements in the Master Address File, to assist in late nonresponse follow-up, or to assist with coverage measurement, would have been more useful. The response mode and incentive experiment examined the use of incentives to increase use of the Internet as a mode of response, but they did not examine other ways to potentially facilitate and improve Internet usage. Finally, the Social Security Number, Privacy, and Notification Experiment did not have any direct bearing on the 2010 design.

It bears repeating that it is an enormous challenge to anticipate what issues will be prominent in census designs for a census that will not be finalized for at least eight years after the census experiments themselves need to be finalized. Since one goal of this panel study is to help the Census Bureau select useful experiments for the 2010 census, our hope is that, when looking back in 2017, the 2010 census experiments will be seen as very useful in helping to select an effective design for the 2020 census.

With respect to the 2000 evaluations, the National Research Council report *The 2000 Census: Counting Under Adversity* provided an assessment of the utility of these studies, with which we are in agreement. The study group found (National Research Council, 2004b:331-332):

> 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. . . . This information is valuable but limited. Many reports have no analysis as such, other than simple one-way and two-way tabulations. . . . Almost no reports provide tables or other analyses that look at operations and data quality for geographic areas. . . . 2010 planners need analysis that is explicitly designed to answer important questions for research and testing to improve the 2010 census. . . . Imaginative data analysis [techniques] could yield important findings as well as facilitate effective presentation of results.

OVERVIEW OF THE 2010 CENSUS

While the 2000 census was still under way, the Census Bureau began to develop a framework for the 2010 census. Originally likened to a three-legged stool, this framework was predicated on three major initiatives:

1. The traditional long-form sample—in which roughly one-sixth of census respondents would receive a detailed questionnaire covering social, economic, and demographic characteristics—would be replaced by a continuing household survey throughout the decade, the American Community Survey, thus freeing the 2010 census to be a short-form-only enumeration;
2. Improvements would be made to the Census Bureau's Master Address File (MAF) and its associated geographic database (the Topologically Integrated Geographic Encoding and Referencing System, or TIGER, database) in order to save field time and costs; and
3. A program of early, integrated planning would be implemented in order to forestall an end-of-decade crunch in finalizing a design for the 2010 census.

Reengineering the 2010 Census: Risks and Challenges reviews the early development of the 2010 census plan, noting an immediate adjunct to the basic three-legged plan: the incorporation of new technology in the census process (National Research Council, 2004a). Specifically, the 2010 census plan incorporated the view that handheld computers could be used in several census operations in order to reduce field data collection costs and improve data quality. Following a series of decisions not to adjust the counts from the 2000 census for estimated coverage errors, the Census Bureau also established the basic precept that the 2010 census coverage measurement program would be used primarily to support a feedback loop of census improvement rather than for census adjustment.

As the 2010 census plan has developed, major differences between the 2010 plan and its 2000 predecessor—in addition to the broad changes already described—include:

- The use of handheld computers by field enumerators has been focused on three major operations: updating the Master Address File during the address canvassing procedure, conducting nonresponse follow-up interviewing, and implementing a new coverage follow-up (CFU) operation.
- The coverage follow-up interview is a consolidation and substantial expansion of a telephone follow-up operation used in the 2000 census, which was focused on following up households with count discrepancies and households with more than the six maximum residents allowed on the census form. While detailed plans for this follow-up operation are as yet incomplete, it appears that the CFU in 2010 will also follow up households with evidence of having duplicate enumerations, with people viewed as residents who possibly should have been enumerated elsewhere, and with people viewed as nonresidents who may have been incorrectly omitted from the count of that household.

- The Local Update of Census Addresses (LUCA) program, which gives local and tribal governments an opportunity to review and suggest additions to the Master Address File from their areas, has been revised to facilitate participation by local governments and to enhance communication between Census Bureau and local officials.
- Nonrespondents to the initial questionnaire mailing will be sent a replacement questionnaire to improve mail response.
- Households in selected geographic areas will be mailed a bilingual census questionnaire in Spanish and English.
- The census questionnaire will include two "coverage probe" questions to encourage correct responses (and to serve as a trigger to inclusion in the CFU operation).
- The definitions of group quarters—nonhousehold settings like college dormitories, nursing homes, military barracks, and correctional facilities—have been revised.
- Continuing a trend from 2000, the Census Bureau will increasingly rely on outside contractors to carry out several of the processes.

THE CPEX PLANNING DOCUMENT

This, the panel's first interim report, provides a review of the current status of the experimentation and evaluation plans of the Census Bureau heading into the 2010 census. As the major input to the panel's first meeting and our work to date, the Census Bureau provided a list of 52 issues, reprinted as Appendix A, corresponding to component processes of the 2010 census design that were viewed either as potentially capable of improvement or of sufficient concern to warrant a careful assessment of their performance in 2010. The list, divided into the following 11 categories, was provided to us as the set of issues that the Census Bureau judged as possibly benefiting from either experimentation in 2010 or evaluation after the 2010 census has concluded:

1. content,
2. race and Hispanic origin,
3. privacy,
4. language,
5. self-response options,
6. mode effects,
7. special places and group quarters,
8. marketing/publicity,
9. field activities,
10. coverage improvement, and
11. coverage measurement.

In addition to the description of the topics themselves, the Census Bureau also provided indications as to whether these topics have a high priority, whether they could potentially save substantial funds in the 2020 census, whether results could conclusively measure the

effects on census data quality, whether the issue addresses operations that are new since the 2000 census, and whether data will be available to answer the questions posed.

This list of topics was a useful start to the panel's work, but, as discussed more below, it is deficient in some ways, especially since it is not separated into potential experiments or evaluations and does not contain quantitative information on cost or quality implications. Also, such a list of topics needs to be further considered in the context of a general scheme for the 2020 census.

GUIDE TO THE REPORT

The remainder of this report is structured as follows. Chapter 2 provides initial views on the 2010 census experiments. There is a first section on a general approach to the selection of census experiments, which is followed by the panel's recommended priorities for topics for experimentation in 2010. Chapter 3 begins with suggestions for the 2010 census evaluations, which is followed by a general approach to census evaluation, and which concludes with considerations regarding a general approach to census research. Chapter 4 presents additional considerations for the 2010 census itself. It begins with technology concerns for 2010, followed by a discussion of the issue of data retention by census contractors. The chapter concludes with a discussion of the benefits of facilitating census enumeration as part of telephone questionnaire assistance. Appendix A provides the Census Bureau's summaries of suggested research topics for experiments and evaluations in 2010. Appendix B summarizes Internet response options in the 2000 U.S. census and in selected population censuses in other countries. Appendix C presents biographical sketches of panel members and staff.

TABLE 1-1 Topic Headings, 2010 CPEX Research Proposals and 2000 Census Evaluation Program

2010 CPEX Proposals	2000 Census Evaluation Topic Reports
Content	Content and data quality
Coverage improvement	Coverage improvement
Address list development	Address list development
Administrative records	*AREX2000 experiment*[a]
Coverage follow-up	*Partial: Coverage improvement*
Residency rules/question development	*AQE2000 experiment*
Be Counted	*Partial: Response rates and behavior analysis*
General	—
Coverage measurement	Coverage measurement
Field activities	
Automation	*Partial: Automation of census 2000 processes*
Training	—
Quality control	*Partial: Content and data quality*
Language	*Partial: Response rates and behavior analysis*
Marketing/publicity/advertising/partnerships	Partnership and marketing program
Mode effects	—
Privacy	Privacy research in census 2000[b]
Race and Hispanic origin	Race and ethnicity
Self-response options	—
Special places and group quarters	Special places and group quarters
—	Automation of census 2000 processes
—	Data capture
—	Data collection
—	Data processing
—	Ethnographic studies[c]
—	Puerto Rico
—	Response rates and behavior analysis

Note: The italics in the entries indicate deviations from the column heading, "2000 Census Evaluation Topic Reports." Some of the entries were not topic reports but were experiments. Also, some of the operations were part of the 2000 Coverage Improvement report.

[a]Described as partial match because the CPEX proposals under automation are oriented principally at one component (handheld computers).

[b]Privacy was also touched on by the Social Security Number, Privacy Attitudes, and Notification (SPAN) experiment.

[c]The 2000 census included several ethnographic studies; administratively, about half were considered part of the experiments while others were formally designated as evaluations (and were the subject of a topic report).

2

Initial Views on 2010 Census Experiments

A GENERAL APPROACH TO THE SELECTION OF CENSUS EXPERIMENTS

The Census Bureau provided the panel with a list of 52 topics for experimentation or evaluation, categorized into 11 general headings (see Appendix A). In addition to the topics themselves, the Census Bureau provided indications as to (a) whether modification of the relevant census processes have a high priority, (b) whether modification of the relevant census processes could potentially save substantial funds in the 2020 census, (c) whether results of an experiment could conclusively measure the effects on census data quality, (d) whether the issue addresses operations that are new since the 2000 census, and (e) whether data will be available to answer the particular questions posed. The panel found these topics and the associated assessments very helpful in focusing our work. The assessments of these topics, in particular, represent a considerable advance over the processes used to select the evaluations and experiments prior to the 2000 census.

However, we think that the Census Bureau can go further, when preparing for the analogous 2020 CPEX program, by providing a more developed context for evaluating various topics for potential census experiments. It is difficult to develop priorities without some sense of the collection of census designs that are under serious consideration. For example, it was not useful, at least from a decennial census perspective, to test skip patterns for the long form in 2000 given that the likely design in 2010 was a short-form-only census (although it may have been useful in support of the American Community Survey). Similarly, it was not useful to test an administrative records census in the Administrative Records Census 2000 Experiment when that was a remote possibility for the 2010 census. We understand that it will not be possible for the Census Bureau to produce a single proposal for the general design of the next census when it is time to select the experiments and evaluations for the current census, but it should be possible to produce a relatively small number of leading alternative designs that are under consideration. To help define possible designs, fundamental questions like the following might be asked:

- Could the telephone or the Internet be used more broadly as an alternative to mailing back census questionnaires for data collection?
- Could administrative records or other data sources be used to better target various operations?
- Could administrative records be used to augment last-resort or proxy enumeration in the latter stages of nonresponse follow-up?

Having a set of designs that are under consideration helps to direct the experimentation toward resolving important issues that discriminate among the designs.

Although we realize that the following are not readily available, in the future it would also be useful to have, for both the current census processes and, to the extent possible, any alternative approaches: (1) estimates of census costs by component operation (and the recent history of costs)[1] and (2) the potential impact on the quality of the collected data by component operation. The attribution of both coverage and characteristics error, to component operations or current processes, let alone suggested alternatives, on a national level, not to mention for demographic subgroups, would have been very difficult to achieve in past censuses. The planned census coverage measurement program in 2010 is hoping to make progress in assessing and attributing component coverage error to various sources. This is an important development because the Census Bureau could better justify priorities in undertaking various experiments by providing information on the impact on costs and quality of various alternatives. Furthermore, even if estimates of costs and impacts on accuracy are difficult to estimate, it should generally be possible to determine the major cost drivers and the leading sources of error.

There are two other modifications to the Census Bureau's list of topics that would have facilitated setting priorities. First, it would have been helpful if the list had been separated into candidates for evaluations and candidates for formal experiments. An experiment is, generally speaking, not possible until a reasonable alternative has been identified. Therefore, the listing of any alternative methodologies along with any knowledge of their potential advantages and disadvantages will facilitate the discussion of which issues should be focused on for either experimentation or evaluation. Second, a summary of the current state of research on some of the issues described would have been helpful (in Appendix A, the column on "new to census" is related to this). While some of these issues are extremely new, some, for example questionnaire design, are topics for which the Census Bureau has a history of relevant research. This information would have supported a more refined judgment of the likelihood that use of various alternative approaches might lead to important improvements.

PRIORITY TOPICS FOR EXPERIMENTATION IN 2010

So, without an overall strategy for the design of the 2020 census, it was difficult for the panel to develop strict priorities for the topics that should and should not be examined through the use of experiments in the 2010 census. T his lack of a strategy could have been overcome to some degree with information on the potential impact on census costs and accuracy of replacing various census component processes with alternative processes. This is so because the overall goal of research on census methods has at its most basic level two main objectives: reducing costs and improving accuracy. However, this information is not available at this point and so the panel developed the following set of priority topics for experiments based on speculations concerning the possible designs of the 2020 census and qualitative information on the potential impact on costs and

[1] It is useful to note here that the cost of the 2010 census is projected to be over $11 billion, which is approximately $100 per housing unit. Therefore, the use of any alternatives that have substantial cost savings is a crucial benefit in looking toward the 2020 census.

accuracy from the use of alternative processes. In the same vein, the primary goal of each experiment that we are recommending for priority consideration is to better understand the impacts on both census costs and census data quality resulting from the use of alternatives to current census methodology.

The three recommendations in this chapter on experimentation should be considered by the Census Bureau as the three highest priority recommendations in this report. Throughout, the panel was mindful of the special context that the decennial census provides for experimentation, and therefore one additional criterion applied was whether experimentation for the topic under consideration would substantially benefit from a decennial census environment.

To start, we put forward two topics for experimentation that were not given sufficient prominence in the list provided by the Census Bureau (see Appendix A).[2] Internet data collection was not mentioned in the list, and the use of administrative records was mentioned very briefly (items A.6 and C.6 in Appendix A) as possibly playing a role in augmenting coverage measurement data collection, in otherwise identifying coverage problems, and in identifying and classifying duplicates. These are both very important mechanisms for improved data collection and improved evaluation.

Before expanding on those two issues, we also mention that research and experimentation on the American Community Survey (ACS) were not mentioned prominently in the 2010 Census Program for Evaluations and Experiments (CPEX) plan. We understand that ACS research and testing are intended to be handled separately, possibly using an experimental methods panel to identify improvements in ACS methodology. However, there are important commonalities between the effectiveness of methods used to collect ACS data and the methods used to collect decennial census data that need to be exploited. It is very likely that more efficient and better research will be possible by combining perspectives from both operations. An explicit recognition of both the crucial need for an ACS research and experimentation program (this is recommended in National Research Council, 2007) and the potential for cross-fertilization of such an ACS program with the CPEX program would be extremely desirable. Furthermore, given that the ACS and the decennial census will be collecting data simultaneously, measurement of the possible impact of the ACS on decennial census data collection, especially coverage follow-up (CFU) and possibly the coverage measurement effort, would be worthwhile. Finally, as we discuss below, the possible impact of the different residence concepts used by the census and the ACS is a major concern that can and should be assessed as part of the 2010 CPEX.

[2] Recall that the Census Bureau typically refers to a census experiment as a study involving field data collection—typically carried out simultaneously with the decennial census itself—in which alternatives to census processes currently in use are assessed for a subset of the population. Census evaluations are usually post hoc analyses of data collected as part of the decennial census process to determine whether individual steps in the census operated as expected.

Internet Data Collection

The Internet is becoming the preferred mode for many households to conduct their banking, shopping, tax filing, and other official communications and interactions. It is anticipated that the Internet will also soon become a major medium for survey data collection. In the decennial census context, the Internet provides important advantages, including alternate ways of representing residence rules, increased facility for the presentation of questionnaires in foreign languages, real-time editing, and immediate transmission of data, which has important benefits for minimizing the overlap of census data collection operations. With respect to the representation of residence concepts, an Internet-based questionnaire could make it easier to display (and link to) additional examples and instructions for determining census residence; it could also guide respondents through a more detailed set of probe questions in order to more accurately determine household counts. An Internet option could provide linguistically challenged respondents with a wider array of questionnaire assistance tools and, perhaps, administration of the actual census questions in more languages than has been feasible under the financial and logistical constraints of paper administration.

The experience in many other countries (see Appendix B for details) is that this alternative mode of response provides important benefits, which are likely to increase as 2020 advances. In particular, the recent 2006 Canadian experience is that the use of the Internet as a response option does improve the quality and timeliness of responses (Statistics Canada, 2007).

As described in Appendix B, the Census Bureau has decided against the use of the Internet in 2010 for two principal reasons. First, it believes that it is unlikely to appreciably improve the rate of response given the results of the 2003 and 2005 National Census Tests. Second, there are issues related to security that need to be considered, including the potential for hackers to disrupt the data collection, in addition to any public perception problems that are related to security concerns.[3]

It is not our charge to evaluate the Census Bureau's decision not to use the Internet for data collection in the 2010 census. However, it is obvious from the discussion in Appendix B that many countries are already strongly moving in this direction. More importantly, given the advantages listed above and the anticipation of greater advantages in the future, the Census Bureau needs to start now to prepare for use of the Internet as a major means for data collection in the 2020 census. An important step in this preparation is the inclusion of an experiment on Internet data collection in the 2010 census.

[3] We note that there is generally little concern about biases in responses received by the Internet, for two reasons. First, there will always be multiple modes for response in the census given the heterogeneous population that is being counted. So mode bias is ubiquitous. Second, mode bias for the questions on the census short form will be relatively modest since there is little room for interpretation, except possibly for residence rules and race/ethnicity.

Regarding possible problems in access to and use of the Internet, the panel thinks that there may be alternative ways of interfacing with respondents that could facilitate Internet response, rather than using the mailed questionnaire as the initiating event. Regarding security concerns, Canada and other countries have been able to successfully mitigate security concerns, and it thus seems likely that the United States should be able to address this issue in time for 2020.

While the testing of an Internet response option does not require a census context, a census context would be very useful, since complex counting rules, needed for unduplicating double counts, are more easily implemented in a complete count operation. Also, response frequency is substantially higher in the census than in test censuses.

We therefore recommend that the Census Bureau include an experiment during the 2010 census that uses alternative mechanisms to facilitate Internet responses and measures the frequency of use for each, along with expeditiousness and quality of response. It may also be possible to ask the respondent if he or she would utilize an online foreign language version if available.

RECOMMENDATION 1: The Census Bureau should include, in the 2010 census, a test of Internet data collection as an alternative means of enumeration. Such a test should investigate means of facilitating Internet response and should measure the impact on data quality, the expeditiousness of response, and the impact on the use of foreign language forms.

Use of Administrative Records to Assist in Component Census Operations

Administrative records are data collected as a by-product of the management of federal, state, or local governmental programs. Key examples for census applications include tax records, welfare records, building permit records, Medicare data, birth and death records, and data on immigration and emigration. Administrative records have a number of potential applications in the decennial census. These applications can be separated into those in which administrative records data are used indirectly and those in which administrative records data are used directly as decennial census data. Applications in which administrative records data are used indirectly include:

- *for improvement of the Master Address File (MAF):* addresses found in a merged administrative records file that were not on the MAF could be visited for field validation.
- *to validate edit protocols*:[4] edit protocols that were used to make decisions about inconsistent information in responses could be based on (or evaluated using) administrative records. For example, a 22-year-old listed as living with his parents and in a prison could have his enumeration moved to the prison address through information found in administrative records.

[4] An edit protocol is an automated rule that either generates an imputed response or changes a collected response based on the values of other responses.

- *for coverage improvement:* for households or individuals found on possibly more than one administrative list who were not enumerated in the census, fieldwork could be instigated at the indicated address; furthermore, addresses identified as being vacant could be checked to see if that assessment agrees with information in administrative records.
- *for coverage measurement and coverage evaluation:* consistent with A.6 in Appendix A, administrative records could be used to improve the information collected in postenumeration survey interviews[5]; furthermore, administrative records could be used to allocate demographic analysis estimates[6] to subnational regions;
- *to help target households for various purposes* (see below).
- *for duplicate search:* administrative records could be used to determine whether two records that have been matched actually represent the same person or to determine where the correct census residence is without resorting to fieldwork.[7]

Applications in which administrative records data are either used directly in the decennial census or in assessing coverage include:

- *as an alternative to last-resort proxy response:* instead of asking a neighbor or landlord for information in situations in which a respondent is not located after six attempts, if information is available from administrative lists, that information could be used for the enumeration.
- *as an alternative to item and unit imputation:* in the situations in which the Census Bureau uses either item or unit imputation (see National Research Council, 2004a, for a discussion of when unit imputation was used in the 2000 census), information from administrative records could be used as input to the imputation.
- *as a means for coverage evaluation*: whereby a person that appears on two or three administrative lists and not in the census is proof of a census omission.

In each of these applications, there could potentially be important benefits for the 2020 census, either in reducing field costs or in improving the quality of census data. We justify our optimism about the potential for applying administrative records to improve the above census component operations, and therefore the need to test those applications in the 2010 census, given the following considerations. First, there is clearly much useful information contained in various administrative records. The nonsurvey nature of the data collection gives a real chance of being able to provide useful information on hard-to-

[5]A postenumeration survey is a survey taken after the census is concluded that is used to measure coverage errors.

[6] Demographic analysis is an accounting scheme, roughly births plus immigrants minus deaths minus emigrants, for estimating the size of national demographic groups.

