The American Community Survey: Summary of a Workshop

DETAILS

72 pages | 6 x 9 | PAPERBACK
ISBN 978-0-309-07315-8 | DOI 10.17226/10051

AUTHORS

Committee on National Statistics, National Research Council

The American Community Survey

Summary of a Workshop

Committee on National Statistics

Commission on Behavioral and Social Sciences and Education

National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C.

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.

Support of the work of the Committee on National Statistics is provided by a consortium of federal agencies through a grant from the U.S. National Science Foundation (Number SBR-9709489). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number 0-309-07315-4

Additional copies of this report are available from National Academy Press, 2101 Constitution Avenue, N.W., Lockbox 285, Washington, D.C. 20055. Call (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area). This report is also available online at http://www.nap.edu

Printed in the United States of America

Suggested Citation: National Research Council (2001) *The American Community Survey: Summary of a Workshop.* Committee on National Statistics. Washington, DC: National Academy Press.

THE NATIONAL ACADEMIES

National Academy of Sciences
National Academy of Engineering
Institute of Medicine
National Research Council

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

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

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

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

COMMITTEE ON NATIONAL STATISTICS

AMERICAN COMMUNITY SURVEY WORKSHOP

Participants

*JOHN ROLPH (*Chair*), University of Southern California
KATHARINE ABRAHAM, Bureau of Labor Statistics, U.S. Department of Labor
BASHIR AHMED, U.S. Census Bureau
*CHARLES ALEXANDER, U.S. Census Bureau
BARBARA BAILAR, University of Chicago
DONALD BAY, U.S Department of Agriculture
*ROBERT BELL, RAND, Santa Monica, CA
*WILLIAM BELL, U.S. Census Bureau
BENNETT BERTENTHAL, Division of Social and Economic Sciences, U.S. National Science Foundation
*JAY BREIDT, Iowa State University
SHAIL BUTANI, Bureau of Labor Statistics, U.S. Department of Labor
WILLIAM BUTZ, Division of Social and Economic Sciences, U.S. National Science Foundation
LYNDA CARLSON, U.S. Department of Energy
STEVEN CARLSON, U.S. Department of Agriculture
JAN CHAIKEN, Bureau of Justice Statistics, U.S. Department of Justice
CYNTHIA CLARK, U.S. Census Bureau
STEVE COHEN, Bureau of Labor Statistics, U.S. Department of Labor
SEAN CREIGHTON, Education Statistics Services Institute, Washington, DC
SCOT DAHL, U.S. Census Bureau
MARSHALL DEBERRY, Bureau of Justice Statistics, U.S. Department of Justice
MIKE DEPIPO, Food and Nutrition Service, U.S Department of Agriculture
VIRGINIA DEWOLF, U.S. Office of Management and Budget
GREGG DIFFENDAL, U.S. Census Bureau
*CATHRYN DIPPO, Bureau of Labor Statistics, U.S. Department of Labor
ALAN H. DORFMAN, Bureau of Labor Statistics, U.S. Department of Labor
LUCY ELDRIDGE, Bureau of Labor Statistics, U.S. Department of Labor
TRENA EZZATI-RICE, National Center for Health Statistics, U.S. Department of Health and Human Services
RON FECSO, Division of Social and Economic Science, U.S. National Science Foundation
STEPHEN FIENBERG, Carnegie Mellon University

*Workshop speaker

CAROL FROST, Congressional Budget Office, U.S. Congress
WAYNE FULLER, Iowa State University
GERALD GATES, U.S. Census Bureau
MARY GOLLADAY, Division of Social and Economic Sciences, U.S. National Science Foundation
SUSAN GRAD, U.S. Social Security Administration
DANIEL HESS, Bureau of Transportation Statistics, U.S. Department of Transportation
HUU HOANG, U.S. Census Bureau
ELAINE HOCK, U.S. Census Bureau
PAUL HSEN, Bureau of Labor Statistics, U.S. Department of Labor
BRADFORD HUTHER, U.S. Census Bureau
TOM JABINE, Consultant, Washington, DC
KEN JOHNSON, Bureau of Economic Analysis, U.S. Department of Commerce
THOM JURKIEWICZ, House Subcommittee on the Census, U.S. Congress
*GRAHAM KALTON, Westat, Inc., Rockville, MD
DANIEL KASPRZYK, National Center for Education Statistics, U.S. Department of Education
CHARLES KINDERMAN, Bureau of Justice Statistics, U.S. Department of Justice
NANCY KIRKENDALL, U.S. Office of Management and Budget
ROBERT KNISELY, Bureau of Transportation Statistics, U.S. Department of Transportation
ROBERT LATTA, Energy Information Administration, U.S. Department of Energy
NANCY LEACH, U.S. Department of Energy
DANIEL LEVINE, Westat, Inc., Rockville, MD
*DENISE LEWIS, U.S. Census Bureau
JIM LIN, U.S. Census Bureau
*RODERICK LITTLE, University of Michigan, Ann Arbor
*TOM LOUIS, University of Minnesota
TERRIANN LOWENTHAL, Consultant, Washington, DC
TIMOTHY MANEY, House Subcommittee on the Census, U.S. Congress
ROSEMARY MARCUSS, Bureau of Economic Analysis, U.S. Department of Commerce
DAVID MARKER, Westat, Inc., Rockville, MD
EDITH MCARTHUR, National Center for Education Statistics, U.S. Department of Education
LARRY MCGINN, U.S. Census Bureau
DAVID MCMILLAN, House of Representatives, U.S. Congress
DAN MELNICK, Dan Melnick Research Inc., Washington, DC

KATHLEEN NEBEKER, National Institute on Aging, U.S. Department of Health and Human Services
JEFFREY NEWMAN, Bureau of Economic Analysis, U.S. Department of Commerce
MARK NORD, Economic Research Service, U.S. Department of Agriculture
WENDE O'NEILL, Bureau of Transportation Statistics, U.S. Department of Transportation
CORDELIA REIMERS, U.S. Council of Economic Advisers
CHRISTEN ROBINSON, National Institute on Aging, U.S. Department of Health and Human Services
ELIZABETH ROGERS, Bernan Associates, Washington, DC
PHILIP L. RONES, Bureau of Labor Statistics, U.S. Department of Labor
PATRICIA RUGGLES, U.S. Department of Health and Human Services
FRANK SAMANIEGO, University of California, Davis
ALLEN SCHIRM, Mathematica Policy Research, Inc., Washington, DC
PAULA SCHNEIDER, U.S. Census Bureau
KATHLEEN SCHOLL, U.S. Government Accounting Office
RONALD SEPANIK, U.S. Department of Housing and Urban Development
GARY SHAPIRO, Westat, Inc., Rockville, MD
*ERIC SLUD, University of Maryland
AMY SYMENS SMITH, U.S. Census Bureau
GEORGE SMITH, Bureau of Economic Analysis, U.S. Department of Commerce
EDWARD SONDIK, National Center for Health Statistics, U.S. Department of Health and Human Services
EDWARD SPAR, Council of Professional Associations on Federal Statistics, Alexandria, VA
ANGELA STACK, U.S. Census Bureau
LYNNE STOKES, University of Texas, Austin
MICHAEL STOTO, George Washington University
RICHARD SUZMAN, National Institute on Aging, U.S. Department of Health and Human Services
CYNTHIA TAEUBER, U.S. Census Bureau
TONY TESINE, U.S. Census Bureau
HELEN TICE, Bureau of Economic Analysis, U.S. Department of Commerce
NANCY TORRIERI, U.S. Census Bureau
CLYDE TUCKER, Bureau of Economic Analysis, U.S. Department of Commerce
DAVID VANDENBROUKE, U.S. Department of Housing and Urban Development
PAUL VOSS, University of Wisconsin, Madison
KATHERINE WALLMAN, U.S. Office of Management and Budget

*LYNN WEIDMAN, U.S. Census Bureau
DANIEL WEINBERG, U.S. Census Bureau
SIGNE WETROGAN, U.S. Census Bureau
JANET WILLIAMS, Bureau of Labor Statistics, U.S. Department of Labor
MARTYNAS YCAS, U.S. Social Security Administration, Washington, DC
*ALAN ZASLAVSKY, Harvard Medical School

Acknowledgments

As workshop chair, I would like thank various individuals that contributed to the success of the American Community Survey (ACS), Workshop. Most importantly, I am extremely grateful to the "thought-piece" authors and the discussants for outstanding papers and presentations that contained very useful suggestions for both how to conceptualize various methodological problems and for suggesting methods that could be used to address these methodological problems. This work was also done under extreme time pressure, which makes the resulting contributions all the more remarkable. The thought-piece authors and discussants were Tom Louis, Rod Little, William Bell, Eric Slud, Alan Zaslavsky, Allen Schirm, Robert Bell, Lynn Weidman, Cathryn Dippo, and Jay Breidt.

I would also like to thank the special guests who contributed greatly to the floor discussion at the workshop: Barbara Bailar, Stephen Fienberg, Wayne Fuller, Tom Jabine, Frank Samaniego, Lynne Stokes, and Paul Voss.

The planning of the workshop was supervised by a subcommittee of the Committee on National Statistics, which provided excellent advice on the agenda, thought-piece authors, discussants, and special guests. The subcommittee members were John Rolph, Rod Little, Tom Louis, Paul Rosenbaum, Keith Rust, and Janet Norwood.

Michael Cohen, as study director for the ACS workshop, did the lion's share of the work in designing the workshop and in securing the participation of thought-piece authors, discussants, and special guests. Most critically, his discussion paper articulated the important issues that will need to be addressed in implementing the ACS. The paper served as a starting point for the contributions of the thought-piece authors and the discussants.

The staff received excellent cooperation from the Census Bureau, especially from Charles Alexander, Robert Fay, and Lynn Weidman for help in identifying the problems that were used to focus the workshop, and most especially from Charles Alexander for writing the response to the staff paper used to provide direction to the thought-piece authors, and for giving several excellent short presentations at the workshop to provide the Census Bureau's current methods and perspective on future methods relative to each issue prior to each session's thought-piece presentation. Charles Alexander also was fully accessible to the staff and was involved in myriad e-mail and telephone conversations concerning all aspects of the workshop. It is clearly the case that the workshop could not have been held without his excellent contributions in each area of the workshop.

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

We thank the following individuals for their participation in the review of this report: James R. Chromy, Statistics Research Division, Research Triangle Institute, Research Triangle Park, NC; Brenda Cox, Mathematica Policy Research, Inc., Washington, DC; James Lepkowski, Institute for Social Research, University of Michigan; Paul Rosenbaum, Department of Statistics, Wharton School, University of Pennsylvania; Rob Santos, Urban Institute, Washington, DC; and Kirk Wolter, National Opinion Research Center, University of Chicago.