[7] An evaluation of A.C.E. Revision II estimates of duplication in Census 2000 using administrative records information demonstrated the potential for use of this information (for details, see Mulry et al., 2007). Administrative records might be used to confirm whether enumerations that are linked by computerized search are the same persons when fieldwork was unable to provide confirmation.

count individuals. This advantage probably motivated the Census Bureau to attempt to use information from administrative records for coverage improvement, as in 1980 with the Non-Household Sources Check, and in 1990 with the Parolees and Probationers Check. Also, the Census Bureau will be using administrative records to generate some of the coverage follow-up interviews in 2010. On the other hand, there are also deficiencies in administrative records, including undercoverage of portions of the population. (See NRC, 1994: Chapter 5 for a discussion of the limitations of administrative records systems for census applications.) Some of the existing research has been on the use of administrative records as an alternative to taking a census, notably AREX 2000, which is not that useful in assessing the value of administrative records for census component operations. As mentioned previously, the population coverage for the more thorough of the schemes tested in AREX 2000 was between 96 and 102 percent relative to the Census 2000 counts for the five test site counties. However, the AREX 2000 and the census counted the same number of people at the housing unit level for only 51.1 percent of the households, and they counted within one person of the census for only 79.4 percent of the units.

However, the Census Bureau has made substantial progress on administrative records since then. For example, E-StARS[8], the Census Bureau's name for a merged and unduplicated list of individuals from several administrative lists, was used to explain 85 percent of the discrepancies between the Maryland Food Stamp Registry recipients and estimates from the Census Supplementary Survey in 2001 (the pilot American Community Survey).

Although there has been much progress in collecting a higher quality merged unduplicated list of individuals, there has been little research on the nine applications listed here, in which the objective is to use administrative records not as a surrogate census but to assist in carrying out specific component operations. The panel's optimism is based not only on the information contained in administrative records, but also on the recognition that some of the component operations, especially last-resort enumeration, are understandably error-prone or are expensive (e.g., the coverage follow-up interview). Given that, administrative records do not have to be flawless to potentially provide a benefit. In addition, looking toward 2020, the quality of administrative records has been steadily improving over time. E-StARS, the Census Bureau's merged list of unique administrative records for individuals and housing units, has about the right number of people. Also, the economic directorate of the Census Bureau has been using information from administrative records directly in establishment surveys for a long time. So there is reason for optimism that some of the applications listed could be substantially improved through the use of administrative records.

---

[8] E-StARS is a nationwide multi- purpose research database, which combines administrative records from a variety of federal and state government sources and commercial databases with micro-data modeling to produce statistics for housing units and individuals that are comparable to decennial census results.

It is therefore important to determine, through either experiments or evaluations, which of the above (and other) applications of administrative records are most likely to be beneficial in the 2020 census, what needs to be done to implement such techniques nationally, and what the risks and benefits are. The basic idea would be to select several counties, merge and unduplicate all the relevant lists that can be collected for both individuals and addresses in those areas, and use the information from the merged file for some of the above purposes in comparison with the current census processes. In some cases, field verification would be needed to produce metrics for comparison—which is the main reason why this might fall into the experimentation rather than the evaluation category. However, in many cases much could be discovered without additional field data collection. Clearly, a census context is extremely helpful or essential for some of the above applications, such as for duplicate search. An additional complication is that administrative records are improving in quality year by year, and therefore any experiment or evaluation should take this possibility into account. (This suggestion is closely related to items C.2 and C.6 on the Census Bureau's list of issues.)

A particular means by which administrative records could be used to reduce field costs, at the price of possibly only a negligible reduction in data quality, is targeting. Targeting is the application of a census procedure to only a subset of the population. This subset of the population is selected through use of an algorithm that attempts to differentiate between people or households that are and are not likely to benefit from the application of the procedure. This algorithm is often supported by some external data source, and, in particular, administrative records should be studied as potentially playing this role. Administrative records offer opportunities to increase the scope and effectiveness of targeting, and in particular they may have important advantages for enumerating hard-to-count populations. (In a sense, the Census Bureau already uses targeting in several respects, including targeting of the advertising campaign, targeting areas for placement of "Be Counted" forms, and targeting areas for so-called blitz enumeration techniques.)

Of course, any time one does not use a census enumeration process on some areas that is used elsewhere, some of the omitted areas may have slightly poorer quality data as a result. So, for example, if a block canvass is not used in a particular block, there is a chance that new housing units there will be missed and that the area will receive a lower count as a result. (It should be noted that the Census Bureau has previously considered targeting for use with block canvassing, but to this point it has rejected this idea.) However, if properly planned and implemented, targeting should increase overall census data accuracy and at the same time reduce costs. This is because, if the targeting is effective, the reduction in data quality due to the selective omission of a census process is likely to be very slight. The resources saved through the use of targeting can then be used in other ways to improve the overall census data quality. Furthermore, sometimes resources are already constrained, and for those situations the question may not be whether to use targeting, but how best to use it. Also, through use of an algorithm, there is no intentional bias against any given area. (It may also be worth mentioning that some suggest that targeting can be perceived as uncomfortably close to sampling for the count. This is clearly an incorrect perception; it is merely the allocation of scarce resources to those cases most likely to benefit from this additional effort at enumeration.)

Clearly, further research (either experimentation or evaluation) is needed before targeting can be used in the decennial census. Given the promise of targeting, the panel thinks that the Census Bureau should prioritize either experimentation or evaluations that assess the promise of various forms of targeting and therefore retain sufficient data to ensure that such evaluations can be carried out. (Targeting is included in items C.3 and E.2 on the Census Bureau's list.) Creation of a Master Trace Sample, discussed in Chapter 3, is likely to satisfy this data need.

RECOMMENDATION 2: The Census Bureau should develop an experiment (or evaluation) that assesses the utility of administrative records for assistance in specific census component processes—for example, for improvement of the Master Address File, for nonresponse follow-up, for assessment of duplicate status, and for coverage improvement. In addition, either as an experiment or through evaluations, the Census Bureau should collect sufficient data to support assessment of the degree to which targeting various census processes, using administrative records, could reduce census costs or improve census quality.

Alternative Questionnaire Experiment

The 1980, 1990, and 2000 censuses have all involved some type of alternative questionnaire experiment in the associated research programs. The reason is straightforward: anything that can be done to increase response to questionnaires when they are sent out will necessarily decrease the amount of work that must be done by enumerators in the field in following up with nonrespondents. Also, to the extent that the initial questionnaire can be made clear, the quality of the collected data should improve. It is therefore of high priority that an alternative questionnaire experiment should be employed in the 2010 CPEX.

The Panel on Residence Rules in the Decennial Census (National Research Council, 2006: Finding 8.2) observed that "the Census Bureau often relies on small numbers (20 or less) of cognitive interviews or very large field tests (tens or hundreds of thousands of households, in omnibus census operational tests) to reach conclusions about the effectiveness of changes in census enumeration procedures." That panel argued for the development of more mid-range, smaller scale tests. We concur; there are numerous questionnaire design issues for which smaller scale tests would be a preferable vehicle compared with a formal census experiment. In thinking about an alternative questionnaire experiment or experiments for the 2010 census, the question is: Which sets of possible changes to the census questionnaire most need (or would most benefit) from being conducted in the census environment?

Race/Ethnicity as a Single Question

On page 1 of the short-form-only questionnaire planned for use in the 2008 census dress rehearsal (see Figure 2-1), the two questions on race and Hispanic origin (questions 8 and

9) take up half of the second column and about 40 percent of the respondent-fillable space on the page. Likewise, the race and Hispanic origin questions take up about half of the space allotted to collect information on persons 2 through 6 in a household (the block for Person 2 is shown in Figure 2-2). In the short-form-only census planned for 2010, then, the largest share of the questionnaire is given to the questions on race and Hispanic origin; therefore, if a viable alternative exists, a major focus of a questionnaire experiment in the 2010 census should be one that focuses on the two questions on race and ethnicity, since the rate of response is typically associated with the perceived ease of compliance.

Information on race is currently requested on the census questionnaire in response to the needs of the Voting Rights Act of 1965. In 1997, the Office of Management and Budget (OMB) developed standards for racial and ethnic classification to be used in the 2000 census, which resulted in 63 possible responses to account for multiple race identification. These standards will continue to apply to the 2010 census. Ethnicity, defined as either "of Hispanic origin" or "not of Hispanic origin," was requested on a separate question in the 2000 census, resulting in 126 total race/ethnicity response categories.

Evaluations have shown that the race/ethnicity questions used in 2000 (and in previous censuses) were associated with substantial confusion of race and ethnicity, often resulting in nonresponse, in some (seemingly) contradictory responses to the decennial census questions, and in high frequencies of response of "some other race" for Hispanic respondents (see, e.g., Census 2000 Topic Report #9, Race and Ethnicity in Census 2000, Census 2000 Testing, Experimentation, and Evaluation Program). Over the past 20 years, the Census Bureau has devoted considerable research to testing various approaches to the design of questions on race and ethnicity, trying alternative question wordings, formatting, and sequencing to elicit quality information (see, e.g., Rodriguez, 1994; McKay and de la Puente, 1995; de la Puente and McKay, 1995).

The Census Bureau has included race/ethnicity as one of their 11 topic groups for possible experimentation or evaluation in 2010. However, the Bureau gives low priority to the issue of developing a combined race and ethnicity question (listed as item B.2 in Appendix A). We disagree with that assessment; race and ethnicity are not really separate notions for many respondents, and the confusion resulting from the use of separate questions might be substantially reduced through the use of a single race/ethnicity question. This notion has been previously tested by the Census Bureau (1997) with generally positive results. Furthermore, the tendency to report "some other race" rather than Hispanic is likely to be reduced through the use of a single question.

The current race and ethnicity questions provide a number of examples of specific groups, including Filipino, Guamanian, or Samoan for race, and Puerto Rican and Cuban for ethnicity. There is no legal obligation stemming from the Voting Rights Act for the census questionnaire to include the mention of these various specific groups on the census short form. The argument in favor of including as many groups as the form will support is that this may increase response given personal feelings of affiliation with very

specific groups. Also, some argue that use of a streamlined questionnaire—that is, one that does not mention these individual groups—will increase the frequency of the mistaken response of "some other race." However, we suspect that the response of "some other race" is much more a function of the separation of race and ethnicity into two questions. Furthermore, we think that the inclusion of the specific groups makes the entire census questionnaire appear more complex, which may lower the response rate. We acknowledge that there is great interest in the relative size of these numerically smaller race and ethnic groups for states and counties, but that information will now be available on the American Community Survey.

We therefore think that the Census Bureau should include, as an experiment, the use of a single question on race and ethnicity. In addition, a streamlined version of this should also be tested, in which the only groups listed are (1) white, (2) black, (3) American Indian or Alaskan Native, (4) Asian, (5) Native Hawaiian or Pacific Islander, and (6) Hispanic, allowing for multiple responses in all of these categories.[9] We think that this is a productive avenue for testing because of its potential improvement regarding data quality. However, progress will be difficult, since the best approach to collecting higher quality data without discouraging respondents is not obvious. Continued experimentation is therefore imperative.

Finally, in addition to the test of a single race/ethnicity question, in-depth follow-up of a small sample of individuals who provide inconsistent responses to the 2010 questions should be planned.[10] Without understanding respondent behavior induced by a given question wording, it is very difficult to come up with hypotheses about how to improve that wording. Therefore, it would be useful to contact 50 or 100 such individuals and through face-to-face interviews determine why they responded the way that they did.

Representation of Residence Concepts

In terms of physical space on the page, the items on race and ethnicity take up the greatest area due to the number of responses permitted. However, the largest single presentation of a question has been Question 1 on recent censuses: the count of residents at the household.

The 2010 census will follow the basic concept laid out in the law authorizing the first census in 1790 of counting people at their "usual place of abode" (1 Stat. 105). Over time, this concept has evolved into one of counting people at their usual residence; this is distinct from counting them at their current residence or the location where they are when reached by the census. The Census Bureau has developed sets of residence rules to

[9] It should be noted that this specific question format runs counter to a provision included in the fiscal year 2005 omnibus appropriations bill (and that was made binding on subsequent years), which requires the Bureau to include a "some other race" option.

[10] Inconsistency is by necessity apparent since the responses for children with parents of different races or ethnicities may not be clear and, more importantly, since race and ethnicity responses are a matter of self-identification that does not need to be consistent. Apparently inconsistent responses include respondents who check a category indicating that they consider themselves to belong to a specific Hispanic group but at the same time also responding that they are not of Hispanic, Latino, or Spanish origin.

determine how to handle cases in which residential location may be ambiguous. Since the switch to reliance on the mail for most census data collection, the phrasing of Question 1 and the instructions that accompany it have been continually revised in order to guide census respondents to reporting their own residential situation in a way that is consistent with the Census Bureau's residence rules.

The National Research Council report *Once, Only Once, and in the Right Place: Residence Rules in the Decennial Census* (2006) comprehensively reviewed census residence rules past and present, assessing their adequacy in light of societal changes that can complicate clear definition of residence. These changes include the growth of both "sunbird" and "snowbird" populations that move to different areas based on seasonal weather changes, the changing nature of family structures (including children in shared custody arrangements), and the emergence of assisted living facilities for the elderly. T he 2006 report also considered long-standing historical challenges to accurate residence measurement, particularly concerning the large share of the nonhousehold (or group quarters) population living in places like college dormitories and correctional facilities.

Based on its review, the study panel suggested additional areas of research. Primary among these was a call to collect "any residence elsewhere" information: allowing respondents to specify a specific street address for another location at which they consider themselves a resident, as well as a follow-up question about whether the respondent considers this other location to be their "usual residence" (National Research Council, 2006:Rec. 6.2). That panel specifically suggested that "any residence elsewhere" be asked of the general household population in a 2010 census experiment and that the resulting data be comprehensively reviewed in an evaluation report (National Research Council, 2006:Recs. 6.5, 8.4). It also suggested that the "any residence elsewhere" question be asked of all group quarters respondents in 2010 (National Research Council, 2006:Sec. 7-D); a similar "usual home elsewhere" question was asked on all group quarters questionnaires in 2000, but they were processed and considered valid only for particular group quarters types.

A major reason for the importance of collection of "any residence elsewhere" information on a test basis for the general population is to help resolve a major outstanding concern about the transition from the traditional census long form to the ongoing American Community Survey. While the decennial census uses a "usual residence" concept, the ACS uses something closer to a "current residence" rule; specifically, residence in the ACS is defined using a "two-month rule" relative to the time of interview (see National Research Council, 2006:Box 8-2 and Sec. 8-C for extended discussion). The differences in census and ACS estimates that may be attributed to their differing residence standards is as yet unknown and is a concern on which solid data are critically important. To that end, National Research Council (2006:Rec. 8.3) suggested the twofold approach of testing the "any residence elsewhere" question in the 2010 census and testing a "usual residence"-type question on the ACS questionnaire as a separate ACS research activity.

In addition to the "any residence elsewhere" query, National Research Council (2006:Rec. 6.5) suggested that additional methods for presenting residence rules and

concepts be included in a 2010 alternative questionnaire experiment. In particular, the panel suggested a shift away from the model of lengthy instructions before Question 1 and instead breaking the resident question into smaller, easier-to-parse questions. This work could build on alternative questionnaire presentations that the Census Bureau tested on a limited basis in its 2005 National Census Test and an ad hoc test in 2006. To be clear—and as is noted elsewhere in this report—National Research Council (2006) argued that the Census Bureau often relies too much on both very small and very large tests, and that some residence-related questions (e.g., specific cues to include on questionnaires or alternative means of developing rosters of household members) may be better handled by other testing means. However, the importance of Question 1, the potential gain in data accuracy, and the potential reduction in the need to dispatch an enumerator to conduct a coverage follow-up interview that could stem from even small changes on the question form all argue strongly for a residence component of a 2010 alternative questionnaire experiment.

Other Content Issues

Other content issues on the 2010 census form are also worth examining and might benefit from an experiment in 2010. The hope is that these various questionnaire wording issues could be folded in with an experiment on race and ethnicity, residence rules, or both. There may be too many issues for a single experiment and therefore there may be a need to further prioritize these issues before finalizing an alternative questionnaire experiment.

- *Coverage probes.* Two coverage probes will be included on the 2010 census questionnaire for the first time. These are: (1) "Were there any additional people staying here April 1, 2010 that you did not include in Question 1?" and (2) "Does Person X sometimes live or stay somewhere else?" This is followed by a listing of situations that are sometimes reported in error. As implemented in 2010, this set of probes is primarily intended as a trigger for inclusion in the coverage follow-up operation, described below. The probes also serve to jog a respondent's memory and prompt them to reevaluate their answer to the household resident count in Question 1 of the census form. It is worth considering whether more specific or differently worded probes are more effective at accomplishing either of these tasks, and whether they can be structured to provide auxiliary information that could be useful in editing census responses. For instance, a more detailed query about whether the respondent is at (or may be counted at) a seasonal residence, or a focused question on the residence of college-enrolled children, may prove to have advantages over the approach planned for 2010.
- *Motivation of respondents.* The 2006 Canadian census questionnaire added brief descriptive statements at key places in order to anticipate respondents' concern about a question's justification in the census. By including these, Statistics Canada thinks that it has achieved some benefits in building respondent motivation to answer questions on the census form. For example, the 2006 census long-form questions on race and ancestry—which, in Canada, are not part of the short-form questions asked of everybody—are prefaced with the explanation:

> The census has collected information on the ancestral origins of the population for over 100 years to capture the composition of Canada's diverse population.

The specific race question includes the reminder that this information is collected to support programs that promote equal opportunity for everyone to share in the social, cultural, and economic life of Canada.

The last page of the Canadian short-form questionnaire includes a paragraph-length section labeled "Reasons Why We Ask the Questions," noting, for example, that "Question 7 on languages is asked to implement programs that protect the rights of Canadians under the Canadian Charter of Rights and Freedoms. It also helps determine the need for language training and services in English and French." It could be useful to measure the impact on the quality of response that would result from various attempts to represent similar motivational messages on the U.S. census form.

- *Group quarters.* Given that some types of group quarters' residences are subject to a high rate of duplication, in particular those in college dormitories (see Mule, 2002), it might be useful to evaluate the benefits of a "usual home elsewhere" question on the census questionnaire for all types of group quarters residences. (This is consistent with Recommendation 6.2 and Section 7-C in National Research Council, 2006.) This might facilitate real-time identification of census duplicates between residents of group quarters and residents of nongroup quarters.

Finally, item G.1 on the Census Bureau's list of research topics proposes administering the 2000 census questionnaire to a group of 2010 census respondents so that some insight can be drawn about the effectiveness of the complete bundle of changes between the 2010 and 2000 forms. This proposal to use the prior census questionnaire as a control group treatment has not always been carried out in past alternative questionnaire experiments. Implementing it is consistent with guidance from the previous National Research Council report (2006:Recommendation 6.8), and we concur that it should be done as part of a 2010 alternative questionnaire experiment.

Deadline Messaging and Other Presentation Issues

Deadline messaging includes a variety of ways of notifying the respondent on mailing materials that in order to be accepted the enclosed questionnaire has to be returned by a given date. By a compressed mailing schedule is meant that, instead of the approach used in the 2000 census, in which the questionnaire was mailed two weeks before Census Day, the households will receive the census questionnaire just a few days before Census Day. In the 2006 decennial short-form experiment,[11] the use of deadline messaging, in conjunction with a compressed mailing schedule, resulted in a higher mail response rate (Martin, 2007). The deadline message was placed on the advance letter informing the

[11] The decennial short-form experiment evaluated several potential improvements to the census mail form. These included a revised instruction about whom to list as Person 1, a series of questions to reduce and identify coverage errors, and a deadline for return of the form.

household of the upcoming appearance of the census questionnaire, on the envelope of the initial mailed questionnaire, on the initial questionnaire cover letter, and on the reminder postcard. However, the 2006 test could not determine whether the increased response was due to a specific form of the deadline message or whether it was due to the compressed mailing schedule. Therefore, some further work attempting to determine the specific cause of the increase in response would be extremely useful. More importantly, since increasing the initial response rates decreases the nonresponse follow-up fieldwork, which reduces census costs, this is important to investigate further. Additional research on the effectiveness of different dates for both the initial mailing of the census questionnaires and the mailing of the replacement questionnaires would also be useful to undertake. Item H.1 on the Census Bureau's list argues that looking at this issue in a census environment is important, and the panel agrees, since response to mail materials differs in a census in comparison to either a test census or a survey environment.

We have described a number of issues that relate to the content and the presentation of the census questionnaire, including race and ethnicity, residence rules, coverage probes, providing a motivation for the cooperation of respondents, collection of alternate address data for residents of group quarters, and deadline messaging. It may be that several of these issues can be jointly addressed in a single experiment by including these issues as separate factors in the experiment. One straightforward way of accomplishing this, which is much more cost-effective with respect to the burden on respondents, is through the use of a fractional factorial design, assuming that some of the higher level interactions between these factors are negligible (see Box and Hunter, 1961).

RECOMMENDATION 3: The Census Bureau should include one or more alternate questionnaire experiments during the 2010 census to examine:

- **the representation of questions on race and ethnicity on the census questionnaire, particularly asking about race and Hispanic origin as a single question;**
- **the representation of residence rules and concepts on the census questionnaire; and**
- **the usefulness of including new or improved questions or other information on the questionnaire with regard to (1) coverage probes, (2) the motivation of census questions, (3) the request of information on usual home elsewhere on group quarters questionnaires, and (4) deadline messaging and mailing dates for questionnaires.**

In such experiments, both the 2000 and the 2010 census questionnaires should be included in the assessments. The Census Bureau should explore the possibility of joining the recommended experiments listed above into a single experiment, through use of fractional factorial experimental designs.

A Possible Additional Experiment: Comparison of Telephone to Personal Interview for Coverage Follow-Up Interview

The current plans are to carry out a coverage follow-up interview in 2010 to collect additional information for six situations for which the number of residents is unclear based on the responses to the initial questionnaire (see Box 2-1). Since a large fraction (probably more than 20 percent) of U.S. households may satisfy one or more of these six situations, the costs of the resulting coverage follow-up interviews could be prohibitive. To reduce these costs, the Census Bureau is planning to follow up these households by telephone only (and therefore only for those households that provide a contact telephone number on the census questionnaire).

This specific implementation of the coverage follow-up interview raises some concerns about the quality of the information received. First, we are concerned that the households that would most benefit from this follow-up will be those not likely to provide valid telephone numbers and consequently will be missed. For example, some of those that are harder to enumerate may make use of prepaid cell phones. Therefore, it would be useful to determine whether other wordings of the request for phone numbers would increase the response to this item. (This relates to the earlier issue of providing motivation for questions on the short form. This suggestion is related to items C.8, C.7, F.1, and F.2 on the Census Bureau's list of issues.)

Another concern stems from the fact that the coverage follow-up interview uses question wording similar to that on the census questionnaire, and there is thus a good chance of generating the same response as was initially received in the case of interviews resulting from coverage probes or from the identification of potential duplicates. One alternative to address this concern that might be worth examining is whether there is a way of communicating to the respondent the circumstances that generated the interview through a series of probes. A second way of addressing this concern is that higher quality answers, possibly using such probes, might be produced through use of a face-to-face interview, rather than a phone interview. While this would clearly be more expensive, knowing the impact on quality would be useful in designing the analogous data collection in 2020. Also, there are ways of reducing field interview costs to permit more face-to-face interviewing. For example, the targeting of households through the use of administrative records might reduce the workload to a manageable level, allowing for face-to-face interviews of selected households.

If the decision is made not to include study of the coverage follow-up interview in a census experiment, the above concerns strongly argue for retention of all relevant information to be able to evaluate this process after the census is completed.

CONCLUSION

These are the panel's suggestions for experiments to be carried out during the 2010 census. We look forward to assisting the Census Bureau in fleshing out more specific study plans for the ideas that are ultimately selected for experimentation in the coming months.

We also think that the Census Bureau needs to increase its in-house expertise in experimental design regarding census experimentation. The panel has seen evidence in the past that some experiments, in both censuses and test censuses, have not been fully consistent with accepted principles of experimental design. This includes the use of preliminary assessments of which factors might affect a response of interest, the use of controls and blocking for meaningful comparisons (see, e.g., National Research Council, 2006: Rec. 6.8), and the simultaneous varying of test factors (including use of orthogonal designs, factorial designs, and fractional factorial designs) for greater effectiveness of test panels. Also, often not enough attention is paid in advance to the statistical power of tests. Certainly some of this can be attributed to the fact that the primary function of a census or a census test is an opportunity to assess the full census operation with the embedded experiments having to make do with various limitations. However, it is important for the Census Bureau to improve its application of experimental design techniques for its experiments, both to reduce the costs of the experimentation and to increase the information contained in the results.

BOX 2-1
Situations Generating a Coverage Follow-up Interview

1. Households with discrepancies between the household counts and the number of individuals for which information is provided

2. Households with more than six residents (which will therefore not fit on the census questionnaire

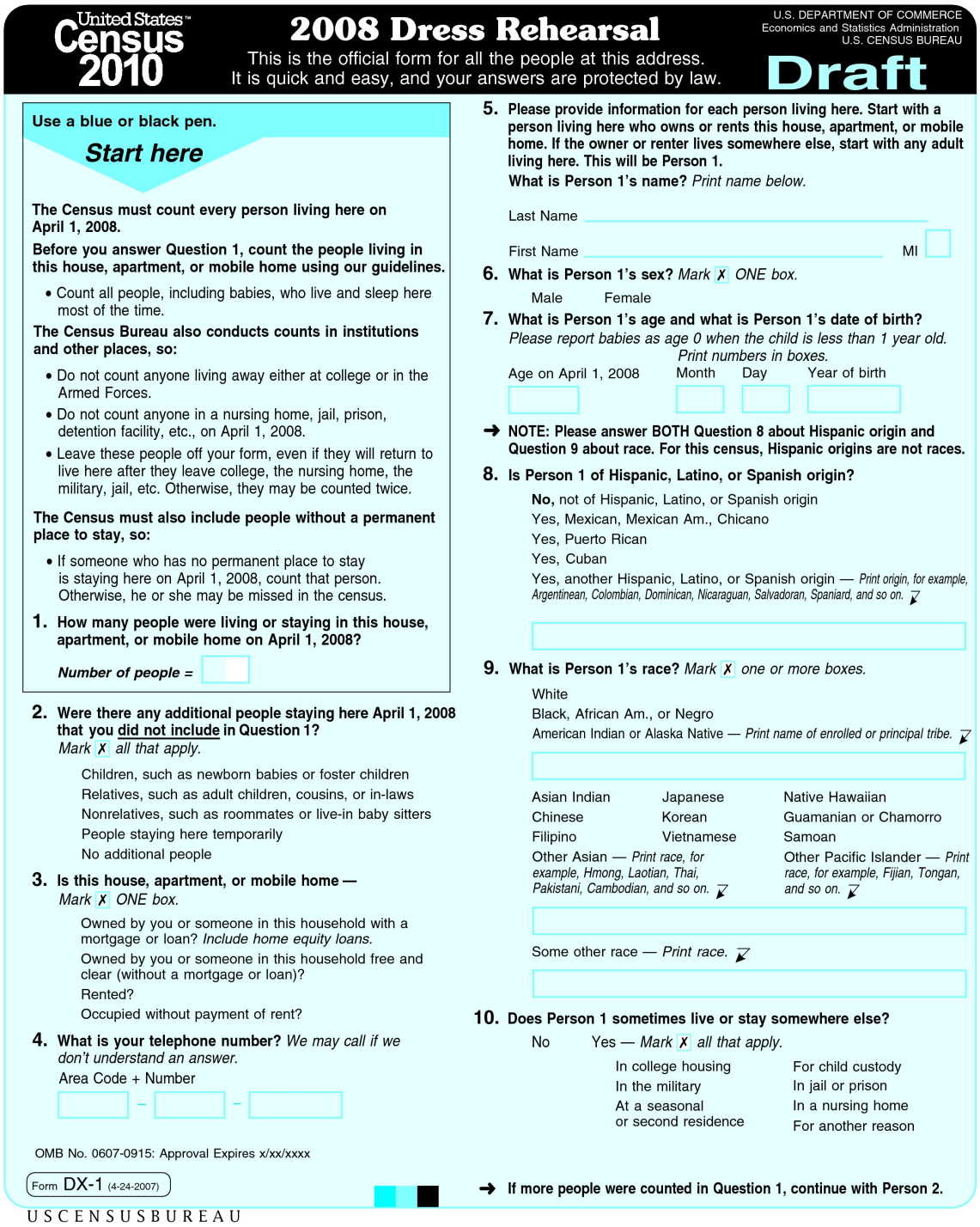
3. Households that indicate on the census questionnaire other households in which the residents might also have been enumerated

4. Households that indicate other people not included in response that sometimes live there

5. Households that are identified as having individuals that might have been duplicated in the census through use of a national computer search for duplicates

6. Households that may have not been correctly enumerated given information from administrative records.

SOURCE: Adapted from information from U.S. Census Bureau; see also National Research Council (2006:Box 6-3).



United States Census 2010

2008 Dress Rehearsal

This is the official form for all the people at this address.
It is quick and easy, and your answers are protected by law.

U.S. DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. CENSUS BUREAU

Draft

Use a blue or black pen.

Start here

The Census must count every person living here on April 1, 2008.

Before you answer Question 1, count the people living in this house, apartment, or mobile home using our guidelines.

- Count all people, including babies, who live and sleep here most of the time.

The Census Bureau also conducts counts in institutions and other places, so:

- Do not count anyone living away either at college or in the Armed Forces.
- Do not count anyone in a nursing home, jail, prison, detention facility, etc., on April 1, 2008.
- Leave these people off your form, even if they will return to live here after they leave college, the nursing home, the military, jail, etc. Otherwise, they may be counted twice.

The Census must also include people without a permanent place to stay, so:

- If someone who has no permanent place to stay is staying here on April 1, 2008, count that person. Otherwise, he or she may be missed in the census.

1. How many people were living or staying in this house, apartment, or mobile home on April 1, 2008?

Number of people =

2. Were there any additional people staying here April 1, 2008 that you did not include in Question 1?
Mark ✗ all that apply.

- Children, such as newborn babies or foster children
- Relatives, such as adult children, cousins, or in-laws
- Nonrelatives, such as roommates or live-in baby sitters
- People staying here temporarily
- No additional people

3. Is this house, apartment, or mobile home —
Mark ✗ ONE box.

- Owned by you or someone in this household with a mortgage or loan? *Include home equity loans.*
- Owned by you or someone in this household free and clear (without a mortgage or loan)?
- Rented?
- Occupied without payment of rent?

4. What is your telephone number? *We may call if we don't understand an answer.*

Area Code + Number

5. Please provide information for each person living here. Start with a person living here who owns or rents this house, apartment, or mobile home. If the owner or renter lives somewhere else, start with any adult living here. This will be Person 1.

What is Person 1's name? *Print name below.*

Last Name

First Name MI

6. What is Person 1's sex? *Mark ✗ ONE box.*

Male Female

7. What is Person 1's age and what is Person 1's date of birth?
Please report babies as age 0 when the child is less than 1 year old.
Print numbers in boxes.

Age on April 1, 2008 | Month | Day | Year of birth

➜ **NOTE: Please answer BOTH Question 8 about Hispanic origin and Question 9 about race. For this census, Hispanic origins are not races.**

8. Is Person 1 of Hispanic, Latino, or Spanish origin?

- **No,** not of Hispanic, Latino, or Spanish origin
- Yes, Mexican, Mexican Am., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, another Hispanic, Latino, or Spanish origin — *Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.*

9. What is Person 1's race? *Mark ✗ one or more boxes.*

- White
- Black, African Am., or Negro
- American Indian or Alaska Native — *Print name of enrolled or principal tribe.*

- Asian Indian
- Chinese
- Filipino
- Japanese
- Korean
- Vietnamese
- Other Asian — *Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.*
- Native Hawaiian
- Guamanian or Chamorro
- Samoan
- Other Pacific Islander — *Print race, for example, Fijian, Tongan, and so on.*
- Some other race — *Print race.*

10. Does Person 1 sometimes live or stay somewhere else?

No Yes — *Mark ✗ all that apply.*

- In college housing
- In the military
- At a seasonal or second residence
- For child custody
- In jail or prison
- In a nursing home
- For another reason

OMB No. 0607-0915: Approval Expires x/xx/xxxx

Form DX-1 (4-24-2007)

➜ **If more people were counted in Question 1, continue with Person 2.**

U S C E N S U S B U R E A U

Figure 2-1 First page (Person 1), draft 2008 dress rehearsal questionnaire.

SOURCE: http://www.census.gov/Press-Release/www/2007/questionnaire_4_24_07.pdf.

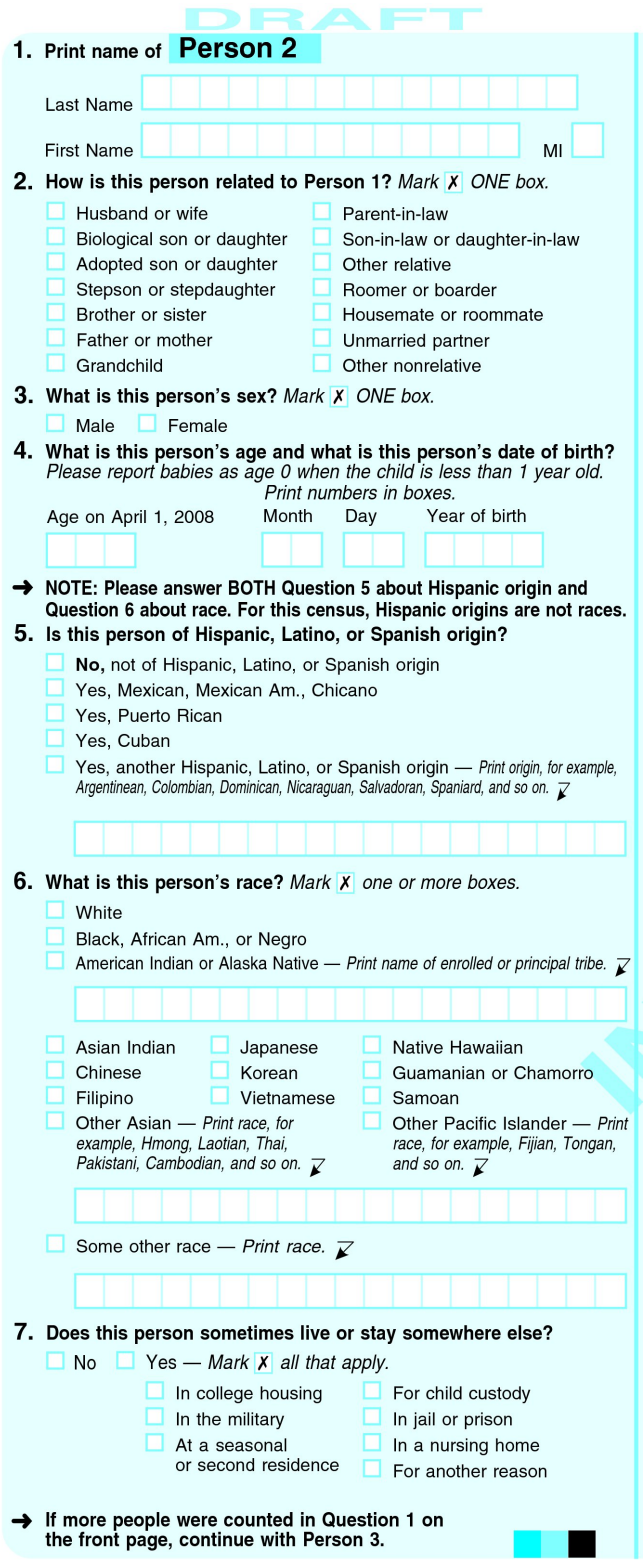


Figure 2-2 Person 2 panel, draft 2008 dress rehearsal questionnaire.

SOURCE: http://www.census.gov/Press-Release/www/2007/questionnaire_4_24_07.pdf.

3

Initial Views on 2010 Census Evaluations

SUGGESTIONS FOR THE 2010 CENSUS EVALUATIONS

The panel's first priority is to provide input to the selection of experiments to be implemented in 2010, since the design of these experiments needs to begin very soon to allow for the development of associated materials and protocols. In addition, the panel has some suggestions relative to the evaluations to be carried out in conjunction with the 2010 census. There is also a time pressure for them since, as stated previously, much of the data collection in support of the 2010 census evaluations needs to be specified relatively early, in particular so that the contractors involved in many of the census processes can make plans for the collection and structuring of data extracts that relate to the functioning of those processes.

Address List Improvement

For the 2000 census, the Census Bureau departed from past practice of building the address list for the census from scratch. Instead, it pursued a strategy of building a Master Address File (MAF), using the 1990 address list as a base and seeking ways to "refresh" the database during the intercensal period. Legislation enacted in 1994 created two major tools for address list improvement. First, the new law authorized the Census Bureau to use the U.S. Postal Service's Delivery Sequence File (DSF; as the name suggests, a master list of mail delivery addresses and locations used to plan postal routes) as an input source. Second, it permitted limited sharing of extracts of the Master Address File (which is confidential information under Title 13 of the U.S. Code) with local and tribal governments. Specifically, this provision led to the creation of the Local Update of Census Addresses (LUCA) program, first conducted in several phases in 1998 and 1999 (see National Research Council, 2004a:62-65).

The Master Address File used to support the American Community Survey during the intercensal period is essentially an update of the 2000 census MAF, revised to include edits to the Postal Service's Delivery Sequence File and new construction. Through these actions, the MAF, heading into the 2010 census, will be certainly more than 90 percent complete but probably not 99 percent complete. (There will almost certainly be a substantial amount of duplication as well.)

The Census Bureau will utilize two operations to increase the degree of completeness of the MAF from its status in 2008 in preparation for its use in the decennial census in 2010. First, it will again use the LUCA program, in which local governments will be asked to review preliminary versions of the MAF for completeness and to provide addresses that may have been missed (or added in error). However, even granting that LUCA will be improved over the 2000 version, it is likely that the participation will be uneven and that a substantial amount of

incompleteness will remain after these addresses are added to the MAF. In anticipation of that, the Census Bureau will carry out a national block canvass, visiting each census block, and adding any missed housing units to the MAF (while collecting information from global positioning systems for all housing units).

It may be the case that for many well-established blocks in the United States a 100 percent block canvass is wasteful, given that there is little possibility in these blocks of addition or deletion of housing units over time. It would be useful to identify such blocks in advance, since then the block canvass could be restricted to the subset of blocks in need of MAF updating (this is consistent with item C.3 in Appendix A). Given the costs of a 100 percent block canvass, identifying a targeting methodology that does an excellent job of discriminating between those blocks that are very stable over time and those blocks that are likely to have recent additions or deletions (or both) would provide substantial cost savings with possibly only a negligible increase in the number of omissions (or erroneous inclusions) in the MAF. It is likely that administrative records, especially building permit records, commercial geographic information systems, and the ACS could provide useful predictors in discriminating between stable and nonstable blocks. Such targeting is already used in the Canadian census; it uses an address register that is updated intercensally, and field verification is restricted to areas where building permit data indicate the presence of significant new construction (Swain et al., 1992).

To support the determination as to whether any targeting methods might satisfy this need—and, indeed, to facilitate a richer evaluation of MAF accuracy than was possible in 2000—the Census Bureau should ensure that the complete source code history of every MAF address is recoverable. In 2000, the MAF was not structured so that it was possible to fully track the procedural history of addresses—that is, which operations added, deleted, or modified the address at different points of time. Therefore, it was not possible to accurately determine the unique contributions of an operation like LUCA or the block canvass; nor was it possible to assess the degree to which various operations overlapped each other in listing the same addresses. Census Bureau staff ultimately derived an approximate "original source code" for MAF addresses, albeit with great difficulty; see National Research Council (2004b:146-147). Redesign of the MAF database structure was included in the plans to enhance MAF and TIGER during this decade; the Census Bureau should assess whether the new structure will adequately track the steps in construction of the 2010 (and future) MAF.

RECOMMENDATION 4: The Census Bureau should design its Master Address File so that the complete operational history—when list-building operations have added, deleted, modified, or simply replicated a particular address record—can be reconstructed. This information will support a comprehensive evaluation of the Local Update of Census Addresses and address canvassing. In addition, sufficient information should be retained, including relevant information from administrative records and the American Community Survey, to support evaluations of methods for targeting blocks that may not benefit from block canvassing. Finally, efforts should be made to obtain addresses from commercial mailing lists to determine whether they also might be able to reduce the need for block canvassing.

Master Trace Sample

The idea of creating a master trace sample, namely designating a sample of households in, say, census blocks, for which the full history of relevant census operations is retained in an accessible manner for subsequent analysis, is extremely important. In each decennial census, there are unanticipated problems that need to be fully understood in order to make modifications to the census design, to partially or completely eliminate their chance of occurring in the subsequent decennial census. A master trace sample provides an omnibus tool for investigating the source of any of a large variety of potential deficiencies that can arise in such a complicated undertaking as the decennial census. Otherwise, the Census Bureau is usually left with evaluation studies that, due to the limited information available, are often univariate or bivariate summaries that cannot inform about even relatively simple interactions between the individuals, the housing unit, and the enumeration techniques that resulted in a higher frequency of coverage (or content) errors.

The value of a master trace sample database or system has been advocated by several National Research Council panels, including the Panel on Decennial Census Methodology (National Research Council, 1985: Rec. 6.3), the second phase of the Panel on Decennial Census Methodology (National Research Council, 1988), the Panel on Alternative Census Methodologies (National Research Council, 1999:Rec. 5.1), and the Panel on Research on Future Census Methods (National Research Council, 2004a: Rec. 8.4, 8.5, 8.6, 8.7). The last cited report contains a useful history of the development of this idea and includes the following recommendation: "The Census Bureau should carry out its future development in this area of tracing all aspects of census operations with the ultimate aim of creating a Master Trace *System*, developing a capacity for real-time evaluation by linking census operational databases as currently done by the Master Trace Sample. Emerging 21st century technology should make it feasible to know almost instantaneously the status of various census activities and how they interact. Such a system should be seriously pursued by the Census Bureau, whether or not it can be attained by 2010 (or even by 2020)." Such a proposal is a straightforward generalization of item A.3 of the Census Bureau's list, though expanding from a focus on the coverage measurement survey to the full set of census operations.