Although the reviewers listed above have 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 John F. Geweke, Department of Economics, University of Iowa. Appointed by the Commission on Behavioral and Social Sciences and Education, 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.

I am also indebted to the staff of the Committee on National Statistics. In particular, I would like to thank Miron Straf for his leadership and commitment to methodologically oriented activities such as this one. Eugenia Grohman not only carried out her usual exemplary technical editing but was also instrumental in restructuring the report, which greatly improved its clarity. I would like to thank Agnes Gaskin for her careful attention to detail with respect to arrangements and her pleasant and professional style. Finally, I would like to thank Meyer Zitter and Michael Cohen, who drafted the initial version of the workshop report, and Dannelle Dessaint for her work on report production.

John E. Rolph, *Chair*
American Community Survey Workshop and
Committee on National Statistics

Contents

1

Introduction

The American Community Survey (ACS), to be run by the Census Bureau, will be a large (250,000 housing units a month), predominantly mailout/mailback survey that will collect information similar to that on the decennial census long form. The development of this new survey raises interesting questions about methods used for combining information from surveys and from administrative records, weighting to treat nonresponse and undercoverage, estimation for small areas, sample design, and calibration of the output from this survey with that from the long form. To assist the Census Bureau in developing a research agenda to address these and other methodological issues, the Committee on National Statistics held a workshop on September 13, 1998. This report summarizes that workshop.

When fully operational (currently planned for 2003), the ACS will provide continuous, small-area information on demographic characteristics, social welfare, education and health status, commuting patterns, crime patterns, and other important attributes of the population of the United States, including the interrelationships of these characteristics. Unlike any other national survey, the ACS will provide information on the American population at substate levels, i.e., counties and cities. Over a 5-year period, the survey's sample size will approximate that of the census long form, supporting the production of estimates, possibly through use of statistical modeling, for small and nonstandard geographical areas, such as school districts and traffic analysis zones. In addition, given the sample size, information will be available for specific demographic groups, including racial and ethnic groups, children, the elderly, people in specific occupations, people with specific health conditions, and people with various levels of educational attainment. The ACS questionnaire may also be able to address current issues of national and regional importance by including supplementary questions (though this raises various complications). Obviously, since the ACS data will be collected continuously throughout the decade, the information will generally be more timely than that from the census long form.

The three primary types of information sources for individuals and households in the federal statistical system are the decennial census, various household surveys, and administrative records systems, e.g., tax and food stamp records. The planned ACS has various advantages in comparison to these other sources. With respect to household surveys, the ACS will have advantages of size and scope: it will sample more households and will provide a wider range of information than is typically found on individual household surveys, which are by necessity targeted to a specific set of issues. With respect to administrative records systems, the ACS will have the advantages of representativeness (e.g., tax records have information only on filers, and food stamp records have information only on participants) and scope—since administrative records do not contain information other than that required for the associated program (e.g., they typically do not include demographic information). Also, the associated programs can change from year to year, making interyear comparisons difficult, and the administration of a program may differ by geographic region, especially by state, which complicates interregional comparisons.

However, these other information sources also have advantages in comparison to the ACS. It is expected that the information from household surveys will generally be of higher quality than that from the ACS for the outputs directly related to the purpose of the household survey. This is due to at least three factors: (1) many household surveys use personal interviews as a primary source of data collection, which tends to result in higher-quality responses than for mailout-mailback surveys; (2) many household surveys visit the same households over time, which can help to improve response; and (3) household surveys targeted on a specific characteristic typically ask a large number of questions concerning that characteristic, which can aid in recall and in the accuracy of response. Therefore, ACS responses are likely to have greater measurement error than household surveys requesting the same information. Also, administrative records systems, which are collected as a byproduct of the associated programs, can provide more detailed information and often provide information for larger numbers of individuals or families than will be possible for ACS.

From the time of the initial planning of the ACS (initially referred to as continuous measurement), the Census Bureau has been aware that its introduction would raise a number of complicated methodological issues. However, given the substantial work entailed in the fielding of such a new, large survey, the focus of the ACS group to date has been on refining data collection, leaving the final answers to the difficult analysis questions for later. Thus, procedures for nonresponse and undercoverage adjustment were modeled, to the extent possible, after current procedures used for the census long form. Now that data collection has matured as the ACS demonstration phase is well under way, the Census Bureau is developing a research plan and initiating research to address all issues related to ACS methodology. Fall 1998 therefore seemed an opportune moment for a workshop to assist the Census Bureau in developing a research agenda to deal with many of these challenging issues. The hope was that the workshop would facilitate a dialogue between Census Bureau staff, interested academics, and other researchers.

DATA COLLECTION AND SAMPLE DESIGN

The idea for the ACS follows a suggestion made nearly two decades ago by Leslie Kish (1981), who used the term "rolling samples." The survey instrument for the ACS is a questionnaire that will be mailed out to households, with 100 percent nonresponse follow-up using computer-assisted telephone interviewing (CATI) in the month following questionnaire mailout, followed by field follow-up (computer-assisted personal interviewing, or CAPI) of a random one-third of the remaining nonrespondents in the month following CATI follow-up. The size of this survey will make direct small-area estimates possible, though the estimates for smaller areas typically will be produced by aggregating information over 2 to 5 years, depending on the size of the area. (At this time moving averages are planned.) The current plan also is that governmental jurisdictions of less than 2,500 population will be oversampled.

Each month's sample is intended to be a self-weighting sample of the population of each area of the United States (except for possibly a few complications such as oversampling small areas). Thus, cluster sampling will not be used to facilitate field follow-up. Also, both to reduce respondent burden and to lower variances of direct estimates aggregated over months and years, a household cannot be in the sample more than once every 5 years. When the ACS is fully fielded, it will use as a sampling frame the Census Bureau's Master Address File, an update of the address list used to conduct the 2000 census. The annual sample will be divided into monthly mailout panels, where each month's panel is a systematic sample across the complete address list.

As mentioned above, the CATI follow-up of mail nonresponse lags 1 month from the questionnaire mailout, and the field follow-up lags 1 month from the CATI operation. To accommodate this schedule and provide timely estimates, a given month's estimates will make use of the totality of information *collected in* that month, which will not be the information corresponding to the systematic sample selected for that month. The information collected in a month will include mail responses received in that month, the CATI interviews relative to the sample for the previous month, and the information from the field follow-up relative to the sample for the month before that. This treatment of nonresponse, along with undercoverage and other complications, raises difficult weighting challenges.

PILOT TESTING

There was a 3-year demonstration pilot test for the ACS in 1996-1998. The survey was administered in Brevard County, FL; Multnomah County, OR; Rockland County, NY; Fulton County, PA; Franklin County, OH; Douglas County, NE; Ft. Bend and Harris Counties, TX; and Otero County, NM. The sampling rate for the first year was 15 percent in most areas with 30 percent in small governmental areas. The rate was lowered to 3 percent during the latter part of the demonstration period. During 1998, the decennial census dress rehearsal site in South Carolina also implemented the demonstration ACS. The goals of the demonstration period were to: (1) illustrate the usefulness of ACS data every year and over time; (2) to improve operations and reduce and understand costs; and (3) to provide a comparison with the dress rehearsal data.

The demonstration period is now to be followed by two comparison studies—in 1999-2001 and 2000-2002—comparing ACS and census long-form information. Full implementation of ACS will begin in 2003, with an ongoing sample size of 3 million housing units a year (a sampling rate of approximately 3%—15% over five years—compared with 17%, on average, for the census long form). The first comparison study was based on an implementation of ACS in 31 comparison sites for 3 years, 1999-2001. An experimental design was used to select the 31 areas with characteristics for which differences between the ACS and long-form responses were anticipated. The sampling rate in the sites was planned to be roughly 5 percent annually so that the sample sizes of the ACS (over the 3-year period) and the long form would be comparable, though budget limitations have reduced this sampling rate in some areas.

For the second comparison study, in 2000-2002, the ACS will have a national sample of 700,000 addresses per year (0.7% sampling rate). This study is designed to make comparisons between the long form and ACS for all states, large metropolitan areas, large substate areas, and population groups.

The objective of the 1999-2001 comparison is to understand the factors associated with the differences between the 1999-2001 ACS and the 2000 long form in the 31 areas, using the second comparison study to develop a calibration model to adjust the 2000 long-form estimates to roughly represent what the full ACS would have yielded in 2000. The adjustment based on this calibration model will reflect differences in question wording, residence rules, reference periods, interviewer training and other field operations, and differences in coverage and nonresponse. Once adjusted, the "calibrated" long-form data for 2000 can be compared with ACS data that are collected following full field implementation in 2003, in order to understand the dynamics over time of such characteristics as poverty and employment.

THE WORKSHOP: PURPOSE AND STRUCTURE

To maximize the interchange of ideas during the limited time available in a 1-day workshop, selected experts were asked to prepare thought pieces to address one of six methodological issues: combination of information cross-sectionally, combination of information across time, the impact of variance on ACS outputs as inputs into fund allocation formulas, weighting to accommodate nonresponse, undercoverage, etc., issues related to sample and questionnaire design, and calibration of the ACS with the census long form. For some of the issues, discussants were also selected. As background for both the writers and discussants, committee staff prepared a document elaborating on the specific problems posed by each of the methodological issues, including research directions that might prove beneficial. Following this, Charles Alexander of the Census Bureau provided a "response" to the background document, and both were provided in advance to the writers and discussants. The staff document, the response from Charles Alexander, the thought pieces, and the discussant papers when completed were made available to all of the thought piece writers, discussants, and special invited guests in advance of the workshop, so that the floor discussion could be more informed as to each presenter's ideas. (These documents are available in the workshop agenda book, "The American Community Survey Workshop: Technical Papers.")

The next six chapters describe, in turn, the methodological areas of focus of the workshop. Although the subjects chosen for focus at the workshop are some of the more important methodological issues facing the Census Bureau, it is important to note that these issues do not encompass all the issues of concern and that there are many others worthy of study. Even in these six methodological areas of focus, the presenters were free to choose to address various subtopics. Some of the remaining issues not examined in the workshop are listed in Chapter 8.

To set the stage, and especially to illustrate concerns for the development of small-area estimates and issues related to the combination of information from different data sources, Graham Kalton started off the workshop by describing the work of the Panel on Estimates of Poverty for Small Geographic Areas (which he chairs). (For more information, see the Appendix and National Research Council, 2000.)