Such a database could be used to evaluate many things, including determining what percentage of census omissions are in partially enumerated households and what percentage of omissions are found on the merged administrative records database. A master trace sample database would be extremely useful in addressing the needs described in the previous section, including understanding the source of duplicates in the Master Address File and evaluating the benefits of LUCA and the block canvass operation. An overall assessment of the workings of the coverage follow-up interview would be feasible if the master trace sample database collected sufficient data so that it was known for each housing unit in the CFU interview what triggered the CFU interview and what the result of the interview was—that is, what changes were made and what information precipitated the change. As indicated, inclusion of the merged administrative records file and relevant data from the American Community Survey in such a database would provide additional information at the individual and local area levels.

Creation of a master trace sample presents a number of challenges. First, there is the retention of the data from the census and affiliated activities. Some modest planning is needed here, especially given the necessity of collecting data from various contractors who are likely not to have planned in advance to provide for such data extracts. In addition, it is necessary to find an effective way of linking the information retained about the enumerators, the housing units, the residents, the census processes, the type of census coverage error made, and contextual information in a way that facilitates a broad range of potential analyses, especially those that examine interactions among these various aspects of the census process. Also, selecting the minimum data to be collected that is included in the master trace sample database is crucial to address early on. This is because while the addition of various sets of variables from different parts of the census and the census management information system provides broader capabilities for investigating various aspects of census-taking, the inclusion of each additional set of variables complicates the formation of the database. This is a hard database management problem, and the Census Bureau should enter into such a project with the recognition of the need for input of considerable expertise in database management to ensure success. (We think that the relative lack of use of the 2000 Master Trace Sample was due in part to its inability to facilitate many types of analysis.)

An additional concern is that the sampled blocks included have to be kept confidential so that the behavior in these blocks is representative of the entire census. Finally, we do not think the size of the master trace sample database is a major concern. A smaller but somewhat analogous database was constructed by the Census Bureau in 2000 and, as noted above, there have been substantial advances in computing memory and speed since then.

RECOMMENDATION 5: The Census Bureau should initiate efforts now for planning the general design of a master trace sample database and should plan for retention of the necessary information to support its creation.

Reverse Record Check

The Canadian Census has successfully employed a reverse record check for the last eight censuses to measure net coverage error. Briefly, four samples are collected: (1) a sample of enumerations from the previous census, (2) a sample of births in the intercensal period, (3) a sample of immigrants in the intercensal period, and (4) a sample of those missed in the previous census. The fourth sample is clearly the most difficult, but by matching those contained in the four samples for the previous reverse record check to the census to determine omissions and continuing this process over several censuses, a relatively useful sample of omissions can be formed over time. Once the four samples are formed, current addresses are determined, and the sample is matched to the census using name, addresses, and other characteristics. In a separate operation, the census is matched against itself to generate an estimate of the overcount, and, using both, an estimate of the net undercount is derived. Characteristics for both the omissions and overcounts support tabulations by age, sex, race, geography, etc.

To date, this procedure has not been used to evaluate the U.S. decennial census, mainly due to the 10-year period between censuses (as opposed to the 5 years between Canadian censuses), which complicates the need to trace people's addresses from one census to the next. This issue

was specifically examined in the Forward Trace Study (Mulry, 1986). However, with administrative records systems improving each year, and given the emergence of the American Community Survey, tracing people over a 10-year period is likely to be much more feasible now in comparison to 1984. Furthermore, a reverse record check has an important advantage over the use of a postenumeration survey with dual-systems estimation in that there is no need to rely on assumptions of independence or homogeneity to avoid correlation bias, a type of bias that occurs in estimating those missed by both the census and the postenumeration survey. There are also more opportunities for validating the reliability of the estimates provided. For example, a reverse record check provides an estimate of the death rate. The key issue concerning feasibility remains tracing, and a useful test of this would be to take the 2006-2007 ACS and match that forward to see how many addresses could be found over the 3.5-year period. In such a test, the ACS would serve as a surrogate for the sample from the previous census enumerations. Either relating this back to a sample of census enumerations and a sample of census omissions, or developing a sample of ACS omissions, remains to be worked out. But certainly, successful tracing of nearly 100 percent of the ACS would be an encouraging first step.

RECOMMENDATION 6: The Census Bureau, through the use of an experiment in the 2010 census (or an evaluation of the 2010 census) should determine the extent to which the American Community Survey could be used as a means for evaluating the coverage of the decennial census through use of a reverse record check.

Edit Protocols

Edit protocols are decisions about enumerations or the associated characteristics for a housing unit that are made based on information already collected, hence avoiding additional fieldwork. For example, an edit protocol might be that, when an individual between ages 18 and 21 is enumerated both away at college and at their parent's home, the enumeration at the parent's home is deleted. (Note that census residence rules are to enumerate college students where they are living the majority of the time, which is typically at the college residence.) This would avoid sending enumerators either to the parent's home or to the college residence, but it would occasionally make this decision in error. The Census Bureau has made widespread use of edit protocols in the past to deal with inconsistent data. For example, there are rules to deal with inconsistent ages and dates of birth. Furthermore, early in 2000, when it became apparent that the MAF had a large number of duplicate addresses, the Census Bureau developed an edit protocol to identify the final count for households with more than one submitted questionnaire (see Nash, 2000).

More generally, edit protocols might be useful in resolving duplicate residences, as well as in situations in which the household count does not equal the number of people who are listed as residents. Again, as with targeting, edit protocols avoid field costs but do have the potential of increased census error. However, given the increasing costs of the decennial census, understanding precisely what the trade-offs are for various potential edit protocols would give the Census Bureau a better idea of which of these ideas are more or less promising to use in the 2020 census. The panel therefore suggests that the Census Bureau prioritize evaluations that assess the promise of various forms of edit protocols and therefore retain sufficient data to ensure

that such evaluations can be carried out. Creation of a master trace sample is likely to satisfy this data need.

Coverage Assessment of Group Quarters

The census coverage measurement program in 2010 will not assess some aspects of the coverage error for individuals living in group quarters. Through use of a national match, as in the 2000 census evaluation, the Census Bureau will be able to estimate the number of duplicates both between those in the group quarters population and those in the nongroup quarters population and the number of duplicates entirely within the group quarters population (see Mule, 2002, for the rate of duplication for various types of group quarters in the 2000 census). However, the number of omissions for group quarters residents will not be measured in 2010, nor will the number of group quarters and their residents who are counted in the wrong place.

Given the variety of ways that group quarters are enumerated, and given the various types of group quarters, coverage evaluation methods will probably need to be tailored to the specific type. We are unclear about the best way to proceed, but it is crucial that the Census Bureau find a reliable way to measure the coverage error for this group, which has been unmeasured for two censuses, going on a third. It is likely that there are sources of information, which if retained, could be used to help evaluate various proposals for measuring coverage error for group quarters residents in 2020.

What is needed is that the list of residents as of Census Day for a sample of group quarters be retained, and for this sample to be drawn independently of the Census Bureau's list of group quarters. Creating such a list probably differs depending on the type of group quarters. One would take the list of residents as the ground truth, and determine whether the residents had been included in the census and at which location. These are ideas are very preliminary, and we hope to revisit this issue prior to issuing our final report. (This general topic was itemA-4 on the Census Bureau's list.)

RECOMMENDATION 7: The Census Bureau should collect sufficient data in 2010 to support the evaluation of potential methods for assessing the omission rate of group quarters residents and the rate of locating group quarters in the wrong census geography. This is a step toward the goal of improving the accuracy of group quarters data.

Training of Field Enumerators

The 2010 census will be the first in which handheld computing devices are used. They will be used in the national block canvass to collect information on addresses to improve the MAF, and they will also be used for nonresponse follow-up and for coverage follow-up. While the implementation of handheld computing devices was tested in the 2006 census test and will be tested further in the 2008 dress rehearsal, there remain concerns as to how successful training will be and whether some enumerators will find the devices too difficult to comfortably learn to use in the five days allotted to training. Given that it will be extremely likely that such devices

will again be used to collect information in 2020 (and in other household surveys intercensally), it would be useful to collect information on who quit, and why they quit, during the training for field enumeration work, who quit and why they quit during fieldwork, and the effectiveness of the remaining enumerators using the devices. In addition, any characteristics information that would be available from their employment applications should be retained as potential predictors for the above. Finally, the Census Bureau should undertake some exit interviews of those leaving training early and those quitting fieldwork early to determine whether their actions were due to discomfort with the handheld devices. This might provide some information either about training that would be useful in adjusting the training used in 2020, or about the ease of use of the devices or about hiring criteria. (This issue is consistent with item D.3 on the Census Bureau's list.)

A GENERAL APPROACH TO CENSUS EVALUATION

The panel also has some general advice on selecting and structuring census evaluations. As mentioned above, the evaluations in 2000 were not as useful as they could have been in providing detailed assessments as to the types of individuals, housing units, households, and areas for which various census processes performed more or less effectively. This is not to say that an assessment of general functioning is not important, since processes that experienced delays or other problems are certainly candidates for improvement. However, evaluations focused on general functioning do not usually provide as much help in pointing the way toward improving census processes as analyses for subdomains or analyses that examine the interactions of various factors. Since the costs of such analyses are modest, we strongly support the use of evaluations for this purpose. This issue was addressed in *The 2000 Census: Counting Under Adversity*, which makes the following recommendation, which this panel supports (National Research Council, 2004b:Rec. 9.2):

> The Census Bureau should materially strengthen the evaluation [including experimentation] component of the 2010 census, including the ongoing testing program for 2010. Plans for census evaluation studies should include clear articulation of each study's relevance to overall census goals and objectives; connections between research findings and operational decisions should be made clear. The evaluation studies must be less focused on documentation and accounting of processes and more on exploratory and confirmatory research while still clearly documenting data quality.
>
> To this end, the 2010 census evaluation program should:
>
> 1. identify important areas for evaluations (in terms of both 2010 census operations and 2020 census planning) 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. focus on analysis, including use of 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. consider ways to incorporate real-time evaluation during the conduct 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.

Item (3) is particularly important, in stressing the need for analysis, not just summaries of the (national) functioning of various census processes.

We think that evaluations should attempt to answer two types of questions. First, evaluations should be used to support or reject leading hypotheses about the effects on census costs or data quality of various census processes. Some of these hypotheses would be related to the list of topics and questions that were provided to the panel, but more quantitatively expressed. For example, such a hypothesis might be that bilingual questionnaire delivery will increase mail response rates in the areas in which it is currently provided in comparison with not using this technique. To address this question, assuming that targeting of mail questionnaires to all areas with a large primarily Spanish-speaking population is used, one might compare the mail response for areas just above the threshold that initiates this process to those just below. While certainly not as reliable or useful as a true experiment, analyses such as these could provide useful evidence for the assessment of various component processes without any impact on the functioning of the 2010 census.

Second, comprehensive data from the 2010 census, its management information systems, the 2010 census coverage measurement program, and contextual data from the American Community Survey and from administrative records need to be saved in an accessible form to support more exploratory analysis of census processes, including graphical displays. Each census surprises analysts with unforeseen problems, such as the large number of duplicate addresses in the 2000 census, and it is important to look for such unanticipated patterns so that their causes can be investigated. Standard exploratory models should be helpful in identifying these unanticipated patterns. Of course, any findings would need to be corroborated with additional testing and evaluation.

INITIAL CONSIDERATIONS REGARDING A GENERAL APPROACH TO CENSUS RESEARCH

The Census Bureau has a long and justifiably proud history of producing important research findings in areas relevant to decennial census methodology. However, the panel is concerned that in more recent times research has not played as important a role in census redesign as it has in the past. Furthermore, there is the related concern that research is not receiving the priority and support it needs to provide the results needed to help guide census redesign. We give four examples to explain this concern.

First, research in areas in which the results were relatively clear has been unnecessarily repeated. An example is the testing of the benefits from the use of a targeted replacement questionnaire, which was examined during the 1990s and also in 2003. The increased response resulting from the use of a targeted replacement questionnaire was relatively clear based on research carried out in the 1970s by Dillman (1978). In 1992 the Census Bureau carried out the Simplified Questionnaire Test (SQT), which examined the use of a blanket replacement questionnaire. Dillman et al. (1993) describe the Implementation Test (IT), also carried out in 1992, which attempted to determine the contribution of each part of the mailing strategy toward improving response. As a result of the SQT and the IT, Dillman et al. (1993) estimated that the second mailing would increase response by 10.4 percent. Subsequently, the Census Bureau also carried out two studies investigating the impact of a second mailing in hard-to-count areas. Dillman et al. (1994) showed that a second mailing added 10.5 percent to the response rate. Given the findings of this research, it is unclear why there was a need to examine the benefits from the use of a replacement questionnaire in the 2003 census test (National Research Council, 2003).

Second, areas in which research has demonstrated clear preferences have been ignored in subsequent research projects, when, for example, the previously preferred alternative was not included as a control (see National Research Council, 2006: Box 5-3). Furthermore, there are some basic questions that never get sufficient priority because they are by their nature long-term questions. The best way to represent residence rules is an obvious example. Finally, the analysis of a test census is often not completed in time for the design of the next test census, therefore preventing the continuous development of research questions.

The Census Bureau needs to develop a long-term plan for obtaining knowledge about census methodology in which the research undertaken at each point in time fully reflects what has already been learned so that the research program is truly cumulative. This research should be firmly grounded in the priorities of improving data quality and reducing census costs. Research continuity is important not only to reduce redundancy and to ensure that findings are known and utilized, but also because there are a number of issues that come up repeatedly over many censuses that are inherently complex and therefore benefit from testing in a variety of circumstances in an organized way, as unaffected as possible by the census cycle. These issues therefore need a program of sustained research that extends over more than a single decennial cycle. Also, giving people more freedom to pursue research issues may reduce turnover in talented staff.

Finally, given the fielding of the American Community Survey, there is now a real opportunity for research on census and survey methodology to be more continuous. These preliminary considerations will be greatly amplified by the panel in its subsequent activities. In the meantime, we make the following recommendation as an indication of the overall theme for which the panel anticipates developing a more refined and detailed message in later reports.

RECOMMENDATION 8: The Census Bureau should support a dedicated research program in census and survey methodology, whose work is relatively unaffected by the cycle of the decennial census. In that way, a body of research findings can be generated that will be relevant to more than one census and to other household surveys.

For example, the Census Bureau can determine what is the best way to improve response to a mailed questionnaire through use of mailing materials and reminders, or what is the best way using a paper questionnaire or the Internet to query people as to their race and ethnicity, or what is the best way using a paper questionnaire or the Internet to query people as to the residents of a household. The objective will be to learn things whose truth could be applied in many survey settings and to create an environment of continual learning, and then document that learning, to create the best state-of-the-art information on which to base future decisions. When an answer to some issue is determined, that information can be applied to a variety of censuses and surveys, possibly with modest adaptations for the situation at hand. This is preferable to a situation in which every survey and census instrument is viewed as idiosyncratic and therefore in need of its own research projects. However, one complication of developing a continuous research program on censuses and surveys is the different environments that censuses and surveys of various kinds represent. We hope to have more to say on how to deal with this in our final report.

As pointed out by the Panel on Residence Rules in the Decennial Census, "Sustained research needs to attain a place of prominence in the Bureau's priorities. The Bureau needs to view a steady stream of research as an investment in its own infrastructure that—in due course—will permit more accurate counting, improve the quality of census operations, and otherwise improve its products for the country" (National Research Council, 2006:271). A major objective of the remainder of the panel's work will be to provide more specifics on how such a research group could develop and carry out a research program in various areas and overall, and how they would make use of the various venues and techniques for research, testing, experimentation, and evaluation.

4

Considerations for the 2010 Census

In carrying out our primary charge regarding the selection of experiments and evaluations for the 2010 census, the panel inevitably had to consider plans for the conduct of the census itself. Moreover, the conduct of every census inevitably affects the Census Bureau's overall research program for the decennial censuses. Thus, in this chapter the panel presents three recommendations concerning some census operations with a view to their contributions to improvement to census methodology. Although we understand that the design of the 2010 census is relatively fixed, we hope that the material in this chapter may still be of use to the Bureau.

TECHNOLOGY

The Census Bureau will be using more technology in the 2010 census than in previous censuses, and this has raised some concerns that the panel would like to see addressed in the final plans for 2010. The concerns involve the functioning of the handheld computing devices to collect field enumeration data and the operation of the management information system for the 2010 census. By management information system is meant the various software systems that manage and monitor, somewhat interactively, the mailout-mailback process, nonresponse follow-up, field enumerator hiring and firing and compensation, questionnaire data capture, and other major census processes. We don't know the full extent to which these systems need to interoperate, but at least some modest degree of interaction is required, for example between the Master Address File (MAF)–TIGER system and the handheld devices in providing electronic maps for the handheld devices to display. The two primary concerns are whether the transmission of data using the handheld computing devices could be compromised in some manner (or could be lost unintentionally through mistakes and technological problems) and whether the needed interoperability of the components of the management information system could be hampered either by the adapting of software or the acquisition of newer software releases for the various components of the system between the dress rehearsal and the 2010 census.

With respect to the security of the transmissions of the handheld computing devices, the motivation to do harm to the census counts may be relatively modest given the lack of a financial incentive, and this may result in less chance for a security breach. However, this argument is not compelling. Furthermore, not only is there interest in reducing the opportunity for a security breach, there is also the matter of being able to assure census data users that the counts are valid. To accomplish this, the Census Bureau should carry out an independent validation and verification of the functioning of the handheld devices. This could be accomplished in the following ways, either in the 2008 dress rehearsal or in the 2010 census:

1. Establish a dual recording stream for all data from mail-in, telephone, or handheld devices: one file to go to the contractors and one to be retained by the Census Bureau. In the event of catastrophic failure by a contractor or a serious challenge to the results, it will be important to have all the raw data in the hands of the Census Bureau.
2. It is practical to develop simple programs, written and run by Census Bureau personnel, that will search large data files for patterns of interest. In this way, unexpected or curious results can be efficiently discovered and checked, and this can contribute to the validation and verification effort.
3. Related to points (1) and (2), the Census Bureau should develop quantitative validation metrics, a priori, to check for data set self-consistency and comparison of redundant data.

Other important general operational measures that we recommend for the 2010 census, either to determine whether any security breaches have occurred or to prove that the 2010 Census was secure (and which are probably already carried out), include:

- Retention of an archive of all raw data with date and time stamps. In the event of serious software failure, it would be important to be able to "replay the census" from these raw data.
- Use, by the Census Bureau and contractors, of dedicated processing systems that run no other applications and have highly secured network connections and secure accounts.
- Use of periodic system checkpoints to monitor and analyze software systems for intrusions or unauthorized manipulations of data.
- Strict control over handheld devices, including their inventory, individual device identification, and permission to operate (turn them on, turn them off, enable data transfer, disable data transfer, etc.).
- Use of methods to prevent and detect bogus data streams, including data that impersonate handheld devices.

With respect to concerns about configuration control of the management information system of the 2010 census, the processing history of the dress rehearsal could be retained and the software systems intended for use in 2010 could be used to "replay" the dress rehearsal soon before the 2010 census to identify any systems that fail to interoperate. That is, assuming that the management information system for the dress rehearsal functions well, saving the processing history would then provide a means for determining whether modifications or updates of components of the management information system between 2008 and 2010 had raised any interoperability problems. (This is referred to as regression testing.) In addition, all information system errors encountered during the dress rehearsal should be captured in a form that allows them to be used during the software development work between the dress rehearsal and the start of the 2010 census.

RECOMMENDATION 9: The Census Bureau should use dual-recording systems, quantitative validation metrics, dedicated processing systems, periodic system

checkpoints, strict control over handheld devices, and related techniques to ensure and then verify the accuracy of the data collected from handheld computing devices.

RECOMMENDATION 10: The Census Bureau should provide for a check to ensure that the subsystems of the management information system used in 2010 have no interoperability problems.

DATA RETENTION BY CENSUS CONTRACTORS

Given the very successful use of contractors to carry out several decennial census processes in the 2000 census, it is expected that the use of contractors will be expanded in 2010. The component processes that will be contracted out in 2010 include (1) the decennial response integration system (DRIS), which involves systems management of the process of questionnaire response and data capture; (2) the automation of field data collection (FDCA); (3) the data access and dissemination system II (DADS II); (4) the 2010 census communications campaign; and (5) the printing contract. The fact that these systems will be operated by contractors raises an additional complication. Any data collected as part of developmental or operational testing of these systems prior to their use in 2010, as well as any data collected in monitoring the operations of these systems while in use in 2010, may be viewed as proprietary. This would limit the Census Bureau's ability to assess the performance of these systems in looking toward 2020. While the contractors themselves may issue their own evaluation studies, this is insufficient given that contractors have a bias in evaluating their own systems. We assume that contractual agreements about the sharing of such data, if they have not already been provided for, are now too late (especially for developmental testing results). In that event, the Census Bureau should try to develop some informal understandings of data sharing with their contractors to address this issue. If it is not too late, such data-sharing clauses should be included in contracts.

CENSUS ENUMERATION AS PART OF TELEPHONE QUESTIONNAIRE ASSISTANCE[1]

The current plans regarding the use of Telephone Questionnaire Assistance (TQA) are for it to function primarily as a means for assisting the public in making correct responses to the census form, in particular for complicated situations involving residence rules or responses to the race and ethnicity questions. In addition, this is a method for people to obtain assistance in filling out the census questionnaire when English is not their primary language. On occasion, this has also been a vehicle for households to provide their responses to the census questionnaire. However, this possibility was not encouraged in 2000.

[1] Telephone Questionnaire Assistance was an operation used in the 2000 census in which people could call a toll-free number to get help in filling out their census questionnaire, to arrange to be sent a replacement questionnaire, to arrange to be sent a language guide, or to provide their census questionnaire information in situations in which they were not provided a census questionnaire.

For the 2010 census, we think the Census Bureau should consider making more transparent to respondents this option of collecting the information for the entire census questionnaire over the telephone once someone calls TQA. Specifically, whenever someone connects to TQA, the willingness of the operator to take the complete information, instead of just providing the specific help requested, should be made known to the caller during the initial part of the interaction. Our understanding is that this was not done in previous censuses due to the resources needed, especially the number of operators, and due to the additional procedural complications, especially of providing this opportunity for those receiving the census long form. However, given that this is a short-from-only census, we think that the need to get the information as soon as possible, when possible, should outweigh other concerns about making this option more frequently used. This could be especially important if the hourly wages of field enumerators increase substantially in 2010, since collection of such information may importantly reduce the cost of the nonresponse follow-up.

If this change is not implemented in 2010, the Census Bureau should collect sufficient information to carry out an evaluation after the census is completed as to the percentage of callers to TQA who ultimately sent back their census questionnaires to estimate the additional nonresponse follow-up costs due to the lack of collection of the entire census questionnaire over the telephone. Also, a possible experiment that should be considered is to sample the callers and ask those sampled if they would mind providing their information at that time by telephone to better estimate the additional resources required.

RECOMMENDATION 11: The Census Bureau should strongly consider, for the 2010 census, explicit encouragement of the collection of all data on the census questionnaire for people using Telephone Questionnaire Assistance. In addition, the Census Bureau should collect sufficient information to estimate the percentage of callers to Telephone Questionnaire Assistance who did not ultimately send back their census questionnaires. This would provide an estimate of the additional costs of nonresponse follow-up due to the failure to collect the entire census questionnaire for those cases. The Census Bureau should also consider carrying out an experiment whereby a sample of callers to Telephone Questionnaire Assistance are asked whether they would mind providing their full information to better estimate the additional resources required as a result of expanding Telephone Questionnaire Assistance in this way.

In conclusion, the panel is enthusiastic about the opportunity to collaborate with the Census Bureau on its plans for selecting and designing productive experiments and evaluations in conjunction with the 2010 census and, more broadly, a more productive research program overall. The Census Bureau has a very proud history of innovation, including the development of punch card machines, the first nonmilitary application of computers, survey sampling, hot-deck imputation, FOSDIC (Film Optical Sensing Device for Input to Computers), to name a few, and we hope to help continue this important tradition.

REFERENCES

Abramson, F. (2004, November 17). *Census 2000 testing, experimentation, and evaluation program summary results.* Census 2000 Testing, Experimentation, and Evaluation Program. Washington, DC: U.S. Census Bureau. Available: http://www.census.gov/pred/www/rpts/TXE%20Program%20Summary.pdf.

Box, G.E.P., and Hunter, J.S. (1961), The 2k−p fractional factorial designs. Part I (Corr: V5 p417). *Technometrics, 3*, 311-351.

Buscher, M., and Stamm, H. (2001). *E-census: The 2000 census on the Internet—Concept and stock-taking*. Paper contributed to the Second Swiss eGovernment Symposium, August 22, 2001, Zurich.

Bye, B.V., and Judson, D.H. (2004). *Results from the administrative records experiment in 2000*. Census 2000 Testing, Experimentation, and Evaluation Program. Topic Report 16. Washington, DC: U.S. Census Bureau.

Caspar, R. (2004). *Results from the response mode and incentive experiment in 2000*. Census 2000 Testing, Experimentation, and Evaluation Program. Topic Report 18. Washington, DC: U.S. Census Bureau.

Cohen, J. (2000, March 9). *2000 census (shhh!) on the net. New York Times*.

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.

Decennial Management Division (2003, May). *Reengineering the decennial census: The baseline design for 2010*. Version 1.5. Washington, DC: U.S. Census Bureau.

de la Puente, M., and McKay, R. (1995). Developing and testing race and ethnic origin questions from the Current Population Survey Supplement on Race and Ethnic Origin. In *Proceedings of the 1994 Annual Meeting of the American Statistical Association: Section on Survey Research Methods*, Vol. 1.

Dillman, D.A. (1978). *Mail and telephone surveys: The total design method*. New York: Wiley.

Dillman, D.A., Clark, J.R., and Sinclair, M.D. (1993). How prenotice letters, stamped envelopes and reminder postcards affect mailback response rates for census questionnaires. In *Annual research conference proceedings*. U.S. Census Bureau. Washington DC: U.S. Department of Commerce.

Dillman, D.A., Clark, J.R., and Treat, J.B. (1994). Influence of 13 design factors on completion rates to decennial census questionnaires. in *Annual research conference proceedings*. U.S. Census Bureau. Washington DC: U.S. Department of Commerce.

Dolson, D. (2006). Efficient multi mode data collection in a census context. *Proceedings of the Survey Research Methods Section*. Alexandria, VA: American Statistical Association.

Duquet, L., and Gilmour, G. (2007). *2006 Census online project*. Presentation at the Statistics Canada IT Conference, April.

Gerber, E. (2003). *Privacy schemas and data collection: An ethnographic account*. Census 2000 Ethnographic Study. Washington, DC: U.S. Census Bureau.

Griffin, D.H., Fischer, D.P., and Morgan, M.T. (2001). *Testing an Internet response option for the American Community Survey*. Paper presented at American Association for Public Opinion Research annual conference, Montreal, May 17-20.

Griffin, D.H., and Obenski, S.M. (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.

Kincannon, C.L. (2006, June 6). *The 2010 decennial census program*. Prepared statement, testimony before the Subcommittee on Federal Financial Management, Government Information, and International Security of the Committee on Homeland Security and Governmental Affairs, U.S. Senate.

Kurihara, Naoki (2004). *Some Issues in the 2005 Population Census of Japan.* Statistics Bureau, Japan.

Laroche, D. (2005). *Evaluation report on Internet option of 2004 census test: Characteristics of electronic questionnaires, non-response rates, follow-up rates and qualitative issues.* Working Paper 23, Work Session on Statistical Data Editing. Ottawa: United Nations Statistical Commission and Economic Commission for Europe and Conference of European Statisticians.

Larwood, L., and Trentham, S. (2004). *Results from the social security number, privacy attitudes, and notification experiment experiment in census 2000.* Census 2000 Testing, Experimentation, and Evaluation Program. Topic Report 19. Washington, DC: U.S. Census Bureau.

Martin, E. (2007). *Final report of an experiment: Effects of a revised instruction, deadline, and final question series in the decennial mail short form*. 2010 Census Memoranda Series #55. U.S. Census Bureau. Washington DC: U.S. Department of Commerce.

Martin, E.A., Gerber, E., and Redline, C. (2004). *Census 2000 alternative questionnaire experiment*. Census 2000 Testing, Experimentation, and Evaluation Program. Topic Report 17. Washington, DC: U.S. Census Bureau.

McKay, R., and de la Puente, M. (1995). Research improves questions. *Civil Rights Journal, 1*(1).

MITRE Corporation (2007, January 17). *An assessment of the risks, costs, and benefits of including the Internet as a response option in the 2010 decennial census.* Version 2.0. Report prepared for the U.S. Census Bureau.

Moraleda, A.G. (2006). *Census data collecting by Internet: Spanish experience in 2001*. Working Paper 12, Work Session on Electronic Raw Data Reporting. Geneva: United Nations Statistical Commission and Economic Commission for Europe and Conference of European Statisticians.

Mule, Thomas (2002), "A.C.E. Revision II: Further Study of Person Duplication," DSSD A.C.E. Revision II Memorandum Series #PP-51, Decennial Statistical Studies Division, U.S. Bureau of the Census, December 31, 2002.

Mulry-Liggan, M. (1986). *Overview and summary of the forward trace study*. Bureau of the Census, Statistical Research Division Report Series, Number Census/SRD/RR-86-24. Washington, DC: U.S. Census Bureau.

Nash, F. (2000). *Overview of the duplicate housing unit operations*. Census 2000 Information Memorandum Number 78, November 7, 2000. U.S. Census Bureau. Washington DC: U.S. Department of Commerce.

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, Commission on Behavioral and Social Sciences and Education. Washington DC: The 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, Commission on Behavioral and Social Sciences and Education. Washington DC: The National Academy Press.

National Research Council (2003). *Planning the 2010 census: Second interim report*. Panel on Research on Future Census Methods, Daniel L. Cork, Michael L. Cohen, and Benjamin F. King, Editors. Washington DC: National Academies Press.

National Research Council (2004a). *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, Division of

Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (2004b). *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.

National Research Council (2006). *Once, only once, and in the right place: Residence rules in the decennial census.* Panel on Residence Rules in the Decennial Census, Daniel L. Cork and Paul R. Voss (eds.). Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (2007). *Using the American Community Survey: Benefits and challenges*. Panel on the Functionality and Usability of Data from the American Community Survey. Constance F. Citro and Graham Kalton (eds.), Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Office for National Statistics (2004, March). *Information paper: The 2011 census: A design for England and Wales.* London: Office for National Statistics.

Rodriguez, C.E. (1994). Challenges and emerging issues: Race and ethnic identity among Latinos. In *Proceedings from the Workshop on Race and Ethnicity Classification: An Assessment of the Federal Standards for Race and Ethnicity Classification*. Committee on National Statistics, National Research Council, February 17-18.

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.

Smith, I. (2006). *Using both Internet and field collection methods for the 2006 Census of Population and Dwellings*. Invited paper for the Population Association of New Zealand Conference. Auckland: Statistics New Zealand.

Statistics Canada (2007). *Census technology: Recent developments and implications on census methodology—Census on the net*. Paper presented for the tenth session of the Economic Commission for Europe, Conference of European Statisticians, Group of Experts on Population and Housing Censuses. ECE/CES/GE.41/2007/9.

Statistics New Zealand (2007). *Implications of the Internet census for the management of field operations.* Paper prepared for the Joint UNECE/Eurostat Meeting on Population and Housing Censuses, Kazakhstan, June 4-6, 2007.

Swain, L., Drew, J.D., Lafrance, B., and Lance, K. (1992). The creation of a residential address register for coverage improvement in the 1991 Canadian Census. *Survey Methodology, 18*, 127-141.

Trewin, D. (2006). *How Australia takes a census*. Report 2903.0. Canberra: Australian Bureau of Statistics.

United Nations Economic Commission for Europe (2006). *Conference of European Statisticians: Recommendations for the 2010 censuses of population and housing.* Prepared in cooperation with the Statistical Office of the European Communities. ECE/CES/STAT/NONE/2006/4. New York and Geneva: United Nations.

U.S. Census Bureau. (1997). *Results of the 1996 race and ethnic targeted test*. 20233-8800, Population Division Working Paper No. 18, May. Population Division and Decennial Statistical Studies Division. Washington DC: U.S. Department of Commerce.

U.S. Census Bureau (2004, January 20). *2003 National Census Test: Contact strategy analysis.* Final report, 2003 National Census Test Analysis. Decennial Statistical Studies Division and Planning, Research, and Evaluation Division. Washington, DC: U.S. Census Bureau.

U.S. Department of Commerce, Office of Inspector General (2006). *Enumerating group quarters continues to pose challenges*. Final Inspection Report No. IPE-18046. Washington, DC: U.S. Department of Commerce.

Whitworth, E. (2002, August 14). *Internet data collection*. Census 2000 Evaluation A.2.b. Washington, DC: U.S. Census Bureau.

Williams, P. (2006). *The Australian 2006 census and the Internet*. Working Paper 9, Work Session on Electronic Raw Data Reporting. Geneva: United Nations Statistical Commission and Economic Commission for Europe and Conference of European Statisticians.

APPENDIX A

THE CENSUS BUREAU'S SUGGESTED TOPICS FOR RESEARCH

The following chart was provided to the panel by the Census Bureau as a partial summarization (augmented by several other reports and presentations) of their deliberations as to the research topics that should be considered for either experimentation during the 2010 census or evaluation shortly after. The leftmost column provides an identification key for each topic along with a short series of either questions or a brief discussion that defines the topic. The next block of columns provides criteria that should be used to help rank these topics, initiated by a high-medium-low ranking of the resulting importance of the topic. The criteria are anticipated impacts on cost, quality of data, whether the topic would require a new census component process, and whether it was accomplishable. Finally, the last block of columns provides information on whether the topic was better suited to 2010 or 2020 and whether a census environment was needed to assess alternatives to current census processes.

TABLE A-1 2010 Census Program for Evaluations and Experiments: Appendix to Summaries of Suggested Research

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
A. Coverage Measurement									
A.1 Census Coverage Measurement (CCM) is the program that will answer the question: How accurate was the coverage of the population?	1	Yes	Yes	No	Yes		Yes	Yes	Yes
A.2 How effective is the CCM interview and subsequent processing in determining the members of the household at each housing unit on CCM interview day and the usual residence of each household member on Census Day? When there are errors in determining household membership and usual residence, what are the causes and what are the possible remedies? What are the effects of recall errors and reporting errors on the CCM interview?	1	No	Yes	Yes	Yes		Yes	Yes	Yes. (Note: Some things can be done outside the census.)
A.3 Can we start to learn if comparing the history of census operations with the CCM results in the sample blocks can help us explain how and when errors occur, and also suggest potential remedies?	1	No	Yes	No	Yes, as a feasibility study in small number of blocks in the CCM sample.		No	Yes	Maybe

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
A.4 Our knowledge about Group Quarters (GQ) coverage is very limited. Efforts to estimate GQ coverage in 1980 and 1990 were limited (in both scope and success), and the GQ population was out-of-scope for the 2000 Accuracy and Coverage Evaluation (A.C.E). Since the GQ population will also be out-of-scope for CCM in 2010, we need to consider that the problems are likely to be different for different types of GQs (e.g., college dorms vs. nursing homes vs. migrant farm worker camps).	1	No	Yes	No	Yes	Always a problem	No	Yes	Yes
A.5 How can we develop a standard of comparison for household membership on CCM interview day and usual residence on Census Day? What are the effects of recall errors and reporting errors on the CCM interview? Candidate methods for developing the standard include ethnographic studies matched to the census and CCM interviews, respondent debriefings following a CCM interview, and an in-depth Living Situation Survey.	2	No	Yes	Yes	Yes, with limitations (scope must be very limited, and any "standard" cannot be expected to get exact truth).		Yes	Yes	Yes, Ideally
A.6 Can administrative records augment CCM fieldwork from telephone follow-up to reduce cost and improve CCM data quality? For example, administrative records information may aid in confirming which enumerations linked in the computerized search for duplicates are the same person when the determination cannot be made in the field. An evaluation of A.C.E. Revision II estimates of duplication in Census 2000 using administrative records information demonstrated potential for improving CCM data quality in this manner.	2	Yes	Yes, unclear	N/A	Yes		No	Yes	Yes

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
B. Race and Hispanic Origin									
B.1 Evaluate alternative race and Hispanic origin questions to include (1) double-banking of response categories and shared write-in spaces, (2) modified examples to follow the Advisory Panel recommendations, (3) separate evaluation of the features of 2005 National Census Test (NCT) panel 6, to better understand how each influences Hispanic and race reporting, and to inform future decisions, (4) modified Hispanic question that allows multiple Hispanic reporting (Y/N, yes multiple types). The latter must be tested in both Nonresponse Followup (NRFU) and the mailout. Samples must adequately represent small groups. Re-interview is needed to assess data quality.	1	No	Yes	No	Yes		No	Yes	Desirable especially for small groups
B.2 Develop a combined race and Hispanic origin question.	3	No	Yes	No	Yes		Yes	Yes	Yes
B.3 Conduct Research to support rules for editing problematic race and Hispanic origin responses (e.g., Y/N responses to Hispanic origin). A goal is to better understand respondent intent of write-in entries in the presence of, and in the absence of, marking checkboxes.	1	No	Yes	No	Yes		No	Yes	Yes
C. Coverage Improvement									
Address List Development									
C.1 How accurate was the final address list?	1	No	Yes	No	Yes		Yes	Yes	Yes
C.2 How should we deal with updating the address frame coming out of the 2010 Census, so that we can avoid a large and expensive address canvassing operation in the future, or so that the operation could be conducted at a much reduced cost?	1	Yes	Yes	Yes	See Below (C.2.a)		Yes	Yes	Not sure
C.2.a How can the quality of the address frame be improved with a more scientific extract process?	1	Yes	Yes	Yes	Yes, we need evaluations to demonstrate that this can be done.		Yes	Yes	Not sure

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
C.2.b How can we use additional information (like the Delivery Sequence File in rural areas, American Community Survey (ACS) Time of Interview data, Carrier Route data, and the National Change Of Address file) to improve address list maintenance?	1	Yes	Yes	Yes	Yes, we need evaluations to demonstrate that this can be done.		Yes	Yes	Not sure
C.3 Can we target Address Canvassing activities better?	1	Yes	No	No	Yes, there are concerns about the ability to reliably match persons with common names and across long distances	Politically acceptable – Can we convince stakeholders that we don't need to do an address canvassing in their jurisdiction, while we do need to do it in the neighboring jurisdiction.	Yes	Yes	No; need 2010 data
C.4 How accurate were the data collected in Address Canvassing? How can we improve Address Canvassing quality?	1	No	Yes	No	Yes		Yes	Yes	Yes
C.4.a How well does automated Global Positioning System (GPS) collection work in terms of completeness and accuracy of GPS coordinate data?	1	No	Yes	No	Yes		Yes	Yes	Yes
C.4.b How can we improve GPS collection—increase human intervention, improve automated collection, both?	1	No	Yes	No	Yes		Yes	Yes	Yes
C.5 How can we improve address list maintenance, operational procedures, and enumeration of small multi-unit structures (2-10 units)?	2	No	Yes	No	Yes		Yes	Yes	Yes

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
Administrative Records									
C.6 How can we avoid the need for followup and use administrative records to: a. Identify coverage problems? b. Identify and classify duplicates? c. Resolve potential coverage problems identified by the coverage probes.	1	Yes	Yes	Yes	Yes	Privacy concerns and issues with file access. Stakeholders have expressed reservations about the use of administra-tive records	Not sure	Yes	Yes

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
Coverage Followup (CFU)									
C.7 Does Coverage Followup actually work? a. How effective is it? b. Is CFU effectively identifying omissions? c. Is it introducing bias? d. How do recall and reporting errors affect its determination of residency, and hence erroneous enumerations (EEs)? e. How can we afford to follow up on more coverage improvement cases? f. Is the expense of CFU worth the coverage gain? g. Can certain categories of response to coverage questions be automatically coded, or field coded by interviewers to reduce follow up workload? . . . h. What recall and reporting problems affect CFU's ability to identify: missed people the respondent had in mind when filling out the undercount question? Are enumerators screening out people who are eligible to be listed? i. How to optimize which cases are coded for CFU?	1	Yes	Yes	Yes	Yes		Yes	Yes	Yes
C.8 Develop and experimentally evaluate alternative designs for coverage followup instruments. Alternative methodologies might involve dependent questions; self-response by all relevant household members; immediate follow-up; and other methodological improvements to facilitate recall and reporting in CFU (i.e., conduct an integrated experiment).	1	Yes	Yes	Yes	Yes		Yes	Yes	Yes

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
C.9 To what extent did nationwide person matching improve the identification and removal of duplicates of housing units and persons in the census? In particular, what improvements can be made in the identification and removal of census duplicates of persons across some distance given the challenges created by chance agreements of names and birth dates?	1	Yes	Yes	Yes	Yes there are concerns about the ability to reliably match persons with common names and across long distances.		Yes	Yes	Yes
C.10 Develop and experimentally evaluate alternative designs of the undercount (and overcount?) questions in the mail form to effectively identify census coverage errors for follow-up. Variations might include format (open vs. closed) and wording of questions and response categories, and placement in the form.	1	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Residency Rules/Questionnaire Design									
C.11 Implement and experimentally evaluate alternative residence rules and presentation of roster instructions in paper and other modes, including the National Academy of Sciences (NAS) recommendations to ask a sufficient number of residence questions to determine residence, and to obtain alternative addresses. Panels would be included in an alternative questionnaire experiment (AQE) and would require a coverage re-interview. Cognitive testing is needed for development, along with research on respondents' reading behavior and use of flashcards or other ways of presenting instructions. Alternative approaches might include: a. de facto approach b. Worksheet approach c. Alternate address elsewhere	2	No	Yes	No	Yes		No	Yes	Yes, for some aspects