The remainder of the report has the following structure: Chapters 2 through 7 treat the individual methodological issues in turn, namely, combination of information across areas, combination of information across time, concerns involved in the use of the ACS as input to funding formulas, weighting issues in the ACS to treat nonresponse and undercoverage, sample and questionnaire design issues, and calibration of the ACS to the census long form. In each chapter, the topics are first introduced, followed by summaries of the presentations and the floor discussion; Chapter 8 provides some final comments.

2

Combination of Information Across Areas

Given its substantial sample size, for many purposes output from the ACS collected directly from a group or an area at a point in time will be adequate to provide useful estimates. However, the utility of the ACS will be greatly enhanced through its use in producing indirect estimates, i.e., estimates derived by combining information from other data sources or from the ACS for other time periods or for other geographic areas, through the use of statistical models. Using the census, administrative records, other household surveys, and now the ACS, statistical models hold the promise of providing timely estimates for smaller areas and groups than would otherwise be possible.

Combining information is an area in which statistics has recently made important advances (see, e.g., National Research Council, 1992). This includes progress in empirical and hierarchical Bayes' modeling (facilitated by the advances in computation provided by Markov chain Monte Carlo methods), variance component modeling, small-area estimation, and time-series analysis, along with advances in generalized linear models (GLM). This is in addition to a greater understanding of how to accommodate complex sample designs using these techniques. While these advances have demonstrated wide utility, each individual application typically presents some novel complications. Especially given the variety and number of sources of information and the variety and number of different responses of interest (information on education, welfare, unemployment, income, health, etc.), understanding how to make use of these techniques in this setting presents a difficult challenge to the Census Bureau.

We separate this topic into two subtopics: combining information across areas for a single time period and combining information across multiple time periods (and across areas). This is mainly for convenience of discussion, since clearly both topics need to be considered simultaneously. This chapter concerns the former issue, and the next chapter concerns the latter.

Models that combine information from these various sources must take account of the following (and other) complications: the information sources to be combined could provide information for different populations (e.g., tax filers are not the same as residents of the United States), represent slightly different reference periods, and make use of different survey or data collection methods. Therefore, to combine estimates from these sources may require techniques that can combine estimates with measurement error and bias that are not well modeled or estimated. In addition, estimates are typically needed at different levels of geographic aggregation, such as national, state, county, and possibly lower (e.g., census tract) levels. Several questions concern the development of such models:

- What types of models are likely to be effective?
- How can estimates be combined with measurement error and other biases?
- At what level of aggregation should the modeling be done? For example, should estimates be modeled at the county level and then aggregated, or should estimates be modeled at the state level and, using simple types of models such as synthetic estimation or modeling county shares, passed down to counties?
- How can Bayes' (or related) methods be used to fold the direct estimates in with the model-based estimates?
- What complications are posed by the sampling weights, nonresponse, and undercoverage for each of the data sources?
- For some large areas, direct estimates from a relevant household survey are likely to be recognized as standard values given their lack of measurement error (but possibly appreciable variance, depending on the area), so agreement of indirect estimates incorporating ACS information with these standards would have the advantage of consistency with an accepted estimate. To address this, one possibility is to control the indirect ACS estimates to the standard estimates. Or one might try to use the ACS information to improve on these standard values. Both approaches, controlling and smoothing,[1] are complicated by the existence of several of these standard values. Should one consider modeling each ACS response separately, or is there some kind of

[1]The term "smoothing" is used to indicate a wide variety of techniques in which two or more estimates are combined through use of weighted averages in order to reduce variance.

very general missing data technique that would put each of the important household surveys together with the ACS, possibly into a multipurpose database that could simply be aggregated to provide estimates?

- An important special problem relevant to this last point is that of providing population estimates for demographic groups within counties. The ACS will provide information as to the size of these populations, and simply controlling ACS estimates to the existing population estimates ignores this important source of information. Therefore, how should ACS and population estimates be combined to produce better small-area population estimates?

RESEARCH DIRECTIONS

In his presentation on the topic, Tom Louis noted that direct estimation for small areas will typically be inferior to reasonable methods in which information is combined using statistical models, and in the ACS context, that will usually mean “borrowing strength” over geography, or time, or both. The determination of how to combine information typically involves a tradeoff of bias and variance. In making this tradeoff, Bayesian formalism effectively structures the integration of information and ensures that all uncertainties are captured by the posterior distribution.

This approach can combine information for relevant data sources and can properly account for missing data. Further, Bayesian methods can address nonstandard goals, which is relevant to a topic addressed in Chapter 3—the use of ACS-based estimates for input into fund allocation formulas, which often have nonstandard forms and therefore implicitly nonstandard loss functions for the associated estimates, e.g., fund allocation formulas that have eligibility thresholds.

Though the Bayesian approach has these and other attractive properties, due to the national importance of the ACS in providing estimates for various official purposes, its use in this context must have good frequentist properties (good objective performance) as well. A large body of literature validates the judgment that Bayesian methods used with care do have excellent, objective properties. In addition, they are no more complex than the application dictates, and they have the advantage of making all assumptions explicit.

Bayesian methods separate the two activities of summarizing information and using it to make inferences. Multiple goals, such as those governed through the use of point estimates, estimation of ranks, and estimation of the cumulative distribution function of the underlying parameters, can be addressed individually, or a single compromise inference or estimate can be used that performs well (but not necessarily optimally) for all goals. This point initiated discussion concerning the distinction between production of estimates for general use and production of estimates for specific purposes. The approach advocated might be more relevant to an estimate needed for input to specific fund allocation formulas. However, if an agency is producing estimates for broad national purposes, it is not clear that this approach is relevant.

In the particular application of small-area estimation involving the ACS, based somewhat on the experience of the small-area poverty estimates panel mentioned above, fixed-effects regression modeling combined with empirical and hierarchical Bayesian and random-effects modeling should be very effective in a wide variety of specific problems. Given that one is simply aggregating lower-level estimates to provide estimates at higher levels of aggregation, a natural concern is that the aggregate estimates will not approximately equal the direct estimates at higher levels of aggregation. (Equality would not be sensible given the variance associated with all sample-based estimates.) Bayesian hierarchical linear models can be developed, in principle, that would have the property that estimates at various lower levels of geographic aggregation would sum to the corresponding estimates at higher levels of aggregation; in addition, these sums would closely approximate the direct estimates for these higher levels of aggregation. To do this it would be appropriate to develop a model at the finest level of geographic/demographic aggregation and let the Bayesian prior to posterior mapping bring in data organized at various levels of aggregation. This can be challenging, since the number of parameters can become large, but approaches to a solution exist. Unfortunately, the property that estimates sum over levels of aggregation and also that higher-level estimates closely approximate direct estimates at that level of aggregation, which is currently possible for Bayes' hierarchical linear models, may be difficult to achieve for generalized linear versions of these models.

In Rod Little's discussion of Tom Louis' presentation, he agreed that Bayes' hierarchical modeling, also known as full probability modeling, was attractive for complex ACS estimation tasks. Having available the full posterior distribution for estimates was important to handle loss functions other than that of mean square error. Of course, most analyses will continue to focus on common summaries, e.g., means and standard deviations, but there are alternatives that should be considered. One important application in which the posterior distribution would play a role is for multiple imputation of missing data, which the ACS will need to accommodate.

When engaged in survey inference, the goal is to create predictive distributions for nonsampled and missing data values in a population. A well-constructed Bayes' hierarchical model can yield these predictive distributions in a manner that gives them many positive features. These models can incorporate information from disparate data sources, they can be used to treat missing data, and they can be used to appropriately reflect variables used in the sample design. Furthermore, these models allow borrowing strength across geography (and time) for good small-area estimates, and they support the display of the use of prior information, which can inform users as to the specific impact of the prior on the posterior distribution. These models have the further advantage of flexibility in that sensitivity to specification of the prior distribution can be readily assessed.

While Bayes' hierarchical modeling can be used to incorporate information from administrative records, the use of such records requires that they be comparable across regions. This can be checked by keeping track of differences in programmatic rules and methods, but it can also be checked by comparing administrative record tabulations with survey data pooled over time, a technique similar to that used in the work by the panel on small-area estimates of poverty. One method for reducing regional biases if they exist is to form strata that cut across regional boundaries.

The criticism that these models are too complex is easily refuted. Markov chain Monte Carlo techniques make the computational complexity of Bayesian models an increasingly minor issue. Similarly, the criticism that the techniques are too dependent on assumptions is refuted since the sensitivity can be assessed. Also, many simple, frequentist techniques rely (at times implicitly) on strong assumptions that may not be supported by the data. What really matters is not the complexity of the algorithm used to generate the estimates, but the complexity of the model itself and the key assumptions on which it relies. It is necessary to identify the mean structure and the variance structure and the hierarchy between the model components. It is then necessary to examine the sensitivity of the model to misspecification.

FINAL POINTS

Bayesian models are likely to provide a natural framework for combining information from the ACS, the census, household surveys, and administrative records. The various advantages of this framework, such as the incorporation of nonresponse, were stated. The identification of particular models for this purpose could not be done, since the ACS has not yet been fully implemented and therefore much about the data structure remains unknown.

3

Combination of Information Across Time

While very large, the ACS sample size collected from the smallest jurisdictions in a year will not be sufficient to support direct annual estimates of acceptable precision. Current plans are to not publish direct estimates for areas with populations of less than 65,000 people. For these areas, instead of the individual yearly estimates, the Census Bureau proposes to report equally weighted moving averages of the ACS yearly estimates for the most recent 2 to 5 years, depending on the size of the area. (The use of the cross-sectional models suggested above and the possible oversampling of governmental units with less than 2,500 population could reduce, but likely not eliminate, the need for some kind of borrowing of information for the smallest areas.)

Assuming linear changes in the true response of interest over time, moving averages will be estimates of the situation in an area 6 months to 2 years in the past, which would still be preferable to use of the decennial census information, which is generally less current and can be as much as 10 or more years out of date. Rather than use a moving average, particularly the equally weighted moving averages under consideration, other time-series approaches are possible, especially when one considers that ACS information (unofficially) can be tabulated on a monthly basis, therefore providing 60 observations over a 5-year period. Alternate forms of time-series modeling (e.g., ARIMA[1]) could reduce the variance of the resulting ACS estimates compared

[1]ARIMA, autoregressive integrated moving average, is a broad class of time-series models.

with the use of moving averages. An additional advantage of such alternatives, in comparison with equally weighted moving averages, is that the resulting estimates could be used as predictions for the current year and therefore would have less time bias. Use of these methods would also reduce some equity concerns when estimates from the ACS are used as inputs into fund allocation formulas, since it would be helpful to provide estimates with as little difference in variance as possible for areas (e.g., counties) regardless of their size.