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
Be Counted									
C.12 What effects does the Be Counted Program have in filling gaps in coverage? a. Is it worth it? b. Is including it better than trying to ensure people are counted in other ways? c. Does it introduce coverage errors?	3	No	Yes	No	Yes	Public perception —of missing people or not giving them an opportunity to be enumerated in the census if not for the Be Counted program.	Yes	Yes	Yes
General									
C.13 How accurate was vacancy/occupancy status of housing units in Census? Are there ways to improve accuracy?	1	No	Yes	No	Yes		Yes	Yes	Yes
C.14 Through ethnographic research, can we learn more about American Indian and Alaska Native households, Hispanic households, and immigrant communities that might result in different methods for enumeration? This research could provide insight into CFU and Census Coverage Measurement (CCM) to understand deficiencies.	2	No	Yes	No	Yes		No	Yes	Yes

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
D. Field Activities									
Automation									
D.1 What was the impact of adding expanded automation to field data collection for Address Canvassing, Nonresponse Followup (NRFU) and Census Coverage Measurement - Personal Interview (CCM-PI)? Did we gain in efficiency? Did we see cost savings? Did automation contribute to operational improvements? Should we use the hand-held computers (HHCs) in operations other than Address Canvassing, NRFU and CCM-PI in 2020(e.g., U/E)?	1	Yes	Yes	Yes	Yes		Yes	Yes	Yes
D.2 What was the impact on field staff of using HHCs to conduct field data collection operations? Did using HHCs help us to improve the effectiveness and efficiency of field staff? Did using the HHC help us improve the productivity of field workers? What impact did the HHC have on field staff training?	2	Yes	Yes	Yes	Yes		Yes	Yes	No
Training									
D.3 How can enumerator training be improved? a. Can we make enumerator training more efficient/effective through the redesign of enumerator training materials and job aids? b. Can we make enumerator training more effective by expanding the use of technology-based training?	1	Yes	Yes	Yes	Yes		Yes	Yes	

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
D.4 How can we better prepare enumerators to ensure they are effective and efficient in their job? What is the optimal contact strategy for NRFU? How many contacts should we make for NRFU? (Requires experimental design).	1	Yes	TBD	No	This research requires an experiment design with different contact strategy in different locations. Implementing such a design in the census environment may not be practical or worthwhile		Yes	Yes	Yes
D.5 How can enumerator training be improved to reduce/minimize errors that may be introduced by interviewers? The focus is on interviewers' contributions to coverage errors (in NRFU, CFU, and CCM-PI) in particular, but also errors in other short form items. Possible research approaches might include: an interviewer variance study, in which interviewer assignments are randomized, or assigning a sample of mail returns to NRFU enumerators for re-interview.	2	No	Yes	Yes	There are studies that could be done to determine interviewer contribution to error. However, we are not sure if training could be improved or changed to address those contributions.		Yes	Yes	Yes

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
Quality Control (QC)									
D.6 How can Global Positioning System (GPS) Technology be used as a QC tool for field work, e.g., to identify curbstoning or inefficient field work?	1	Yes	Yes	Yes	Yes	Privacy concerns—GPS Tracking of employees	No	Yes	Yes
D.7 How can the QC design for field operations be improved to be more effective/efficient? a. How much does the QC improve the quality of the census operations? Does the QC have a high probability of identifying data falsification and/or violation of procedures? b. Is there an efficient way to verify the QC work? Is it worth it to verify the QC work?	2	No	Yes	No	Yes		Yes	Yes	Yes
D.8 Can a batch level approach to re-interview sampling improve efficiency and/or effectiveness of field re-interview operations?	3	Yes	Yes	No	Yes		No	Yes	Yes
E. Language									
E.1 Can an alternative design for the bilingual English/Spanish questionnaire result in improved data?	1	Yes	Yes	Yes	Yes	Backlash; referring to lowered mail response from non-Spanish-speaking populations	Yes	Yes	Yes
E.2 Is there a better or more efficient way to stratify the mailing of the bilingual forms?	2	Low	Low	Yes	Yes		No	Yes	No

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
E.3 What information can systematic observations yield about how census enumerators are obtaining information from households with little or no understanding of English? Are there changes we can or should make to our methodologies and practices to improve these interviews?	2	No	Yes	No	Yes, but may present complex stratification issues		Yes	Yes	Yes
E.4 Can we obtain better mail response and/or higher quality data by mailing a Language Assistance Guide Booklet that depicts the questionnaire in Five (5) languages?	1	Yes	Yes	Yes	Yes		No	Yes	Yes
F. Mode Effects									
F.1 What is the magnitude of the effects of mode on responses to 2010 census questions? This research would compare the 2010 mail mode content to the adaptation of the specific content items for other modes used in the 2010 census. An example would be comparing the 2010 mail form relationship question, which shows all 14 categories, to the proposed 2010 telephone-adapted version, which asks an open-ended question. In general, this study would examine response distributions (or reliability and other data quality measures) for the 2010 mail items compared to the adapted versions used in 2010 for other modes; comparable random samples would be ideal to avoid self-selection confounds.	1	No	Yes	No	Designing an experiment that reduces/ eliminates the self-selection bias is complex and may not be feasible from a field/budget/ schedule standpoint.		Yes	Yes	Yes

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
F.2 What are the effects of mode for alternative adaptations of the 2010 census questions in non-mail modes? This research would compare the 2010 mail mode content to alternative adaptations of the specific content items for other modes. These alternative versions for the non-mail modes are adaptations, which show promise in terms of providing comparable data to the mail form, but were not used in the 2010 census. We would examine response distributions (or reliability and other data quality measures) for the 2010 mail items compared to the alternative adapted versions for other modes; comparable random samples would be ideal to avoid self-selection confounds.	2	No	Yes	No	Designing an experiment that reduces/ eliminates the self-selection bias is complex and may not be feasible from a field/budget/ schedule standpoint.		Yes	Yes	Yes
G. Content									
G.1 What are the combined effects on the data of all questionnaire changes made in the 2010 mail questionnaire?	1	No	Yes	No	Yes		Yes	Yes	Yes
G.2 What are the consistency and reliability of reporting in the 2010 census?	1	No	Yes	No	Yes		Yes	Yes	Yes
G.3 How well do questions perform in interviews?	2	No	Yes	No	Yes		Yes	Yes	Yes
G.4 How comparable are Census 2010 data and American Community Survey (ACS) data?	1	No	Yes	Yes	Yes		Yes	Yes	Census data required but not environment
G.5 Do current methods for identifying the householder/Person 1 perform well in all modes? If not, can improved method(s) be developed?	2	No	Yes, possibly	No	Yes		No	Yes	No

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
H. Self-Response Options									
H.1 How can we improve alternatives for increasing mail response? An experiment for 2006 showed that a deadline plus delayed mailing of questionnaires improved the mail response rate by two percentage points? This should be replicated in the 2010 Census as an experiment to see if results hold up in a census environment, and to get good data on timing of returns under a deadline. If design permits, effect of deadline messaging and compressed schedule could be teased out.	1	Yes	Yes	No	Yes		No	Yes	Yes
H.2 Experiment testing an additional contact reminder after replacement questionnaire. This contact would contain stronger language, relative to the reminder postcard and replacement questionnaire, indicating that failure to comply would mean inclusion in the Nonresponse Followup (NRFU) workload (more expense, etc). Different types of contacts could be tested such as postcard, full size letter, phone message, etc.	1	Yes	Yes	No	Yes		No	Yes	Census environment is optimal but at least site test is required to utilize NRFU message.
I. Special Places/Group Quarters									
I.1 Did the revised Group Quarters (GQ) definitions improve the identification and classification of GQs (GQs versus housing units, and by type)?	1	No	Yes	No	Yes		Yes	Yes	Yes
I.2 Evaluate methods for improving GQ data collections by: 1) assessing the yield from the various sources used to update the MAF/TIGER database (MTDB), as well as assessing how the various census operations update the MTDB; 2) studying effects of allowing a Usual Home Elsewhere in more types of GQs; and 3) collecting additional information to assist with unduplication of college students.	2	No	Yes	Yes	Yes		Yes	Yes	Yes

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
I.3 The National Academy of Sciences recommends, "The U.S. Census Bureau should participate in a comprehensive review of the consistency of content and availability of prison records. The accuracy of prisoner-reported prior addresses is uncertain, and should be assessed as a census experiment. A research and testing program, including experimentation as part of the 2010 census, should be initiated by the Census Bureau to evaluate the feasibility and cost of assigning incarcerated and institutionalized individuals, who have another address, to the other location." (National Academy of Sciences (NAS) report entitled "Once, Only Once, and in the Right Place – Residence Rules in the Decennial Census (September 2006), pp 9-10)."	3	No	No	Yes	Maybe—the scope of this comprehensive review of prisoner reported address information would be massive if all levels of correctional facilities (e.g., federal, state, local, and private) were included in the review. Additionally, the Census Bureau has concerns about the feasibility of actually collecting the prisoner address information.	There could be possible implications for other types of GQs. This topic has generated much discussion and varying views among census stake-holders.	Yes	Yes	Yes
J. Marketing/Publicity/Paid Advertising/Partnerships									

Topics and Questions	Criteria						Considerations		
	Rank 1 = H 2 = M 3 = L	Cost (Big Payoff)	Quality	New to Census	Accomplishable	Other Criteria	For 2010?	For 2020?	Census Environment Required
J.1 How effective was the communication strategy for improving response and accuracy of the census? a. How do the separate components of the communications strategy contribute to the improvements (e.g., advertising, partnerships, etc.)? b. How effective were the targeted messages at reaching specific audiences? c. Did the communications strategy change attitudes or behavior toward and/or increase awareness of participation in the census?	1	Yes	Yes	No	The efforts in 2000 to evaluate the Advertising/ Marketing approach was inconclusive. Whether or not we can evaluate the 2010 approach depends on the solution put forth by the vendor and our ability to develop reliable technology.		Yes	Yes	Yes
K. Privacy									
K.1 Test alternative presentation and placement of privacy messages in cover letter, etc.	1	Yes	Yes	No	Yes		No	Yes	Yes
K.2 Monitor public concerns about privacy and confidentiality in a series of quick-turnaround surveys conducted during the census to provide U.S. Census Bureau executives with timely information about emerging concerns and issues. Data from monitoring surveys can also augment (or replace) traditional outreach evaluation surveys, which are slow and do not provide useful information on a timely basis.	1	No	No	Yes	Yes	Can help decision makers respond to emerging issues, crises	Yes	Yes	Yes

Criteria and Considerations for Assessing Proposed Research Topics and Questions

Criteria:

Cost (Big Payoff) – **[Yes/No]** Will results potentially lead to substantial cost savings in the 2020 Census?

Quality – **[Yes/No]** Could results conclusively measure effects on data quality?

New to Census – **[Yes/No]** Does the question address operations that are new since Census 2000, experienced significant procedural change, or experienced significant issues during Census 2000?

Accomplishable – **[Yes/No]** Will data be available to conclusively answer the question? Will there be a high demand of resources to address and answer the question? Are complex or untested methods foreseen to address and answer the question?

Considerations:

For 2010 – **[Yes/No]** Is this research question intended to assess an operation in the 2010 Census?

For 2020 – **[Yes/No]** Is this research question intended to assess a 2010 Census operation to inform the 2020 Census?

Census Environment Required? **[Yes/No]**

SOURCE: 2010 Census Program for Evaluations and Experiments—Appendix to Summaries of Suggested Research (planning document shared to the panel by the U.S. Census Bureau, April 13, 2007).

APPENDIX B

INTERNET RESPONSE OPTIONS IN SELECTED POPULATION CENSUSES

In this appendix, we briefly describe provisions for an online response option in past (and upcoming) censuses of population. We begin by describing the use of online response in the 2000 U.S. census and subsequent tests before describing experiences in other countries.

THE INTERNET AND THE U.S. CENSUS

The 2000 Census

The Internet response option was implemented in the 2000 census without the benefit of prior large-scale testing. Online response was considered for the 1998 dress rehearsal but ultimately abandoned "due to security concerns" but was revived in late 1998 by a Commerce Department directive (Whitworth, 2002:1). Due to insufficient time, online response was restricted to the 2000 census short-form questionnaire only and a single language (English). Programming of the form was kept as simple as possible for compatibility with different web browsers; JavaScript was avoided because it was deemed "unstable in some environments" (Whitworth, 2002:1). As a consequence, the online form was essentially presented as a single screen page rather than walking through separate questions in different web pages; hence, real-time editing and confirmation steps were not used, nor were skip patterns to move respondents through the questionnaire.

To access the electronic questionnaire, respondents needed to have the paper questionnaire that they received in the mail in hand. Following a link from the main census web page, they were asked to enter the 22-digit Census ID printed on the paper form's label (thus ensuring a linkage to a specific mailing address). If the 22-digit ID was confirmed as valid, then the questionnaire appeared onscreen. No publicity was given to the Internet response option.

During the time span between the opening of the online questionnaire site and the cutoff for nonresponse follow-up workload (March 3 to April 18, 2000), 89,123 submissions of Census ID numbers were made on the web site. Of these (Whitworth, 2002:5):

- 74,197 (83.3 percent) were valid Census IDs; however, only 71,333 resulted in a questionnaire submission. The other 2,864 may have been instances in which a respondent made an error entering the ID but inadvertently entered a valid number; they could have then broken off the interview and subsequently rekeyed their ID correctly. After some reconciling for unique address identifications,

questionnaire data from 66,163 of the 71,133 submissions were ultimately sent on for processing; about 1,500 online submissions are unaccounted for in the Bureau's tallies, with "no apparent explanations for this discrepancy" (Whitworth, 2002:???6).

- 14,926 (16.7 percent) attempts to enter a Census ID were failures. That this proportion matches the approximate 1-in-6 coverage of the census long-form sample is perhaps telling: "since [the Census Bureau] did not advertise the Internet response option, respondents would have also had no idea that long-form households were ineligible." Hence, "it is quite possible that many, if not most, of the submission failures" were attempts to use the Internet to answer a long-form questionnaire.

Although the vast majority of the Internet responses (98.4 percent) were each associated with only one ID number, there were some repeats of ID numbers: specifically, 1,090 ID numbers had to account for 2,853 responses. Most of these were incidents of 2 or 3 entries per ID and involved a pure replication of the same data; most likely, this was caused by a respondent clicking on the "Submit" button multiple times waiting for the browser page to load. The extreme case was a single ID associated with 17 entries; "many of these were on different days, and many with different data" (Whitworth, 2002:8-9). After final processing, 63,053 households representing 169,257 persons were included in the census through the Internet form.

The Census Bureau evaluation of the Internet response option in 2000 (Whitworth, 2002:17) deemed it "an operational success" and argued for further research:

> Obviously, the Internet is here to stay. The software and hardware developed for this program could have handled **tens of millions** of records instead of the **tens of thousands** it did handle. It is our recommendation that future research focus not necessarily on **how** to implement the form itself, but how to promulgate the Internet form as an option and convince the public that there is sufficient data security. Future research should also focus on how to use it as a tool to increase data quality by implementing real-time data feedback and analysis.

Response Mode and Incentive Experiment

Conducted as an experiment in the 2000 census, the Response Mode and Incentive Experiment (RMIE) gauged response rates to the 2000 census questionnaire by paper, interactive voice response (IVR, a fully automated telephone interview), or the Internet. In addition, the test considered whether the offer of an incentive (specifically, a 30-minute telephone calling card) influenced the response rates. The test (including a print of the Internet census form) is documented by Caspar (2003). The Internet usage survey component of the RMIE yielded relatively small numbers of online returns (with or without the incentive of a calling card), and some respondents noted a preference for paper. However, Caspar (2003:21) argued for further work on an online response option:

> Based on conservative assumptions and the data from RMIE, one might save between one and six million dollars in postage costs alone if between three percent and 15 percent of the sample uses the web rather than the mail survey. . . . This savings would more than offset the costs required to design, develop and maintain the web survey. Of course, the web survey would also produce savings related to reduced processing (receipt and scanning). Given this crude calculation, it is anticipated that the Internet would be cost-effective even if a relatively small proportion of respondents used it.

The 2003 and 2005 Tests

The 2003 National Census Test was designed as a mailout-only test: no fieldwork for nonresponse follow-up was planned or conducted. The mail sample was divided into 16 panels, 7 of which tested revisions of the census questions on race and Hispanic origin and 8 of which included different packages of response modes and contact strategies (e.g., sending a replacement questionnaire or a telephone reminder call, responses by telephone or the Internet). The Census Bureau concluded that offering the option of responding by telephone or the Internet along with the mailout of a paper questionnaire neither increased nor decreased the response rate. However, attempts to "force" respondents to use either of the electronic response modes by not including a paper questionnaire resulted in lower response rates. In terms of data quality, item nonresponse rates were significantly lower for the Internet responses than for paper returns for almost all items.[1]

A second mailout-only National Census Test in 2005 made another attempt to implement the telephone and Internet response modes, having made interface improvements in both. Illustrative screens—of the respondent log-in section and the race question—from the 2005 online instrument are shown in Figure B-1. Apparently, this test performed comparably to the options used in 2003 and did not yield major gains in response.

In November 2000–January 2001, the Census Bureau also conducted a test using 10,000 addresses on an Internet response option for the American Community Survey (ACS), the replacement for the traditional census long-form questionnaire in 2010. The recent report *Using the American Community Survey: Benefits and Challenges* describes ACS methodology in greater detail (National Research Council, 2007). In brief, the sample of households selected in one month is first contacted by mail and asked to return their questionnaire by mail. If they do not respond by mail, a telephone interview is attempted in the second month; if that fails, then enumerators attempt a personal visit in the third month. The hope of an Internet response option would be to supplement mail responses in the first months so that the follow-up steps in months 2-3 need not occur. Griffin et al. (2001) found that only about 2 percent of the respondents in the experimental group used the Internet response option (compared with about 36 percent by mail). The data showed some attempts to access or partially fill out the questionnaire, but they did not result in a

[1] The 2003 test was summarized (albeit without specific numbers) at http://www.census.gov/procur/www/2010dris/web-briefing/dris-tel-int.html.

full online form being submitted and were not enough to explain the low response rate. Although the response was low, the quality of the resulting data (in terms of whether subsequent editing was required) was found to be slightly better in the Internet responses than the mail responses.

Decision for 2010

An initial planning framework for the 2010 census (Decennial Management Division, 2003:3) noted among the major improvements planned for 2010 that "expanded use of Internet and telephone systems (using Interactive Voice Response) will provide new opportunities for using technology to make it easier for people to complete their questionnaire." The strategy document elaborated (Decennial Management Division, 2003:5-6):

> Fundamental to the 2010 census is expanding the ways people can be counted. Following a widespread awareness campaign, households will receive an advance letter in the mail before April 1, 2010. The letter will tell them about the census and the ways they can participate, using English or other language methods. . . . We will also use technology to build on this strategy by combining these mailings with Internet and telephone contacts. These technologies will provide respondents with additional options for receiving and submitting their census questionnaires. Our expectation is that we can increase the response rate even further by developing and implementing the optimal mix of contacts and response options. By taking advantage of the Internet and the telephone we can significantly increase the number of forms that move directly into data capture without needing to be scanned in a data capture center. . . .
>
> Despite all efforts to encourage everyone to provide information, we project that we will not obtain mail, Internet or telephone IVR responses from as many as 31% of the addresses to which we deliver a questionnaire. Many of these addresses will be vacant or nonexistent, but many will be occupied. Therefore, we must still conduct a nonresponse follow-up operation. . . .

Indeed, the initial scope of work for the Census Bureau's Decennial Response Integrated System (DRIS) for 2010 included requirements to facilitate census responses by three modes: paper, telephone, and Internet. The first two objectives suggested for the DRIS solution were to "Enable the Public" to "Obtain assistance or request an English or foreign language questionnaire or language guide using the telephone or Internet" and "Complete their 2008 Dress Rehearsal and 2010 Census questionnaire via the telephone, Internet and paper."[2] The DRIS contract was awarded to Lockheed Martin in October 2005.

[2] http://www.census.gov/procur/www/2010dris/web-briefing/dris-goals-objectives.doc.