Borrowing ACS information across time also raises a broader combination-of-information challenge than represented in the discussion of the cross-sectional models discussed above, since the ACS, most household surveys, and most administrative records data are collected annually (and sometimes monthly). These inputs for other time periods could be used in several ways to improve the above cross-sectional models.

RESEARCH DIRECTIONS

Bill Bell organized his presentation into three pieces: (1) general borrowing of information over time in repeated surveys, using time-series models, (2) the specific use of moving averages or ad hoc smoothing models for the purpose of borrowing information over time in repeated surveys, and (3) methods used in the project on small-area estimates of poverty to combine information across time and geography and the relevance of these methods to ACS.

Borrowing Information Over Time in Repeated Surveys

The original work on borrowing information over time in repeated surveys was by Scott and Smith (1974). They assume that $y_t = \theta_t + e_t$, where y_t is the time series that is observed, θ_t is the true process of interest, and e_t is the sampling error. In this situation, data and models are needed for both the time series of the sampling errors e_t—whose distribution is primarily determined by the sampling autocovariances—and that of the true response θ_t, which is assumed to have a stochastic (error) term (e.g., a common assumption is that $\theta_t = \lambda\theta_{t-1} + \varepsilon_t$) and may depend on regression variables. In this context, the stochastic term for the true process is generally assumed to be correlated over time and nonstationary. Best linear unbiased prediction (based on multivariate normal conditional expectations and variances) is used to estimate the model's parameters. One difference between this and cross-sectional models is that it is more difficult to recognize uncertainty in model parameters. Variance estimation of the estimates is complicated, but there are simulation approaches to this problem. Bell was aware of two implementations of this method: by Dick Tiller at the Bureau of Labor Statistics on the state labor force time series and by the Australian Bureau of Statistics to estimate regional time-series estimates. This approach has been researched for some years, but it has been difficult for researchers to demonstrate substantial gains for the following reasons: when sampling error is low, substantial gains are not possible; when sampling error is high, although there is the potential for substantial gains, estimating the model's parameters is more difficult.

A key problem in this area concerns the need for a model of the sampling errors, especially the autocovariances. Current plans do not exclude the possibility that in the ACS design the sampling errors will be approximately uncorrelated. If that turns out to be the case, it would permit a great simplification. Another issue is the consequences of uncertainty about the variances and time-series parameters, in particular the signal-to-noise ratio (i.e., model error variance relative to sampling error variance). Furthermore, there has been little study of the robustness of the resulting estimates to model misspecification.

A related topic is that of benchmarking. This is the adjustment of estimates from, for example, a monthly survey so that the estimates for a given year, when summed, agree with annual data produced from another survey or a census. This adjustment would be supported by the assumption that the annual survey had less sampling error (which is reasonable) and possibly less nonsampling error. If so, benchmarking can reduce both sampling and nonsampling error. In the small-area poverty work (National Research Council, 1998, 1999), the Census Bureau made the assumption that the CPS contained *less* nonsampling error than the census, which is the reverse of the more usual situation. As a result, the Census Bureau did not constrain the CPS results to agree with the census. The interesting technical question was therefore how to use the census data to reduce variances without substantially increasing nonsampling error. This was accomplished by using the census data to define regression predictors in the CPS equation and using the fitted CPS equation to carry out empirical Bayes' smoothing, which effectively calibrates the estimates to a CPS basis. This issue will need to be addressed with the ACS with respect to surveys that are considered to be highly reliable for various outputs.

USE OF MOVING AVERAGES FOR BORROWING INFORMATION OVER TIME IN REPEATED SURVEYS

One suggested method for borrowing information that is under consideration for the ACS is the use of moving averages, e.g., estimating the true series at time t by averaging values for the observed series for the k closest time periods (where k is some small integer). Moving averages are a simple way to achieve reduction in the size of sampling errors associated with an estimate, assuming that the sampling errors are relatively uncorrelated, with the downside that one is not then directly estimating the quantity of interest. So the use of moving averages results in a particular bias-variance tradeoff. Also, moving averages have a time delay for contemporaneous estimates, though asymmetric moving averages can provide estimates with less time delay. An interesting complication is that the application of moving averages to survey data later to be used as inputs in a model (e.g., for use in regression models as a dependent variable) may be problematic, as the moving averages will alter the statistical properties (particularly the autocorrelations) of the data.

With model-based smoothing over time, one is typically producing minimum mean square error estimates of a response at the current time, assuming that the model is correct. Moving averages, as ad hoc procedures but still model-based, tend to disguise the underlying model that one is assuming. In various applications, the underlying model may or may not be sensible.

For the specific situation where the true series follows a random walk (i.e., $\theta_t = \theta_{t-1} + e_t$) and the sampling errors are uncorrelated over time, the resulting optimal time-series smoothing weights depend only on the signal-to-noise ratio. If one has 5 years of data, with a small signal-to-noise ratio, the optimal method approaches equal weighting. As the signal-to-noise ratio increases, one gets close to using only the direct estimate. Similar results could be obtained from other models, and further study is needed for other situations.

COMBINING INFORMATION CROSS-SECTIONALLY AND ACROSS TIME

The third part of Bell's presentation concerned how the Census Bureau might put together ACS, household survey, decennial census, and administrative records data over several years to produce small-area estimates, based on the work of the small-area poverty estimates program at the Census Bureau (described in the Appendix). The estimation strategy for small-area poverty estimates starts with a base model, representing the use of direct estimates for small areas, which is simply true process plus sampling error, as in the model by Scott and Smith (1974). It also uses a regression model for the true process, with an additive model error. The regression variables come from administrative records data. In the county model, the census data are also brought in as an additional covariate. In the state model, the Census Bureau incorporates information from the decennial census by adding the residuals from the analogous regression fit using census data as the dependent variable. The Census Bureau is examining a more recent approach for the county-level model, referred to as the bivariate model, which uses two linked equations, one for the census estimate and one for the CPS estimate. They are both true process plus sampling error models in which true process is modeled using multiple regression and the error terms for these models are assumed to be correlated, which links the two models. There are formulations of this approach in which the census residuals show up in the CPS regression, but with a coefficient that varies according to the sampling variance in the census. Since the sampling variances in the state model are small, this approach makes little change to the state model, but it has an effect in the county model. A related approach would be to include the use of a measurement error model to link aggregate CPS and census responses, which can be thought of as a restricted form of the bivariate model.

The small-area poverty estimates have not yet incorporated multiple years of CPS data into the model, though the Census Bureau has experimented with state-level models that use up to 5 years of CPS estimates using a multivariate generalization of the bivariate model. The problem, similar to that discussed above, was in developing a time-series model for the model errors. Further, given the high-level of sampling error for counties and most states, it was unlikely that including previous years of CPS data would be that helpful.

When considering the ACS as an added source of information, most likely as a replacement for the census in the above models, especially the bivariate model or multivariate generalizations, its large sample size could result in at least two modifications to small-area poverty modeling. First, the census may be less useful as a covariate when the ACS is included. Also, for the same reason, using past years of ACS data may have greater value than using past years of CPS data.

Clearly, this area is only beginning to be explored. The most troublesome possibility is when one cannot conclude that the regression model is stable over time, because in that case one has many fewer data points with which to work. Another difficulty occurs when modeling discrete outcomes, since there the theory is even less well developed.

Discussant Eric Slud stated that there are two distinct types of information involved in this area: statistical variation and the variation of the signal over time. Borrowing strength over time involves understanding the models that generate both, otherwise the bias could be substantial. Understanding the autocovariance of the sampling errors is extremely important. More generally, this borrowing of information over time is a highly model-dependent activity. Therefore, it is important to emphasize model-checking assessments and whether the validity of assumptions can even be assessed.

Another key question is how to model nonsampling error, which does not appear to be treated in the literature, except in applications with longer time series. A key nonsampling error here is that the CPS, the census, and the ACS present different approaches to measuring various quantities—for example, in the primary example cited, for measuring poverty. In order to make full use of these measurements in combination, one would benefit greatly from the use of a measurement error model. To support development of such a model, more matching studies are needed between the CPS and the census, and, when possible, three-way match studies involving all three data sources. These studies might also help in measuring the degree of cross-correlation among the measurements, which would be useful in a Kalman filter[2] approach to this problem, and they might help support simple weighted combinations of direct estimates and model-based estimates.

Generally speaking, every small area will follow its own time-series model. The only hope is that one can model these collectively in a simple form, in which case it will be possible to share information across small areas through a regression model that includes a shared random-effects component. This is a multiple time-series problem, and it has not been well researched.

The development of time-series methods that use auto- and cross-correlations that do not exist or are poorly estimated is an important point. It supports the need for some kind of study of whether there are substantial divergences from these estimated or assumed auto- and cross-correlations and, if so, what the implications are. A first step would be checking the sensitivity of the results to the use of alternate auto- or cross-covariance forms. This could be a component of a larger study of model form that is related to nonsampling error. This step will be needed in the effort to calibrate the long form with the ACS (discussed in Chapter 7) and in calibrating the ACS with various household surveys. Such studies will also help in the interpretation of standard errors that would be produced from time-series analysis.

The floor discussion raised additional modeling ideas. One possibility would be to use direct estimates at a higher level of aggregation and moving averages of shares over time to allocate these estimates to smaller areas. The question was raised whether the ACS would release monthly estimates, which would facilitate time-series modeling. Other issues raised included the benefits from incorporation of spatial autocorrelation structure in these models and ways of addressing the effects of census undercoverage in these modeling problems.

FINAL POINTS

The various requirements for the development of time-series models for combining information in this context, including estimating the sampling autocorrelation structure, were put forward. The difficulties of parameter

[2]Kalman filters involve a "state space" representation of a time series, which assumes that a linear model (with an additive error term) represents the relationship between the observed series and a set of state space variables (representing aspects of the time series such as trend or seasonality), and a second linear model represents the relationship between the state space variables at time t and at time t - 1.

fitting and model validation were emphasized. As a result, the borrowing of information across time could be difficult, and even if accomplished might not yield substantial gains. In particular, it may not be easy to find an estimation procedure that provides a substantial improvement over the current decision by the Census Bureau to use moving averages. However, the potential remains for improvement, and methods were discussed that might be used to move forward.

4

Funding Formulas

Many funding formulas currently use as input census long-form data. However, this information is often several years out of date when utilized, which is a problem since the funding formulas are used for programs that are intended to address current problems. The ACS data therefore will have an advantage over long-form data for this application in that they will be produced on a more timely basis. As mentioned above, the ACS initially plans to use moving averages to provide estimates for very small areas. As input into funding formulas for these small areas, one could use the ACS equally weighted moving average estimate, an asymmetric moving average that gives more weight to the current time period (such weighted averages may not greatly increase the variance and will provide information that has less time bias), or one could use the (direct) estimate based only on the current year's information.[1] This is a tradeoff between variance and bias and their effect on the funds that areas would and should receive. How should this choice be evaluated?

It is easy to see that the use of moving averages could have an important effect on allocations. Consider a fund allocation formula that has an eligibility threshold, for example, an area receives benefits only when estimated per capita income falls below $17,000 per year. Also consider an area that typically

[1]Use of a more elaborate time-series model approach, as discussed in Chapter 2, could reduce the problem discussed here. However, that would be relatively complicated to apply for all ACS responses. Furthermore, the general issues discussed in this chapter would still be relevant.

has an (estimated) per capita income above the threshold (and hence is often not eligible) but has wide variation in this estimate (due to either real change or the variance in the estimate) from year to year. The use of moving averages in producing this area's estimated per capita income would reduce the number of highs and lows, which will, in turn, reduce the years in which estimated per capita income will fall below the threshold. Alternately, smoothing could also result in more money going to an area. For example, smoothed estimates for an area with an average value below the threshold having substantial variability from year to year will serve to keep that area eligible for this hypothesized program for more years.

Fund allocation formulas often have hold-harmless provisions, which is another common feature that can cause an area to have substantially different allocations as a result of that area's estimate having a larger or smaller variance. Hold-harmless provisions guarantee an area a high percentage (often 80% or 90%) of the funds they received in the previous time period regardless of their inputs for the current time period. These provisions have the goal of protecting areas against large decreases in funding from one year to the next, so they can undertake multiyear fiscal obligations. Assume now that an estimate for an area, to be input into a formula with a hold-harmless provision, has a large variance. Assume also that larger estimates are associated with larger allotments. Because the estimate for the area has a large variance, at some point the estimate will be relatively high due to random variation, and the area will receive a much larger allocation than it would have without the random variation. Then, as a result of the hold-harmless provision, that area could continue to receive "undeserved" higher benefits for several years.

The overall question to address is how to evaluate the performance of direct versus moving average estimates as inputs into fund allocation formulas. A first step would be to understand how variance affects allocations.

RESEARCH DIRECTIONS

Charles Alexander described the Census Bureau's current plans. For areas of more than 65,000 population, the input into fund allocation formulas from the ACS will be the direct estimate. For areas with population between 30,000 and 65,000, the input will be the average of the most recent and previous years' data. There are further thresholds for the use of moving averages of 3 and 4 years. For the smallest areas of less than 15,000 population, the input will be the average of the direct estimates for the most recent 5 years. These cutoffs are designed for typical uses of census data (point-in-time observations) and are based on a roughly equal coefficient of variation criterion for sampling error. However, it is important to distinguish between uses of estimates that require an assessment of the situation at a given time and uses of the estimates for forecasting, since forecasting would favor shorter moving averages.

Alan Zaslavsky's presentation, based on joint work with Allen Schirm, focused on the relationship among data, estimation methods, and funding formulas. His goal was to show that funding formulas with such features as hold-harmless provisions, eligibility thresholds, using estimates (especially nonlinear estimates) with substantial bias and variance, along with dynamics in the quantity of interest, often have unintended consequences.

Funding formulas often need three inputs: the number of people that are categorically eligible, the rate of incidence, and the total population of interest. These quantities are estimated in a variety of ways, sometimes directly and sometimes indirectly. The indirect estimates can include averaging over time, but more involved methods include small-area estimation models, using regression and empirical or hierarchical Bayes' methods to combine data over time and space. Much of the following is exemplified by two main examples: the allocations to states under the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) and Title I allocations to counties and school districts (discussed in the Appendix).

Typical data sources as inputs into fund allocation formulas are the most recent decennial census—generally the long form, which has nonnegligible sampling error for small areas; household surveys, which have smaller sample sizes but good measurement properties; and administrative records, for which the content is often not what one desires, the definitions are often inconsistent, and access and use can be complicated. The ACS as a data source has many advantages (and some disadvantages) over these existing data sources. It is more current than, say, the last census (use of the census for input to an allocation formula implies a great deal of stability in what is being measured). The ACS has a larger sample size than all current household surveys, and its content is better tailored to meet current data needs than administrative records. Of course, it is expected that the output from household surveys will have considerably less measurement error than the ACS.

The use of data from the ACS will likely have the following implications: more frequent recalculation of formula inputs due to the timely availability of current data, the availability of useful direct estimates for more variables and smaller areas, and improvement of the quality of indirect estimates due to currency and uniformity.[2] The direct estimates from the ACS will have reduced variance relative to current surveys, but an increase in variance relative to the census long form. Most important, there will be a reduction in bias due to the increased timeliness of the information. Also, changing from a census-

[2]It is possible that the production of ACS estimates could generate interest in reviewing the allocation formulas themselves.

based or CPS-based measurement process to an ACS-based process will very likely introduce considerable measurement error.

It is important to understand that the estimation procedures and the formula are not separable. For example, if one computes this year's inputs through averaging the data for the previous 3 years, that is equivalent to simply requiring as the input the average of the previous 3 years. In one case, it is a choice for an estimate, and in the other it is what the formula requests.[3] (Of course, the best estimate of the moving average of the previous 3 years' true means, which a fund allocation program might ask for as an input, might not be equal to the simple 3-year moving average of the individual annual estimates.)

The presentation focused on the impact of the variance of estimates used as input for fund allocation formulas, with the understanding that bias also has an extremely important and possibly more widely appreciated impact. If an estimation procedure is linear and if the allocation formula is linear, variance in the inputs will not affect the long-range allocations in expectation and therefore are unlikely to matter very much. However, in the nonlinear case, the variance of an estimate for an area can have a large effect on how much money an area receives. This poses a problem, since an area has no effect on how much variance its estimates have.

A simulation study was described in which four factors were varied: (1) the sampling standard error; (2) the estimation method (single year, 3-year moving average with equal weights ending with the current year, or 3-year exponentially weighted moving average also ending with the current year); (3) the formula (use of hold-harmless or threshold provision, or both); and (4) population trends (constant, trending upward, trending downward). The assumption was also made that the general allocation formula, possibly aside from these two provisions, is proportional. The simulations were over a 4-year period and were repeated (except for a few exceptions) 10,000 times.

The simulations showed that with an eligibility threshold and no trend, sampling variability serves to smooth out the effects of the threshold, which may not be a bad thing from a public policy point of view. However, the degree of smoothing depends on the sampling variance, which is related to the size of the area for many household surveys, with the result that an area's allocation depends on the sampling variance of its inputs, which is not rea-

[3]There is a distinction between providing different estimates that measure conceptually distinct items and providing different estimates for the same conceptual item as a result of using different loss functions. The former is easily explained, while it is difficult to explain the latter. It would be useful to better communicate the costs of using the wrong loss function to users. More generally, point estimates are not the best way to communicate information about the estimates of a quantity. It might be more useful to have some idea of the distribution of estimates for that quantity, which address more questions relevant to fund allocation.

sonable. Examining an 80 percent hold-harmless provision with no trend, the larger the sampling variance, the larger the allocation bias,[4] given the asymmetrical nature of the effect of the hold-harmless provision. If one uses moving averages, the allocation bias is reduced. Specifically in this study, a bias of 20 percent is reduced to one of 3 percent using a 3-year moving average. If rapid responsiveness to fluctuations is not that important, this smoothing helps to reduce the allocation bias that occurs from a having a hold-harmless provision in the allocation formula. Combining a hold-harmless provision with a threshold is even more worrisome, since the threshold results in larger jumps in allocation from year to year. With a downward trend, 3-year accumulations cause a bias, since one is using data for years that do not reflect the latest changes. With a hold-harmless provision, coupled with a downward trend, the effect of the provision is more extreme, since the smoothing of the trend and the hold-harmless provision are both (occasionally) keeping the allocations too high. With an upward trend, the effects from the provision are less worrisome.

The remaining results from the simulation were relatively intuitive, once the patterns were examined, with three main points: (1) the data sources, the estimation procedures, and the allocation formulas must be considered as an integrated whole; (2) when a change is made to one piece, the others must be kept in mind, going back to consider the original intentions of the program, if necessary; and (3) linear procedures have more predictable consequences, and methods that produce smoother estimates also have more predictable consequences. New data sources such as the ACS require reevaluation of funding formulas in light of the original intentions of the program, not simply replicating previously used procedures.

The floor discussion raised a number of additional issues. The importance of the level of geography at which estimates were needed in these programs was stressed. In response to the comment that use of the ACS might result in an increase in the variance of estimates in comparison with use of the census long form, the value of using the ACS in combination with, say, the CPS, to reduce sampling variance (and also measurement bias) was stressed. This approach, which would provide high-quality, timely small-area estimates, might provide an incentive to consider distributing funds using federally controlled allocations to smaller areas. Also, it was noted that once the hold-harmless allocations are computed, the total amount allocated would be affected, necessitating a recalculation of shares.

[4]With respect to these simulations, allocation bias is defined as the average difference, over replications, between the allocation received using estimates with some assumed variance and the allocation received using estimates with zero variance.

The problem of thresholds and hold-harmless clauses in fund allocation formulas was also addressed 25 years ago by the Federal Committee on Statistical Methodology (Office of Federal Statistical Policy and Standards, 1978) in its first report, although the number of fund allocation programs was much smaller then. Considering the expansion in programs with allocation formulas, it is a good time for reconsideration of these issues. The formulas could be reviewed, along with the estimation methods, the data sources, and what statistical problems might be involved. The three different elements in the allocation process—the formula, the data source, and the estimation procedure—are in the domains of many different individuals and groups, and they rarely interact.

TerriAnn Lowenthal commented on fund allocation formulas and the ACS from a congressional perspective. The concern is whether the designers of these formulas realize the implications from their use, especially the unexpected consequences of hold-harmless provisions and eligibility thresholds. She described how formulas come about in the legislative process. She suggested that this discussion needs to be communicated to the people that are developing these formulas. There needs to be more interchange so that the developers understand and avoid these unintended consequences. She added that the project on small-area estimates of poverty is an illustration of how to bring together the interested parties on the intent of legislation. One needs to keep in mind that it is important to find the right language to communicate with members of Congress.

FINAL POINTS

The effects of the variance of estimates used as inputs to fund allocation formulas with features, such as hold harmless provisions and eligibility thresholds, is complicated and can have unintended consequences. This needs to be more widely understood. The impact of the bias of estimates used as inputs to fund allocation formulas also needs to be examined, especially how one might trade off of bias and variance in comparing competing estimators for purposes of equitable fund allocation.