However, the perceived low Internet response rates in the 2003 and the 2005 tests—combined with concern over inherent risks and the lack of guaranteed major cost savings—led the Census Bureau to reverse course. The Bureau's decision not to pursue online enumeration was formalized in a July 2006 decision memorandum. Earlier, on June 6, census director Kincannon (2006) offered the following argument in testifying before a U.S. Senate subcommittee:

> We have also considered other data collection methods, including Internet data collection. Based on our research, as well as our own experience and knowledge of the experiences of other countries, we do not believe Internet data collection would significantly improve the overall response rate or reduce field data collection. The Census Bureau offers an electronic response option for the Economic Census and other economic surveys and we generally obtain high response rates. It is altogether different, however, when we consider household and population surveys and censuses. The 2003 and 2005 Census Tests offered an Internet response option, and in both cases, the response rates were low, and offering an internet response option did not increase the overall response rate. We have also consulted the statistical offices of Australia, Canada, and New Zealand. Each of these countries utilized the Internet in their most recent censuses. The Internet response rate ranged from 7 to 15 percent. Each of the statistical offices indicated that it was not possible to accurately anticipate the response rate, and that ultimately using the Internet did not affect the overall response rate. Anticipating the response rate has important operational considerations. Because they were unable to accurately anticipate the Internet response rate, the other countries were unable to reduce the paper data capture operations out of concern they would not have the capacity to fully process the census responses. This would be true for the Census Bureau as well. Moreover, the Internet response option did not reduce the overall cost of data collection, and the cost for some specific activities, such as security and server capacity, increased.
>
> We have seriously considered the lessons our colleagues have learned. We are also concerned that utilizing the Internet could jeopardize other planned improvements. At this point in the decade, efforts to develop an Internet response option would divert attention and resources from tested and planned improvements such as the second mailing—which we know can increase the overall response rate by several percentage points. It is also important to keep in mind that the 2010 Census utilizes only the short form. There are very few questions in this form, and most can be answered by checking a box.

The major risks perceived by the Census Bureau—summarized in a commissioned report from the MITRE Corporation (2007)—are as follows:[3]

- Above all, the Census Bureau is concerned that something gone awry in an Internet response option—publicity of the census site being hacked or establishment of a "phishing" site appearing to be related to the census, for example—could cause voluntary response to the census to decline. This would tax nonresponse follow-up capabilities and raise the overall cost of the census.
- The Bureau's DRIS contractor concluded that it could not provide an Internet response facility in time for testing in the 2008 dress rehearsal, so that it would have to go into the main 2010 census without a large-scale test (as happened with the 2000 census online response option).
- A problem faced by any Internet site is a "denial of service" attack: deliberate bombardment with hits in order to shut down a site's operations.

(The MITRE evaluation also expresses concern that census data might be captured from individuals' computers through the use of spyware.)

In evaluating the Census Bureau's work on group quarters enumeration, the U.S. Department of Commerce, Office of Inspector General (2006:20-21) acknowledged the Bureau's decision not to use the Internet for main data collection in 2010. However, the review strongly suggested that the Bureau consider use of Internet methods for one traditionally hard-to-count population: college students. One reason for the selection of parts of Travis County, Texas, as a census test site in 2006 was a large college student population. Yet only 719 college student census report forms were returned during the test while expectations were that more than 6,700 should be found. In the inspector general's review, this suggested that online response options might appeal to the Internet-savvy college generation. Reacting to this recommendation, the Census Bureau reiterated its opposition to online enumeration generally.

USE OF THE INTERNET IN FOREIGN CENSUSES

In offering guidance to member countries on the 2010 round of population and housing censuses, the United Nations Economic Commission for Europe (2006) concisely summarized the basic rationale and concerns for permitting an Internet response option; this summary is presented in Box B-1. Stopping short of recommending that countries adopt an online version, the commission observed that online response is becoming an increasingly attractive option.

[3] The MITRE report was circulated on some technology blogs in July 2007, following a Senate subcommittee hearing at which the Census Bureau restated its intent not to pursue online enumeration. At the same hearing, Sen. Tom Coburn (R-OK) issued a public "Census Challenge" for ideas to use technology to reduce the costs of the 2010 census. See, e.g., http://www.fcw.com/blogs/archives/editor/2007/07/the_census_inte.asp, which contained a link to the MITRE report and references an interview with a former Census Bureau official.

In this section, we profile the use of the Internet as a response mode in selected censuses around the world, focusing almost exclusively on countries that still perform a traditional census rather than rely on a population register or other methods. Online enumeration has been performed in most of these cases; however, we also describe one census that ruled out Internet enumeration in its most recent census (Japan) and another that has not yet used the Internet in the census or in a major census test but intends to do so (United Kingdom).

One common theme to several of these profiles—particularly Canada, Australia, and New Zealand—is that the drive to allow the Internet as a response option came about through longer standing commitments to making government services electronically accessible. The Canadian "Government On-Line" initiative began in 1999, with the objective of making most government services accessible online by 2004-2005. The Canadian government also has an initiative to maintain a common visual theme on its websites, and the 2006 census website observed these basic standards (Laroche, 2005). The Government On-Line effort also included study of security and encryption protocols—an infrastructure on which Statistics Canada was able to piggyback. Similarly, the Australian Electronic Transaction Act of 1999 required agencies to permit electronic communications between citizens and the government (Trewin, 2006). In New Zealand, the "e-government strategy" adopted the goal of making the Internet "the dominant means of enabling ready access to government information and services" by mid-2004 (Smith, 2006).

Australia

In 2006 (as in previous years), the Australian quinquennial census was conducted on a drop-off–pick-up basis: enumerators delivered forms on the designated Census Night and returned within the next three weeks to pick them up. (Respondents were urged to complete the questionnaire on Census Night, as Australia uses a de facto residence concept.) The questionnaire package delivered to households also included a Census Form Number on the printed questionnaire and a 12-digit eCensus Number in a sealed envelope. Both numbers were needed to use the eCensus application on the Internet. The Australian Bureau of Statistics contracted with IBM to develop its eCensus web application and support systems.

Because of the drop-off–pick-up strategy used for the Australian census, designers needed to provide a mechanism for advising field enumerators that questionnaires in their districts had already been returned online, so that they did not need to do a follow-up visit. Ultimately, the Australian Bureau of Statistics (ABS) settled on notification by text message to enumerator cell phones;[4] this messaging system was part of a larger communications scheme connecting census field staff, central coordinators, and members of the public (who called with inquiries).

[4] A text message was also generated and sent to enumerators if a questionnaire was received by mail and processed.

Williams (2006) observes that "the 2006 eCensus system was opened to the public just after 8pm on 27 July, with enumerators due to commence delivery of forms on 28 July. The first eCensus respondent submitted their online form at 20:29 on 27 July." In total, ABS experienced a estimated 9 percent response rate via the Internet, representing 775,856 household forms; this slightly exceeded the system's performance in dress rehearsal, in which 7.9 percent of dwellings responded via the Internet. Due to the de facto nature of the census and the encouragement to complete the questionnaire upon receipt, 40.4 percent of all responses received by the Internet came in between 6 pm and midnight of the designated Census Night.

Prior to use in 2006, the Internet response option was tested in field tests in 2003 and 2004 and in the 2005 dress rehearsal. Based on the preliminary testing, ABS anticipated—and built its systems to accommodate—a surge of entries on Census Night. Contingency plans, including temporary service interruptions on the eCensus site and public relations messages, were also developed. As it turned out, "the capacity of the system was never really put to the test—with peak load on census night reaching only 15 percent of capacity" (Williams, 2006). ABS also developed contingency plans for malicious denial of service attacks on the census site—deliberate attempts to flood the system in order to shut it down. Mechanisms for monitoring the Internet service providers of incoming log-in attempts were put in place and, "in cases where these attacks could not have been dealt with quickly, public relations messages would have firstly assured the public that their census information is secure and secondly provide information about alternatives such as delaying use of the eCensus system or using the paper census form." However, no such denial of service attack was detected.

It is useful to note that Australia is effectively a long-form-only census—using only one questionnaire—rather than a distinction between short- and long-form samples or the 2010 U.S. census short-form-only model.

Canada

The 2006 Canadian census was the first to offer an online response option.[5] Every paper questionnaire sent by mail or dropped off by enumerators bore a 15-digit Internet Access Code (five groups of three digits) at the upper right of the questionnaire. A banner instruction immediately before "Step A" of the questionnaire read "COMPLETE YOUR FORM ON-LINE OR ON PAPER," and the first question advised respondents that they could complete the form online at a website (http://www.census2006.ca) using the Internet Access Code printed on the form.[6] A follow-up instruction to that option reminded online respondents, "Do not mail back your paper questionnaire."

[5] Dolson (2006) describes the multiple response modes offered in the 2006 Canadian census: "Respondents had a choice to respond [to a paper questionnaire sent by mail or dropped off by an enumerator] by either Internet or mail. Some data were collected by personal or CATI interviews. As well, respondents to the long[-form] questionnaire could either reply to the income questions or give Statistics Canada permission to link to their tax records to obtain these data."

[6] Respondents who lost the paper form could call a Census help line to request a new paper questionnaire or an Internet Access Code; alternatively, help line operators could also administer the questionnaire during the phone call. Responses generated though the help line—whether paper, Internet, or direct interview—

Online response was permitted for both the census short-form (8 questions) and long-form (53 questions) instruments. The online questionnaire could be rendered in either English or French, and the two languages could be toggled back and forth during the course of completing the online form. The Internet form was designed so that "no software trace (footprint) was left on [a respondent's] computer" once they had submitted it online. However, persons replying to the Canadian long-form questionnaire could indicate that they wished to pause and resume the questionnaire later; they were prompted to create a password and—upon logging back onto the census site—could resume the questionnaire where they left off. If they did not resume the form within some set period of time, though, the partial form was submitted for processing (Statistics Canada, 2007).

Prior to Statistics Canada's designated cutoff date to begin nonresponse follow-up activities, 22 percent of returned questionnaires had been returned online; overall, by the end of August 2006, the online response rate stood at 18.5 percent. Large households (5 or more people) were more likely to invoke the online option (26 percent) than smaller households, including single-member households (of which only 13.5 percent returned the form online). Online response rates did not seem to vary by form type (short or long form), but did vary by province: Alberta experienced the highest online response rate (21.4 percent) and the Northwest Territories and Nunavut—both of which are principally enumerated by personal visit rather than mail—the lowest (13.6 and 0.0 percent, respectively).

The 18.5 percent overall online response rate was consistent with expectations developed based on a 2004 census test using an Internet response option in parts of four provinces, as well as an Internet response experiment conducted as part of the 2001 census. Based on these pretests, Statistics Canada anticipated a 20 percent Internet share in 2006. Significantly, the 2004 test also led Statistics Canada to expect—and plan for—temporal patterns in questionnaire response. Like the U.S. census, Canadian census forms are delivered a few weeks before a designated reference date (Census Day); in the case of the 2006 Canadian census, Census Day was May 16. Based on the testing, Statistics Canada anticipated an early peak in online returns upon the first mailout in early March, with declining amounts until Census Day itself, at which point heightened publicity could be expected to create another response spike. Consistent with expectations, about 15 percent of responses received via the Internet came in on May 16 itself; system managers were able to devise a "graceful deferral" system on Census Day itself to limit the load on census servers.

In terms of data quality, Statistics Canada determined that Internet questionnaires produced much lower item nonresponse rates than did paper questionnaire responses: item nonresponse for paper questionnaires was 102 times higher than Internet questionnaires for short-form responses and 10 times higher for long-form responses. It was also determined that the Internet responses had lower failure rates during basic data

---

incurred an extra processing step: matching against an address register to determine the link to a geographic location (Dolson, 2006).

editing than the paper forms (Duquet and Gilmour, 2007). In part, this may be due to the use of data confirmation steps that are not possible on a paper form. The Internet short-form questionnaire (as well as computer-assisted forms used in nonresponse follow-up) prompted respondents to confirm the age of household members based on what had already been entered as their dates of birth (rather than answer both questions separately and potentially have a mismatch). The section of the Internet long-form questionnaire on household income also compiled the answers that had already been collected and presented them to the respondents for review and—if necessary—correction.[7] Use of the Internet option may also have saved costs in nonresponse follow-up due to the inherent limitation of space on the paper form: the version of the Internet instrument tested in 2004 permitted listings of up to 36 people, compared with the paper form's limit of information for 6 household members and names only for an additional 4 persons (Laroche, 2005).

During the conduct of the 2006 census, Statistics Canada also performed an experiment on targeting the Internet response option to particularly receptive audiences. This study—somewhat similar to the U.S. census tests in 2003 and 2005—was intended to suggest whether households "in geographic areas with a very high Internet penetration rate" might best—and less expensively—be contacted with only a letter and an Internet Access Code (but no questionnaire). As summarized by Statistics Canada (2007:12):

> A model was developed to identify *a priori* areas that include a significant number of dwellings likely to answer the Census online. Households in this study, called the Push Strategy, received only a letter instead of a paper questionnaire. These households were asked to complete their questionnaire online. The letter also included a 1-800 telephone number, which respondents could call for information about the study or to request a paper questionnaire. A preliminary sample of 40,000 households in mail-out areas was selected for this study. This sample was split randomly into two groups of 20,000 households each in order to create a control group [which received a paper questionnaire]. . . . The method was quite effective since the Internet response rate of the Push sample was 2.6 times more than the control group and 3.4 times more than the general population.

The Internet questionnaire used in the 2004 Canadian census test differed significantly from its paper counterpart in its approach to obtaining the basic resident count at a household. The paper questionnaire presents respondents with a set of detailed instructions of who should and should not be included in a household count and then asks for a roster of names. However, the Internet version asked respondents to complete a roster first and then used three follow-up questions—based on the instructions from the paper form—to guide respondents through the process of excluding temporary residents or foreign visitors from the final roster (Larouche, 2005). Whether this feature was also implemented in the final 2006 census Internet instrument is unclear.

[7] These editing steps are described in Statistics Canada summary of changes in the 2006 census, available at http://www.statcan.ca/english/sdds/document/3901_D17_T9_V1_E.pdf.

Deemed a success in 2006, the online response option is slated for use in the 2011 Canadian census, with the hope of boosting online response to as much as 40 percent. Though definitely not a set policy, Duquet and Gilmour (2007) suggest Statistics Canada's eventual vision for Internet collection in the census, in which an invitation to complete the census online (presumably with an Internet Access Code or the like) and in which a paper questionnaire is mailed only if the household specifically requests one or fails to respond to the initial invitation. Toward that end, Statistics Canada (2007) suggests that it may use its Push Strategy—tested in 2006—on a somewhat larger basis in 2011.

Japan

Alone in these examples—save for the U.S. 2010 census—Japan elected not to allow online response in its 2005 quinquennial census. For 2005, Kurihara (2004) reports that the Japanese Statistics Bureau sought to improve the information technology infrastructure of the census by rebuilding its internal geographic information system, testing the use of optical character recognition (OCR) of handwritten responses, and redesigning the user interface to obtain and work with small-area census data.[8]

New Zealand

Like the Australian census, the New Zealand quinquennial census is collected primarily by enumerators dropping off questionnaires and returning at a future date to collect them. Since 1996, New Zealand census questionnaires have been made available in an English-only or bilingual (English/Maori) version, the latter of which uses a "swim-lane" design that is a model for the bilingual English/Spanish form the Census Bureau plans to use in some areas in 2010. For 2006, to better meet perceived user needs, Statistics New Zealand planned an Internet response. However, it purposely did so without "attempt[ing] to leverage efficiency gains in any of the traditional census processes" or forecasting a desired Internet response rate target: plans were made to complete the census using traditional methods, and such responses by the Internet as were completed were deemed "a longer-term investment in improving participation" in later censuses (Smith, 2006). Furthermore, "it was recognized that there would not be financial savings in its implementation in the 2006 Census" (Statistics New Zealand, 2007).

In implementing the Internet response option, Statistics New Zealand (2007) decided not to aggressively promote the option. Instead, the agency chose to rely on limited promotion "through selected high-usage Internet sites only" and—principally—on

[8] On the second of these points—optical character recognition—it is worthwhile to note that this was a major test built into the conduct of the census itself. The specific objective was to determine whether completely automated OCR was sufficiently reliable or whether clerical checks of each questionnaire were still needed. One question—the destination of one's commute to work—was chosen for the automated testing since the seemingly "free" responses to this category were actually limited to the names of about 3,000 municipalities, making quality comparisons easier. Ultimately, it was concluded that "the accuracy of recognition was not sufficiently high" and that research on fully automated recognition would have to continue (Kurihara, 2004:4).

advocacy from the enumerators assigned to drop off the census forms. As part of their training, census enumerators were allowed to go through the online response questionnaire themselves; this was deliberately done so that they would be familiar with the requirements and so could accurately inform people in their household workload of the capability to complete the form online. When they visited the households to drop off the questionnaire, they also offered an envelope containing an ePIN identification number in order to use the Internet response option.

The online questionnaire allowed respondents to use either English or Maori. As with the enumerator-dropoff-and-return Australian census, mobile phone text messages were sent to individual enumerators after Internet responses were received, so that those households could be removed from the enumerator's visit workload.

Statistics New Zealand (2007) concluded that "despite very low promotion . . . the online option was very successful, not only in terms of the uptake" (7 percent of responses, or about 400,000 forms, via the Internet) "but an almost completely trouble-free operation." The agency plans to use the Internet response option again—with more active promotion—in 2011.

Prior to implementing the online response option in 2006, the Internet option was included in field tests in March and November 2003 as well as the 2005 dress rehearsal.

Singapore

In 2000, Singapore transitioned from a traditional census model to a register-based approach. The Household Registration Database (HRD) was developed in 1996 from administrative records as well as 1990 census returns. Hence, the 2000 Singapore census became a sample survey, intended to cover 20 percent of the population, to ask for information not included in the basic register data. These data items included relationship between members of a household, religion, and transportation/commute mode. To carry out this smaller scale survey, the Singapore Department of Statistics adopted a multimode approach. Sample households were invited to complete the form online; if they did not do so by a particular cutoff date, then computer-assisted telephone interviewing (CATI) was attempted. Barring that, trained enumerators were sent out to conduct face-to-face interviews with households that were not reached by either electronic means.

As summarized in a discussion paper for a 2003 census conference,[9] the online response option required respondents to log in using a user id and password, presumably provided in a mailing or through other contact. Once logged on, "basic data already available in the pre-Census database would be displayed" and "the respondent would then proceed to fill up the rest of the questionnaire on-line." Provision was made for respondents to pause the interview, save their results, and return at a later time to complete the questions. "Simple on-line checks were included and respondents would be prompted to re-enter the data if the information is incorrect or inconsistent."

[9] http://www.ancsdaap.org/cencon2003/Papers/Singapore/Singapore.pdf.

Ultimately, about 15 percent of the households in the sample completed the 2000 census form online,[10] and the multimode approach was considered a success.

Spain

The 2001 Spanish decennial census incorporated two main technological developments in the area of response methodology. One was preprinting of some questionnaire items—including name, sex, birth date, and place of birth—based on entries in Padrón, the Spanish Population Register. Hence, for these questions, respondents confirmed or updated the entries rather than working from purely blank spaces. The second was an Internet response option.

The two technical changes interacted in defining the way respondents were authenticated in order to use the online questionnaire. Those users with no changes to make in the pre-printed Padrón data could enter two personalized "keys" included in the mailing with the census form; alternately, they could access the form if their web browser was equipped with a certain "electronic certificate"—essentially, a digital signature obtained through another agency of the government. Users who wished to update the Padrón information had to have this type of electronic certificate in order to use the Internet form (Moraleda, 2006).

The need for an electronic certificate played some role in dampening the response rate via the Internet. Only 1 percent of households (13,818) completed the form online, of which 29.9 percent authenticated using the certificate. More than this number of households—16,238—attempted to use the Internet census questionnaire to update their Padrón information but gave up because they lacked the requisite certificate (Moraleda, 2006).

The Internet questionnaire application was designed to accommodate completion of the form at multiple sittings: partial information could be saved and then revisited later before submitting a finished questionnaire. The Spanish Internet response option was also available in Spain's co-official languages as well as English, French, German, and Arabic.

Switzerland

Along with Spain, Switzerland was the other European census to permit online responses as part of its e-Census initiative for the first time in 2001. Buscher and Stamm (2001:1-2) credited the creation of a government "Service Centre" for managing information technology as a final impetus for allowing online responses—a decision made even though Swiss census officials knew that "only a minority of the Swiss population currently have Internet access." The Swiss Federal Statistical Office reasoned that "electronic communication options are increasingly expected by potential users" and that the "PR and advertising impact of an Internet solution would be highly beneficial for the Census." As in the New Zealand experience, the move was also made with gaining

[10] http://www.singstat.gov.sg/pubn/popn/c2000sr4/coverage.pdf.

experience with new technology as the guiding goal: "the purpose was to see how far using the Internet could boost the efficiency of data entry and data quality while possibly cutting costs."

Because the Swiss e-Census relied on Service Centre networks, eligibility to file under the e-Census was limited to those communes or regions that had already opted to use the Service Centre equipment; this represented about 90 percent of the total population. Online questionnaires could be administered in German, French, or Italian.

The Swiss online response form was launched on November 27, 2000, and was operated until March 25, 2001; Census Day in the 2000 Swiss census was December 5, 2000. Buscher and Stamm (2001:5) report that "apart from two minor down-times during the first few days of operations, due to high visitor numbers and a server configuration which had not yet been optimized, the e-census ran smoothly, with no security problems throughout the four-month operating period." In all, 281,000 questionnaires (4.2 percent return rate) were completed via the Internet—just under 90 percent of those received during the first three weeks of operation. However, Swiss census officials also found that the form had a curiosity factor: about 20 percent of hits on the questionnaire site seemed to be "tourists" who "wanted to have a quick look at the e-census without attempting to enter their data." Demographically, Internet responses from younger middle-class men were more likely than from other groups but not so much so as to suggests "a major 'digital divide' in Swiss society" (Buscher and Stamm, 2001:7). About 10 percent of visitors to the site were unable to successfully log in to fill out the data: Buscher and Stamm (2001:6) do not describe the log-in procedure, noting only that "while it guaranteed maximum security, was also fairly complicated."

United Kingdom

The initial design document for the 2011 decennial census of England and Wales (Office for National Statistics, 2004) signaled the intent to use an online response option. Adding the Internet option is considered a useful step in improving the overall response rate, but the Office for National Statistics (2004:10) recognizes that the option will not immediately cut the cost of the census:

> By increasing the take-up of Internet completion, real cost and time savings could be made by reducing the quantity of paper forms to be captured and processed. Although we would seek to maximize the Internet response in order to realize the potential savings there is no guarantee of success, particularly since among the hard-to-count populations (such as the elderly) there would be significantly lower levels of take-up.

The Office for National Statistics conducted its first major pre-2011 field test in May 2007 with a sample of about 100,000 households. A major focus of the test was to evaluate new residence and national identity questions. However, the 2007 test did not include an Internet response option. A "frequently asked questions" list for the 2007 test

posted on the Office for National Statistics website explained that, "as this is a Census Test, resources are limited especially for the large expense to provide a facility to complete the questionnaire online." Nonetheless, the user was reassured that "it is proposed that a facility to complete the questionnaire online will be available for the Census in 2011."[11]

[11] http://www.statistics.gov.uk/census/2011Census/2011Project/pdfs/2007TestFAQsEnglish.pdf.

BOX B-1 United Nations Economic Commission for Europe Comments on Internet Data Collection in the 2010 Round of Censuses

Using the Internet as a collection method means that the census collection methodology will need to be self-enumeration rather than interview based. The Internet option can be incorporated into any of the traditional methods of delivering and collecting census forms (for example drop-off/pick-up, mail-out, mail back). The key factor is managing collection control operations—that is ensuring that every household and individual is counted once and once only. This requires the ability to provide each household and individual with a unique code linked to a geographic location. An added complication for those countries where forms are collected by census enumerators (rather than mailed back) is to have adequate and timely feedback to enumerators so that they can update their own collection control information so that they do not visit households that have already returned forms.

The potential level of take-up of an Internet option should be considered by assessing the proportion of the population who can access the internet from home, the proportion who use broadband services and the general use of the Internet for other business purposes (for example on-line banking, filing tax forms, shopping). The use of the Internet is likely to increase the cost of the census, at least initially. As it is not known in advance who is likely to use the Internet, there will be a need to deliver a paper form to every household including those who will subsequently use the Internet. Systems and processes that allow for Internet return of census forms will also need to be developed. These will increase costs. On the other side there are potential savings in data capture costs. However, scanning and Intelligent Character Recognition are in themselves cost efficient. Therefore, savings in data capture costs are likely to be considerable less than the costs of developing and implementing the internet system.

Security is an important consideration. Industry standard encryption (SSL128) offers two-way encryption (that is it encrypts data flowing both from and to the user's computer) and has been accepted by nearly all countries as adequate to protect the census information. Security should be a key consideration in designing the infrastructure. A physically separate infrastructure should be set up to collect the census information. Completed individual census forms should be moved behind firewalls and then into infrastructure that is completely separate from the collection infrastructure.

A downloadable on-line form requires much less infrastructure than for forms that are completed on line. However, downloadable forms require a greater level of computer literacy than on-line forms. They will not necessarily work in thousands of different computer configurations and there will be an expectation that the census agency will be able to deal with each individual problem. From the respondents' point of view, they are much more likely to prefer completing the form on-line. For these reasons it is expected that most countries will adopt on-line completion of census forms.

An electronic form offers the possibility of interactive editing to improve response quality that is not possible on a paper form. People using electronic forms have a certain

level of expectation that a certain amount of guidance will be offered—at a minimum that they will be sequenced through the form and not asked questions that are not relevant to their situation. How far other editing or on-line coding is built in to the form needs to be carefully considered. Some limited studies indicate that forms returned by the Internet are of higher quality than paper forms. More work is required in this area to determine whether this is a function of the type of people using the Internet or the technology itself.

Providing an Internet option may contribute to improving the quality of the census by making it easier for some hard-to-enumerate groups to respond. Most countries report difficulties in enumerating young adults and people living in secured accommodation where access is restricted. Some people with disabilities will also find it easier to complete an Internet form than a paper form. These groups are also more likely to be using the Internet and, if available, this option should be promoted to these groups as a means of encouraging participation in the census.

Provision of sufficient infrastructure provides one of the major challenges for offering an Internet option. The census occurs over a relatively short period of time and affects the whole population of a country, and it is unlikely that the census agency will have adequate infrastructure to cope with the peak demands of a census. It is therefore likely that this component, at least, of the Internet solution will be outsourced. It may be necessary for collection procedures to be modified to constrain demand. For example, requiring people outside predetermined target populations/areas to contact the census agency before they can use the Internet form may be a means of restricting use of the Internet form. Census agencies need to assess how they wish to promote the use of the Internet. Promotion of the Internet option should be determined by the capacity of the service to handle the expected load and should be coordinated with the collection procedures. The public relations strategy will need to encompass assurance about security of information supplied via the Internet. Assuming that the Internet option is targeted to the whole population, the public relations strategy should encompass managing public expectations about the ability to access the site during periods of peak demand. Simple messages advising people to use the internet option at “off peak” times should be prepared and used if necessary on the census internet site itself and through the census telephone inquiry service, radio and print.

SOURCE: Excerpted from United Nations Economic Commission for Europe (2006:Paragraphs 119-125).

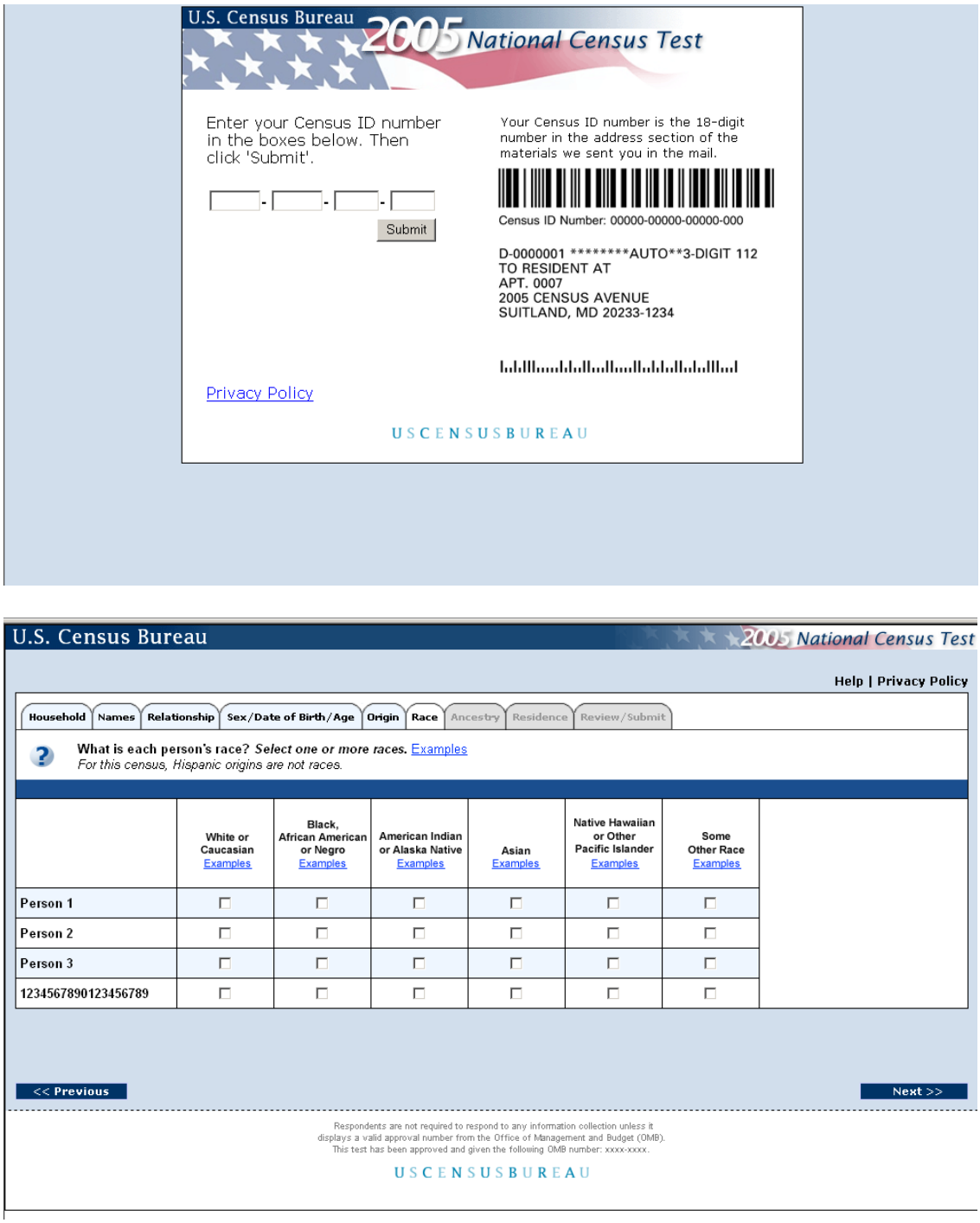


Figure B-1 Housing Unit ID log-in screen and race response screen, Internet questionnaire, 2005 census test.

SOURCE: http://www.census.gov/procur/www/2010dris/omb-person-based-screens.doc.

APPENDIX C

BIOGRAPHICAL SKETCHES OF PANEL MEMBERS AND STAFF

Lawrence D. Brown (*Chair*) is the Miers Bush professor in the Department of Statistics of the Wharton School of Business at the University of Pennsylvania. His research has been in the areas of statistical decision theory, statistical inference, nonparametric function estimation, foundations of statistics, sampling theory (census data), and empirical queueing science. He is a member of the National Academy of Sciences. At the National Research Council, he was a member of the Committee on National Statistics and served on its Panel to Review the 2000 Census; he also served on the Commission on Physical Sciences, Mathematics, and Applications and its Board on Mathematical Sciences. He is a fellow of the American Statistical Association and is a fellow and past president of the Institute of Mathematical Statistics. He has a B.S. from the California Institute of Technology and a Ph.D. from Cornell University.

Richard Berk is professor of criminology and statistics at the University of Pennsylvania. He is active regarding a range of methodological concerns, such as causal inference, statistical learning, and methods for evaluating social programs. His main areas of research include the inmate classification and placement systems, law enforcement strategies for reducing domestic violence, the role of race in capital punishment, detecting violations of environmental regulations, claims that the death penalty serves as a general deterrent, and forecasting short-term changes in urban crime patterns. Currently, he is working on a project on the development and application of statistical learning procedures for data sets in the behavioral, social, and economic sciences. He was elected to the Sociological Research Association and is a fellow of the American Association for the Advancement of Science, the American Statistical Association, and the Academy of Experimental Criminology. He was awarded the Paul S. Lazarsfeld Award for methodological contributions from the American Sociological Association. For the National Research Council, he has served on the Committee on Applied and Theoretical Statistics, the Panel on Monitoring the Social Impact of the AIDS Epidemic, the Working Group on Field Experimentation in Criminal Justice, and the Panel on Sentencing. He has a B.A. from Yale University (1964) and a Ph.D. from Johns Hopkins University (1970).

Eric Bradlow is the K.P. Chao professor of marketing, statistics, and education in the Wharton School of Business of the University of Pennsylvania. He previously held positions at the Educational Testing Service and at E.I. DuPont de Nemours and Company. He serves as associate editor for the *Journal of Computational and Graphical Statistics*, *Marketing Science*, *Quantitative Marketing and Economics*, and *Psychometrika*, and as senior associate editor for the *Journal of Educational and Behavioral Statistics*. He has won numerous teaching awards, and his research interests include Bayesian modeling, statistical computing, and developing new methodology for unique data structures. His current projects center on optimal resource allocation, choice modeling, and complex latent structures. He is a fellow of the American Statistical Association. At the National Research Council, he was a member of the Panel to Review the U.S. Department of Agriculture's Measurement of Food Insecurity and Hunger. He

has a B.S. from the University of Pennsylvania (1988) and A.M. (1990) and Ph.D. (1994) degrees, the latter in mathematical statistics, from Harvard University.

Michael L. Cohen (*Co-Study Director*) is a senior program officer for the Committee on National Statistics, currently serving as study director for the Panel on the Functionality and Usability of Data from the American Community Survey and the Panel on Coverage Evaluation and Correlation Bias in the 2010 Census. 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 in 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 has a B.S. in mathematics from the University of Michigan and M.S. and Ph.D. degrees in statistics from Stanford University.

Daniel L. Cork (*Co-Study Director*) is a senior program officer for the Committee on National Statistics, currently serving as study director of the Panel to Review the Programs of the Bureau of Justice Statistics. Previously he served as study director of the Panel on Residence Rules in the Decennial Census, co-study director of the Panel on Research on Future Census Methods, and program officer for the Panel to Review the 2000 Census. His research interests include quantitative criminology, particularly space-time dynamics in homicide; Bayesian statistics; and statistics in sports. He has a B.S. in statistics from George Washington University and an MS. in statistics and a joint Ph.D. in statistics and public policy from Carnegie Mellon University.

Ivan Fellegi is chief statistician of Canada and head of Statistics Canada. In 1992 he was made a member of the Order of Canada and was promoted to officer in 1998. He has served as president of the International Association of Survey Statisticians, the International Statistical Institute, and the Statistical Society of Canada. Besides being an elected member of the International Statistical Institute, he is also a fellow of the Royal Statistical Society and the American Statistical Association. He has provided advice on statistical matters to his native Hungary following its transition to democracy and, in 2004, was awarded the Order of Merit of the Republic of Hungary. He has chaired the Conference of European Statisticians of the United Nations Economic Commission for Europe. He has been awarded the Gold Medal by the Statistical Society of Canada and the Robert Schuman medal by the European Community and the Order of Canada. He has published extensively in the areas of census and survey methodology, in particular on consistent editing rules and record linkage. At the National Research Council, he was a member of the Panel on Privacy and Confidentiality as Factors in Survey Response, the Panel on Census Requirements in the Year 2000 and Beyond, and the Panel on Decennial Census Methodology. He has a B.Sc. from the University of Budapest (1956) and an M.Sc. (1958) and a Ph.D. in survey methodology (1961) from Carleton University.

Linda Gage is the liaison to demographic programs at the California Department of Finance. She represents California in federal and profession forums and evaluates the effect of various demographic and statistical programs on the state. Previously, she served as the California state demographer for two decades and in other positions in the Department of Finance since 1975. She has served on the U.S. secretary of commerce's Decennial Advisory Committee since 1995.

At the National Research Council, she was a member of the Panel on a Review of Statistical Issues in the Allocation of Federal and State Program Funds. She has B.A. and M.A. degrees in sociology, with emphasis in demography, from the University of California, Davis.

Vijay Nair is the Donald A. Darling professor of statistics and professor of industrial and operations engineering at the University of Michigan. He has been chair of the Statistics Department since 1998. He was a research scientist at Bell Laboratories for 15 years before joining the faculty at Michigan. His area of expertise is engineering statistics, including quality and productivity improvement, experimental design, reliability, and process control. He is a fellow of the American Association for the Advancement of Science, the American Statistical Association, and the Institute of Mathematical Statistics, and an elected member of the International Statistical Institute. He is a former editor of *Technometrics* and *International Statistical Review* and has served on many other editorial boards. He is currently the chair of the board of trustees of the National Institute of Statistical Sciences. At the National Research Council, he is a member of the Committee on National Statistics and served on several study panels, including the Panel on Statistical Methods for Testing and Evaluating Defense Systems and the Assessment Panel on NIST's Information Technology Center, and he chaired the Oversight Committee for the Workshop on Testing for Dynamic Acquisition of Defense Systems. He has a Ph.D. in statistics from the University of California, Berkeley.

Jesse Poore holds the Ericsson/Harlan D. Mills chair in software engineering in the Department of Computer Science at the University of Tennessee. He is also director of the University of Tennessee–Oak Ridge National Laboratory Science Alliance, a program to promote and stimulate joint research between those two organizations. He conducts research in cleanroom software engineering and teaches software engineering courses. He has held academic appointments at Florida State University and the Georgia Institute of Technology; he has served as a National Science Foundation rotator, worked in the Executive Office of the President, and was executive director of the Committee on Science and Technology in the U.S. House of Representatives. He is a member of the Association for Computing Machinery and the Institute of Electrical and Electronics Engineers and is a fellow of the American Association for the Advancement of Science. At the National Research Council, he was a member of the Panel on Statistical Methods for Testing and Evaluating Defense Systems and the Oversight Committee for the Workshop on Testing for Dynamic Acquisition of Defense Systems. He has a Ph.D. in information and computer science from the Georgia Institute of Technology.

Stanley Smith is a professor of economics and director of the Bureau of Economic and Business Research at the University of Florida. He is also director of its population program, which produces the official state and local population estimates and projections for the state of Florida. He is Florida's representative to the Federal-State Cooperative Program for Population Estimates and Projections and a past president of the Southern Demographic Association. He has also served on the U.S. Census Bureau's Decennial Advisory Committee. His research interests include the methodology and analysis of population estimates; he has done particular work on the measurement of seasonal populations. At the National Research Council, he was a member of the Panel on Alternative Census Methodologies. He has a B.A. in history from Goshen College and a Ph.D. in economics from the University of Michigan (1976).

Roger Tourangeau is director of the Joint Program in Survey Methodology at the University of Maryland and a senior research scientist at the University of Michigan. Previously, he was a senior methodologist at the Gallup Organization, where he designed and selected samples and carried out methodological studies, and founded and directed the Statistics and Methodology Center of the National Opinion Research Center. His research focuses on attitude and opinion measurement and on differences across methods of data collection; he also has extensive experience as an applied sampler and is known for his work on the cognitive aspects of survey methodology. He is a fellow of the American Statistical Association and has served on the editorial board of *Public Opinion Quarterly*, as well as on Census Bureau advisory panels. At the National Research Council, he was a member of the Panel on Residence Rules in the Decennial Census. He has a Ph.D. in psychology from Yale University.

Nora Cate Schaeffer is professor of sociology at the University of Wisconsin, Madison. Her areas of expertise include respondent behavior and interviewer-respondent interaction. Her past research has concentrated on a number of different areas in survey methodology dealing with nonsampling error, both nonresponse and response errors of various kinds. She was on the editorial board of *Public Opinion Quarterly*, *Sociological Methodology*, and *Sociological Methods Research*. At the National Research Council, she is a member of the Committee on National Statistics. She has an A.B. from Washington University and a Ph.D. in sociology from the University of Chicago.

Allen Schirm is a senior fellow and associate director at Mathematica Policy Research, Inc. Previously, he was Andrew W. Mellon assistant research scientist and assistant professor at the University of Michigan. His principal research interests include small-area estimation and sample and evaluation design, with application to studies of child well-being and welfare, food and nutrition, and education policy. At the National Research Council, he was a member of the Panel on Research on Future Census Methods, the Panel on Formula Allocations, and the Panel on Estimates of Poverty for Small Geographic Areas. He has an A.B. in statistics from Princeton University and a Ph.D. in economics from the University of Pennsylvania.

Judith A. Seltzer is professor of sociology at the University of California, Los Angeles. Previously, she was on the faculty of the University of Wisconsin, where she contributed to the development and implementation of the National Survey of Families and Households. Her research interests include kinship patterns, intergenerational obligations, relationships between nonresident fathers and children, and how legal institutions and other policies affect family change. As part of a cross-university consortium of researchers, she is developing new models for explaining family change and variation, in which family dynamics and residence patterns will be important components. She has also participated in the Los Angeles Family and Neighborhood Survey. She has collaborated on research to improve the quality of data on children's living arrangements, transfers, and contact with nonresident parents using information from surveys and administrative data. At the National Research Council, she was a member of the Panel on Residence Rules in the Decennial Census. She has master's and Ph.D. degrees in sociology from the University of Michigan.

Kirk Wolter is senior fellow and director of the Center for Excellence in Survey Research at the National Opinion Research Center, where he has also served as senior vice president for statistics

and methodology. He is also professor of statistics, part time, at the University of Chicago. During his career, he has led or participated in the design of many of America's largest information systems, including the Current Business Surveys, the Current Employment Statistics program, the Current Population Survey, the 1980 and 1990 decennial censuses, the National 1997 Longitudinal Survey of Youth, and the National Resources Inventory. He is a fellow of the American Statistical Association, an elected member of the International Statistical Institute, and past president of the International Association of Survey Statisticians and of the Survey Research Methods section of the American Statistical Association. At the National Research Council, he was a member of the Panel on Conceptual, Measurement, and Other Statistical Issues in Developing Cost of Living Indexes and is currently a member of the Panel on Measuring Business Formation, Dynamics, and Performance. He has an M.A. and a Ph.D. in statistics, both from Iowa State University.