5

Weighting and Imputation

Nonresponse and undercoverage in the ACS will require various weighting and imputation schemes to produce annual and monthly estimates that accommodate both of these sources of incomplete data. The Census Bureau's current plans for the ACS involve the use of as many as 11 factors in an overall weighting scheme (see Alexander et al., 1998, for a more detailed description). The factors are designed to account for:

- oversampling of small governmental units,
- computer-assisted personal interviewing (CAPI) subsampling,
- monthly variations of the percentage of the population using different response modes,
- noninterviews (two separate factors called "noninterview factors"),
- mode bias,
- differences in nonresponse in individuals and households, so that housing unit counts agree with the totals on the master address file (two separate factors),
- differences between marginal population totals and totals based on demographic analysis resulting from census undercoverage (referred to as a person post-stratification factor), and
- household-level undercoverage.

As is obvious from these descriptions, some weights need to be applied to individual records and some are applied to household records.

While the justification for many of these factors is relatively straightforward, for others, such as noninterview adjustment or mode bias factor, there are clearly a variety of ways to define the weights and how they are applied. How should alternatives be judged? What are the evaluation criteria? Also, the adequacy of some of the controls, such as population controls (person poststratification factors), need to be considered; perhaps methods could be developed in which ACS data would be used to improve various control factors.

Charles Alexander suggested that the weighting scheme currently used in the ACS pilot testing might include some unnecessary factors, since it was designed before data were available in order to have something in place for the July 1997 deadline for 1996 data. This timing resulted in a number of weighting factors that in practice are very close to one, so their utility, at least for the initial application, has been minimal. The overall approach that the Census Bureau adopted is somewhat old-fashioned, mimicking that used in the decennial census, but it is known to work. At least one of the weighting factors results from the Census Bureau's decision to have a given month's estimates make use of data collected during that month, rather than data originating from the sample selected for that month. This was done to provide data that would likely have less mean square error, based on the reasoning that it was easier for the respondent to provide reliable information for the month of data collection (reducing recall bias), but it has complicated the weighting factors needed to combat nonresponse.

The Census Bureau is not wedded to the current methodology, and there is still an opportunity to make modifications to the weightings that are used with the 1999 ACS. In addition, after data have been collected for a few years of full implementation, it is hoped that the weighting scheme can be changed to reflect specific aspects of the data. One area that is likely to change is the controlling of ACS population counts to county-level population estimates. The Census Bureau is uncomfortable assuming that the county-level population estimates are so reliable that they cannot be improved through combination with the ACS estimates, especially at the county level disaggregated by age, race, and sex. (Differences in residence rules also may make these controls questionable for this application.) For the future, the Census Bureau is even considering, through use of ACS data, development of population estimates at the tract level.

RESEARCH DIRECTIONS

Robert Bell's presentation on weighting intentionally raised questions more than it provided answers. There are at least three purposes of weighting: (1) for nonresponse adjustment and related issues, (2) for poststratification to accepted external values (such as dealing with undercoverage), and (3) to treat differential sampling rates. This discussion focuses mainly on weighting for treatment of nonresponse (and mode bias) or for differential sampling rates, and the few general points made relate to the use of weighting for those purposes. However, some minor points are also made concerning a specific use of poststratification in the ACS.

The typical purpose of weighting to adjust for nonresponse is to reduce bias without greatly increasing variance. Biases can occur when three conditions hold: (1) sampling probabilities vary (designed or otherwise, as through nonresponse) with some background variable, x, such as geography, tenure, or demographic characteristic; (2) an outcome of interest, y, also varies with x; and (3) the sampling probability is correlated with y. In this way, housing units with a high (or low) probability of being sampled are likely to have a response that is higher (or lower) on average, which will cause a bias. If the nonresponse is properly modeled as a missing-at-random process (i.e., the value of the outcome, y, is not involved in the probability of response conditional on x), weighting could theoretically eliminate the bias. Although this assumption is generally not true, weighting can generally serve to reduce bias.

At the same time, weighting tends to increase variance, which raises weighting choices involving tradeoffs of bias and variance.[1] The ideal solution would be to compute an estimate that is a weighted mean of predicted values for strata—the predicted values making use of a model for nonresponse within the strata—weighted by the population size of the strata. Then the variance could be controlled by "shrinking" the predicted values through the type of models discussed in Chapter 2. For example, an analysis of variance model could relate stratum estimates based on the demographic or other characteristics used to define the strata.

Practically, however, many users either cannot or would not perform this type of modeling. The Census Bureau therefore needs to use a method that takes into consideration how the data will actually be analyzed, which is to make use of an output file that is the result of the application of weights.

There are a lot of decisions involving weighting, such as whether to try to address a specific set of problems or to try to apply a more general weighting framework, what level of geography to use in forming weighting cells, what factors to take into account, and whether to cross-classify these factors or to do some sort of raking.[2] Most of the decisions (both explicit and implicit) made by the Census Bureau in its weighting scheme for the ACS seem very reasonable. However, there are a lot of things that could have been done in other ways. It is worth mentioning that bias correction is elusive, in the sense

---

[1]While this statement is true when there is no use of additional, external information, there are applications of weights when variances are reduced through the use of external information. See, e.g., Rosenbaum (1987), which provides a technique that might be applicable to the ACS when using external information.

[2]Raking refers to a constant multiplicative adjustment within a row or a column so that the rows or columns of the revised table add up to a given marginal row or column total.

that correcting for bias at a given level of aggregation (e.g., geographic) is no assurance that a bias does not exist at a lower level of aggregation. So, for example, correcting a bias at the county level does not necessarily correct for bias at the tract level.

Because weighting involves a tradeoff between bias and variance, how should this tradeoff be considered? One criterion that makes sense is to minimize mean square error (though there are alternatives). This criterion still leaves two important issues: (1) For which outcomes is one minimizing mean square error and (2) at what level of geography and for what period of time should the focus of the weighting be? With respect to outcomes, it is important to decide which outcomes have the highest priority and therefore require weighting treatment. Should the focus be on demographic characteristics or attributes captured on the census long form? Reflecting the second point above, what is the relative importance of accuracy at different levels of geography or for shorter versus longer time periods? If one tries to minimize mean square error for each month of the survey, that is going to put too much emphasis on reducing variance, and as a result there is going to be bias that will tend to replicate month after month, which will dominate the variance for longer moving averages, such as annual estimates. Setting up this criterion is not easy, but it is very useful to do, since it puts emphasis on the outcomes, and it makes explicit which interactions one believes are important to consider.

An ACS feature worthy of attention is that the monthly estimates are based on responses collected in that month, which is quite different from the questionnaires collected from the panel selected for that month. Furthermore, respondents are asked to provide answers corresponding to the month of data collection and not corresponding to the month of panel selection. One major advantage of this is that the data are available sooner, since, for example, data from households in the March sample that respond by mail in March are immediately available for combination with data from the January and February sample households that also responded in March by telephone and in person, respectively. In contrast, for sample-based cohorts, one would have to wait until May for the complete data collection for March. The Census Bureau supports this with a second advantage, that of reducing recall bias. While there are advantages, it does necessitate additional weighting, referred to as "variable monthly sampling weights,"[3] which are used to address the biases that might be caused by the decision to base estimates on the data collected during a month.

Consider a case in which the March mail response was 40 percent of the sample, the CATI response was 30 percent of the February panel, and the

[3]The phrase "variable monthly sampling rates" will refer in the following specifically to weights that are applied to address this potential source of bias.

TABLE 5-1 Illustrative ACS Data Collection Percentages by Mode

	Mode and Month of Data Collection			
Month of Mailout	Mail	CATI	CAPI	Total
January	Jan 55%	Feb 25%	**Mar** **20%**	100%
February	Feb 45%	**Mar** **30%**	April 25%	100%
March	**Mar** **40%**	April 30%	May 30%	100%

NOTE: Boldface type indicates the data collected in March.

CAPI follow-up was 20 percent of the January panel (see Table 5-1). Then the 20 percent and 30 percent need to be jointly reweighted to represent the 60 percent missing from the March sample, so the individuals in those cells should get a weight of 1.2. This is reasonable if there are no major month-to-month differences. However, suppose the reason there was only a 40 percent response in March was because there was a flood in March, and the mail responses were from people who responded early in the month, who had a high propensity to respond. In this case it is not clear that the January CAPI and the February telephone responses are necessarily representative of the same groups for March. The question is whether variable monthly sampling weights can fix systematic variations in response rates.[4]

A second set of weighting factors that might benefit from an alternative approach is the noninterview weights. These are two-stage weights in which the first stage involves weighting all respondents to account for nonresponse and the second stage accounts for the different modes of response. (These stages are clarified in the example that follows.) For purposes of this set of weighting factors, there are three categories of people in terms of response: (1) mail or CATI respondents, (2) CAPI respondents, and (3) CAPI nonrespondents (see Table 5-2). One approach to this problem would be to assume that the CAPI respondents provide the best information about nonrespondents, which would argue for weighting the CAPI respondents to take CAPI nonresponse into account within each tract (see Table 5-3). The Census

[4]A simulation study is currently under way by the Census Bureau to examine whether using data collected in a given month (rather than from the sample selected for each month) is causing too much of a bias to support continuation of that procedure.

TABLE 5-2 Illustrative Responses for a Pair of Census Tracts: Basic Data

Mode	Tract I	Tract II	Total
Mail/CATI	60	80	140
CAPI	40	20	60
None	10	20	30
Total	110	120	230

TABLE 5-3 Illustrative Weights for Nonresponse Addressed by Reweighting CAPI

	Weights		Weighted Counts		
Mode	(1)	(2)	(1)	(2)	Totals
Mail/CATI	1.00	1.00	60	80	140
CAPI	1.25	2.00	50	40	90
Total			110	120	230

Bureau is concerned that this could generate relatively high weights in some tracts if the number of CAPI responses is small. Instead, the Census Bureau decided to make use of a two-stage procedure. In the first stage, all respondents are weighted up to account for the CAPI nonresponse (see Table 5-4). In doing so, one essentially imputes complete records through weighting using the records corresponding to mail/CATI respondents more than the CAPI respondents, so it is necessary to reweight to remedy this.

In the second stage, the mode bias factor downweights mail/CATI respondents by the ratio of the original weights across tracts to the weights taking nonresponse into account (see Table 5-5). The difficulty with this procedure is that the weights for each tract no longer match the number of (weighted) housing units in the sample; this is treated using another weight in a later stage of the weighting scheme, as demonstrated below (see Table 5-6).

In the example provided, the two-stage weights for this factor have the advantage of considerably less variability (a range of 0.91 to 1.66) compared with that for the procedure that reweighted only the CAPI respondents (a

TABLE 5-4 Example of Noninterview Factor in a Pair of Census Tracts

	Weights		Weighted Counts		
Mode	(1)	(2)	(1)	(2)	Total
Mail/CATI	1.10	1.20	66	96	162
CAPI	1.10	1.20	44	24	68
Total			110	120	230

NOTE: To reweight Mail/CATI and CAPI respondents to equal total, apply ratio of total to number of respondents. Therefore, apply the weights of 1.10 = 110/100 and 1.20 = 120/100.

TABLE 5-5 Example of Mode Bias Factor for a Pair of Census Tracts

	Weights		Weighted Counts		
Mode	(1)	(2)	(1)	(2)	Total
Mail/CATI	0.95	1.04	57.0	83.0	140
CAPI	1.46	1.59	58.2	31.8	90
Total			115.2	114.8	230

NOTE: To make the CAPI percentage equal the observed rate, apply the four weights as follows: 0.95 = 1.10 (140/162), 1.04 = 1.20 (140/162), 1.46 = 1.10 (90/68), and 1.59 = 1.20 (90/68).

range of 1.00 to 2.00). In a sense, instead of a cell-based model, one is using a type of raking. The reduction in the variability of the weights is not a compelling argument to use this procedure, especially if there are strong effects that are particular to tract-mode combinations. However, in the absence of strong tract-mode combination effects, it is difficult to design alternatives to this weighting procedure that have obvious advantages and that also keep the variability of the weights from being too large. Therefore, depending on the data, this approach may have a considerable advantage over the procedure described above.

TABLE 5-6 Illustrative Example for a Pair of Census Tracts: NIF and MBF Multiplied by Housing Unit Poststratification Factor

	Weights		Weighted Counts		
Mode	(1)	(2)	(1)	(2)	Total
Mail/CATI	0.91	1.08	54.4	86.8	141.2
CAPI	1.39	1.66	55.6	33.2	88.8
Total			110.0	120.0	230.0

NOTE: To make the housing counts in each tract equal to the observed number, 0.91 = 0.95 (110/115.2), 1.08 = 1.04 (120/114.8), 1.39 = 1.46 (110/115.2), and 1.66 = 1.59 (120/114.8).

Another factor with some interesting alternatives is the person poststratification factor. The idea of this weighting factor is to try to correct for undercoverage by controlling the ACS population counts to independent population estimates based on age, sex, race, and Hispanic origin at the county level. The key question is whether the population estimates provide better information than the ACS data at the county level and for demographic subgroups. The estimates from the ACS probably provide valuable additional information on changes since the last census at low levels of aggregation, but for higher-level totals the direct ACS estimates may be inferior. In addition, even controlling to better demographic data may not improve the estimates of population characteristics. For example, if ACS has inferior estimates because there is differential undercoverage, weighting probably will help; however, if the estimates are inferior as a result of the different manner in which people respond to the ACS questionnaire in contrast to the census questionnaire, then the ACS estimates may be harmed through use of this weighting.

A related point is that the ACS provides an estimate of the average population over the year, while population controls represent a point-in-time estimate of population, which could make a difference in areas with seasonal populations. In reaction to this possibility, the Census Bureau has developed a different question on the 1999 ACS form about the seasonal population. (This is one of potentially several areas in which ACS could be controlled to estimates that are conceptually distinct.)

The following points require further consideration. First, given that there are both individual and household weights, any inconsistencies between person and household weights may cause problems for users. Second, any weighting method needs to take the size of counties into consideration since they vary tremendously in size, which has an effect on the size of weighting cells.

In addition, there are other issues to examine. First, as the ACS accumulates enough data over time, it might be possible, using some simple time-series techniques, to model mail response (and other kinds of considerations) and thereby identify unusual differential mail response patterns to determine when certain weighting methods might have advantages.[5]

Second, the variable monthly sampling weights should be checked to see if there is an interaction of response mode with various characteristics, e.g., a vacancy rate. There is some evidence of the following situation: if a housing unit is vacant, it will obviously not respond by mail or telephone; it will be interviewed 2 months later with CAPI, at which point it may be occupied. However, if it had been occupied initially, there might have been a response, so there is a potential bias. This is true for any characteristic that differs between movers and nonmovers. The variable-mode sampling weight factor was an attempt to address this problem; its success is not clear.

Third, there is the possibility of different responses to the ACS for race and ethnicity questions. The calibration that is discussed in Chapter 7 should provide information as to the degree of this problem.

Fourth, weighting rules do focus on reducing bias without explicitly considering possible increases in variance. If these estimates are going to be used as inputs to nonlinear allocation formulas, the work in Chapter 4 indicates that with a high variance even unbiased estimates can result in strongly biased allocations. Therefore, the appropriate emphasis may be less concern with bias and more with getting good estimates.

Fifth, the county-level controls for race and Hispanic ethnicity are not now very reliable. Implicit in the methodology that produces these estimates is the assumption that anything at lower levels of aggregation within a state changes in the same way as the entire state. Therefore, information since 1990 at substate levels is not used to produce the estimates. Either the ACS should be used to improve these estimates or administrative record information at that level of aggregation should be incorporated into the estimates.

FINAL POINTS

Some alternative approaches to weighting and imputation methods were examined in comparison to the current plans put forward by the Census Bureau. An examination of these alternatives on ACS pilot data will determine which techniques to apply when ACS is fully implemented.

[5]It was mentioned that the average ACS nonresponse rate was around 2 percent, probably because response is currently required by law and therefore it is unlikely to matter how one treats nonresponse if the mandatory status is retained.

6

Sample and Questionnaire Design

The development of the ACS raises a number of issues concerning sample and questionnaire design, both for the ACS and for current household surveys. The potential uses of the ACS that involve sample or questionnaire design include: (1) modifying the ACS questionnaire to enable its use as a screener for oversampling selected populations (and for asking them further questions) to support current household surveys; (2) providing information for oversampling areas in other household surveys, which in turn could be used to make the sample design for these surveys more efficient (e.g., oversampling areas associated with higher variances for the National Crime Victimization Survey, possibly through use of variables associated with areas having more criminal activity); (3) using the ACS to help determine when redesigns of household surveys are needed and to support those redesigns;[1] and (4) using responses to the ACS questionnaire to effectively increase the sample size of current household surveys through regression-type modeling, with the possible redesign of those surveys as a result of this modeling. With this last possibility, assume an ACS question that, when aggregated, has a given correlation (e.g., at the state level) with an aggregate output of a household survey of interest. Then, using a regression-type model to combine information, how much could the variances of estimates from the household survey be reduced? How would one change the

[1]By survey redesign, we mean at least a new sample size, a new sample allocation, and reselection of primary sampling units and households within primary sampling units.

design of the household survey as a result of such modeling? For example, if such a model reduced the variances differentially by area, the sample design could be modified to concentrate samples in those states for which the model was less effective.

These proposed uses of ACS responses bring up several policy issues, of which one is key: What will be the process for adding questions of interest to other federal agencies to the ACS questionnaire? Other important questions include: Who will decide the priorities and allocate the costs of these additional questions? What is the effect of added questions on the quality of response for the remainder of the ACS questionnaire? The policy for determining ACS content after 2002 will require the interaction of the Office of Management and Budget (OMB), Congress, and other interested and involved agencies.

Denise Lewis of the Census Bureau described the National Crime Victimization Survey (NCVS) within the context of the broad issues under discussion, especially with respect to sample and questionnaire design. The NCVS[2] is a household survey that collects data on the amount and types of crime in the United States and measures the incidence of personal crimes of violence and theft and other household crimes, such as burglary and motor vehicle theft. The Bureau of Justice Statistics uses the NCVS to publish annual estimates of the nation's crime rate for various demographic groups.

The potential benefits of the ACS acting in concert with the NCVS lie in four areas: (1) improving the effectiveness of the data collection by increasing the use of CATI; (2) improved weights and improved control totals for use at the state level; (3) use of output from the ACS (assuming a limited number of crime-specific questions were added) and the NCVS in statistical models for developing improved state estimates, possibly even making many more state estimates reliable enough for release (analogous to the modeling accomplished for small-area estimates of poverty); and (4) as a screening device, possibly including use of the ACS to screen for crime victims, to screen for nontelephone households, and to screen for rare events. Before proceeding in these directions, several important questions need to be addressed. First, what is the likelihood of adding questions to the ACS? Second, what priority does the production of state-specific estimates have? Third, what are the cost implications of using the ACS as a screening mechanism?

[2]The sample design for the NCVS is a stratified, multistage cluster sample that collects information on all persons 12 and over in about 60,000 housing units. Each sample consists of six rotations. Sample units in a given sample rotation are interviewed once every 6 months for 3 years. Each rotation is further divided into six panels, and each subsequent panel is interviewed in successive months, so one-sixth of a rotation is interviewed each month during a 6-month period.

Alexander focused on the policy issues that need to be addressed to make full use of the ACS as discussed. First, with regard to adding a question to the ACS questionnaire, the content currently consists only of questions that are required by law. For a separate section of questions that are voluntary, one hopes that the addition of such a section would not adversely affect the quality of the required responses. Second, if an agency is planning on using its survey as a follow-up to the ACS, a problem might arise if the ACS responses are considered confidential under the privacy provisions of Title XIII.[3] Third, the charge for additional questions is yet to be worked out. Many of the agencies interested in making use of the ACS do not have a great deal of discretionary funds for this purpose. Finally, there are a number of pragmatic, more focused questions, such as how quickly the ACS should pick up new construction. Some users may need this to be done more expeditiously than others. Answering these questions will require the interactions of several parties, including OMB, Congress, statistical agencies, and user groups.

RESEARCH DIRECTIONS

The main purpose of Lynn Weidman's presentation was to raise some operational questions and discuss these in terms of what the Census Bureau would like to do. The Census Bureau manages many household surveys, which are redesigned after each decennial census to take advantage of the latest available data. A hope is to use the ACS for help in redesigning these surveys more frequently. Prominent examples of these surveys are the Survey of Income and Program Participation (SIPP), which asks about income, employment, and participation in various governmental aid programs; the National Crime Victimization Survey, discussed above; the Current Population Survey, which deals mainly with employment but has many supplements dealing with a wide range of subjects (particularly each March's demographic supplement); the Consumer Expenditure Survey, which details how people spend their money; and the National Health Interview Survey, which requests information on a wide variety of health questions.

In the past, new construction would be added to the address list in each primary sampling unit (PSU) at the time of the decennial census. The hope now is to more continuously update the master address file. The ACS is also considered for use as either a screener or to oversample areas that have a large percentage of people with certain characteristics, and as a source of covariate

[3]Title XIII of the United States Code provides detailed regulations concerning the activities of the U.S. Census Bureau.

information for regression models to reduce the variance of household surveys. Weidman discussed each of these in turn.

Oversampling Groups

Household surveys often have different coefficient-of-variation stipulations for specified demographic groups. Two ways of achieving better performance for subgroups is oversampling of areas and screening persons and housing units to include more persons with those characteristics in the sample. For various household surveys there is potential interest in oversampling young children, the elderly, high- or low-income groups, and racial groups. Until now, one could use the decennial census to identify areas that could be oversampled to collect data targeted to these groups. As the decade progresses, those areas are less and less useful for this purpose. The ACS will be collecting information that could be used to retarget the sample.

This approach raises several problems. One problem is the size of geographic aggregation at which this process would operate. The larger the area, the greater the travel for household survey interviewers if they are using CAPI. It is not clear at what geographic levels ACS will be informative. Certainly, oversampling individual blocks will not be feasible with ACS. For higher levels of geographic aggregation, one might need to accumulate data over several years to inform targeting, but then the information becomes somewhat dated.

A Screening Tool

With respect to screening, the current situation is analogous to targeting, since one now uses the decennial census, even when it is less and less current. The demands raised by screening might require more sampling (or oversampling) since one has to find a match to the characteristic(s) of interest. However, screening is more efficient, since one gets the households or people one wants to interview. The greatest problem is that one is now asking people to answer two completely different questionnaires, the ACS questionnaire and the follow-on questionnaire, at the same time. The ACS questionnaire alone takes 30-60 minutes to complete. As a result, the Census Bureau has to worry about low response rates for the ACS and especially any follow-on survey. The issue of nonresponse, possibly involving the groups of interest, also complicates this possibility.

Redesigning Household Surveys

The third area of interest is the periodic redesign of household surveys. Based on ACS information, one might want to reselect PSUs or to resort housing units within a PSU (when using systematic sampling) to incorporate more up-to-date information. This could happen either more frequently than every 10 years or when ACS data indicated that it would be useful. However, there is a cost in making such changes. There is also an unknown amount of time that is required to implement a redesign. For example, if the counties (PSUs) are changed, then new interviewers would have to be trained in those areas. So a key question is whether more frequent redesign would be cost-efficient.

Supplying Covariates for (Variance-Reducing) Regression Models

Finally, there are complications concerning the addition of questions to the ACS questionnaire to support (combination of information) modeling. First, the ACS will be broad-based with respect to subject matter; that is, it will have relatively few questions for any specific area, e.g., income or health. For some models for some areas of interest, it may not be necessary to add many questions to the ACS questionnaire. However, it is likely that additional questions could be useful for models of many subject-matter areas. It is not possible to add hundreds of questions to the ACS, so some process for selecting additional questions will be needed. An important issue is whether these additional questions will be a permanent part of the survey. Other issues include where on the questionnaire these additional questions are placed. They could be incorporated in with other questions related to that subject area, or they could be placed at the end of the ACS questionnaire. The answer will affect nonresponse and quality of response of the other questions. Another key issue is the extent to which the information collected is increased with a question that is administered on a self-response basis (except for the CATI nonresponse cases).

Discussant Cathryn Dippo focused mainly on the practical issues raised if one were to use the ACS for redesigning current household surveys between decennial censuses. The most important feature for this discussion was the development of the master address file (MAF). Before turning to this topic, Dippo discussed some background issues.

The Bureau of Labor Statistics conducts some surveys where the units of interest are persons and some where the units of interest are households. Also, for the CPS, the real interest is in estimates of change over time, not estimates of level. In the second stage of the design of many household surveys (selecting housing units within PSUs), there is some implicit stratification based on the short form. The Consumer Expenditure Survey (CEX) is one example: before the 2000 census, it made use of information from the short form on household rent (the contracted rent amount for rental housing or the corresponding value of owner-occupied housing). Some surveys use area sampling at the second stage. If this is performed using sampling proportional to size, these size estimates have to be useful at low levels of geographic aggregation. Other related issues are the treatment of new construction and input into surveys that use random digit dialing.

A major issue concerns the operational aspects of changing PSUs between censuses and the costs associated with this. Every change to a PSU requires firing some interviewers and hiring and training the new ones. New interviewers, on average, obtain a lower response rate than experienced interviewers. With the consumer price index (CPI), for example, there are three interacting surveys that all use the same PSUs to take advantage of the benefit of experienced interviewers.

The real potential for ACS to provide assistance to household surveys is with respect to the MAF. Currently, the census address list, not comprehensively updated between censuses, is used until it is as much as 16 years out of date. For example, the CEX will use (essentially) the 1990 census address list until about 2006, when it begins implementing a new panel. The introduction of a continuously updated MAF is therefore a real advance. However, several questions can be raised: How good is the MAF going to be between censuses for seasonal housing units? How good is it going to be for new construction?

For the CPS, another consideration is that the ingoing and outgoing rotation groups are in neighboring segments, which helps in terms of panel correlations in composite estimation. If using the ACS means using different address lists, it may be difficult to maintain these correlations. One benefit may be to use symptomatic information from the ACS at a slightly aggregate level, such as blocks, for aiding in the efficiency of the sample design. Of course, the variance of and overall benefit from use of these symptomatic variables would be important to assess for this purpose.

If the ACS is going to be used to screen for various subpopulations, for example, to expand the CPS to increase the sample for specific race groups, the currency of the information collected when used for this purpose is an important consideration.[4] Another consideration, mentioned above, is that the 3-month window used in the ACS for data collection from a sample may miss some recent renters, which is an important issue for the CPI Rent Survey. In addition, any information on the ACS nonsampling error structure would be useful to have. Toward this goal, it would be useful to include the ACS with the proposed CPS-Census match, especially to understand within-

[4]The use of ACS for this purpose raises the hard problem of the sampling weight those people or households should get in subsequent analyses. Also, one needs to consider in the sample design that there will be both false positives and false negatives in the responses to the screening question.

household coverage and the characteristics of the population that are undercovered relative to the decennial census.

Possibly the next most important step was to address the relevant policy issues more seriously now that large-scale data collection is a reality, since these issues are extremely difficult. The two most difficult and important policy issues are data sharing and the use of the ACS for screening. An example is the National Health Interview Survey, which uses screening to support an area sample, which is used since the Census Bureau could not provide access to its address list for a list survey. This lack of access is a serious problem. Second, with respect to the process of adding a question or a set of questions to the ACS questionnaire and the related costs, if the decennial census is the guide, this approach will raise a substantial problem since adding questions to the long form has proved to be a flawed process. Furthermore, the costs are a worry. The decision on which questions to include is one that requires years to plan. One would need to start addressing this problem right now for the 2003 date of full ACS implementation.

FINAL POINTS

A number of opportunities for ACS and current household surveys to jointly benefit from each other were suggested. The extent to which these advantages can be obtained depends on various things, e.g., the correlations between ACS responses and responses on household surveys, and the impact of the addition of new ACS questions on the quality of the response to existing ACS questions. Therefore, the benefits cannot be determined before more is learned about the ACS. Also, the benefits of an updated MAF for redesigning household surveys is complicated by several factors; thus, the degree to which ACS can be beneficial is an empirical question and needs further work.

7

Calibration of the Long Form to ACS Output

Although there could certainly be additional uses of the output from the American Community Survey over time—some of which are described earlier—in the short term the ACS output is primarily intended as a (more timely) substitute for the decennial census long form. Given this, it is important to determine the effects that would be expected when switching from the long-form estimates to those from the ACS on various applications of long-form data. To analyze this, the Census Bureau is developing a "calibration" model of the long form in 2000 to the hypothesized full implementation of the ACS in 2000, based on ACS data collections prior to full implementation (described later), and the long-form output in 2000. The development of this calibration model will be challenging, since not only will individual-level matching of the long form to the ACS not be possible in 2000 due to the designs of the long form and the ACS data collections, but also the ACS sample sizes prior to full implementation will be substantially smaller than for the full implementation. This calibration model, in addition to clarifying ACS/long-form differences, will be used to understand the dynamics of change in various subject-matter areas (e.g., income, employment, health, welfare, education) between 2000 and later years in the decade by providing an analogue to a full implementation of the ACS in 2000.

Related to the issue of the effect on long-form applications of switching from the long form to the ACS, users have a need to understand the quality of output from the ACS (or any major data collection), since this understanding assists in the utilization of the estimates. Users need to have estimates of bias (often relative to some gold standard), variance (and associated confidence intervals),[1] and other summary error measurements of ACS at various levels of geographic aggregation and over time to best understand its utility for various applications. This type of information is useful, for example, in the development of models in which ACS information is combined with information from other sources, as discussed in Chapter 2. In addition to informing users about the quality of the ACS information, an evaluation of the quality of ACS output would help direct Census Bureau efforts toward improving the ACS over time. To understand the quality of the ACS output, a variety of evaluation methods will need to be identified. This issue was not directly addressed at the workshop, except that it was pointed out that, with calibrated long-form data, examining mean square errors for the ACS relative to the long form, possibly at some temporal or geographically aggregated level within various categories, could provide substantial information about the bias of the ACS. This was an approach taken in the panel study of small-area estimates of poverty. The remainder of this chapter focuses on the ACS/long-form calibration model.

The ACS data collections prior to full implementation in 2003 involve several steps. The Census Bureau is now in the field in 31 comparison sites, chosen on the basis of expected differences between the two data collection schemes, to examine ACS/long-form differences. This data collection began in 1999 and will continue through 2001. The Census Bureau plans on using these data to support site-specific analysis of the ACS/long-form differences. In addition, an annual 700,000 household national ACS sample will be collected; it started in 2000 and will end in 2002. In 2000, by design, no housing units will receive both the long form and the ACS questionnaire. The 31 comparison site data collection is designed to understand the factors that are associated with ACS/long-form differences. Then the calibration model will use these factors as covariates, along with the long-form responses, in models fitted using the annual 700,000 ACS sample collected from 2000-2002. One problem with this basic approach is that due to the size of the 2000-2002 ACS sample, it may be difficult to model the ACS/long-form differences in very small areas.

Charles Alexander stated that the Census Bureau needs substantial assistance in calibrating the long form to the ACS. It needs help both in understanding whether and how to do the calibration and in understanding the proper role of the calibrated numbers. In selecting a model to calibrate the 2000 long form to the ACS, methods similar to those used for small-area estimation using either the 1990 postenumeration survey or the 2000 integrated coverage measurement plan are under consideration. One important

[1]The variance estimates from the ACS will be complicated by the various weighting schemes discussed in Chapter 5.

difference is that the ACS problem has a large number of dependent variables of interest, rather than just the single one of undercoverage. He mentioned that there was some recent interest in comparing the results of various microsimulation models using long form and ACS information, and he supported more efforts in this direction to help understand the implications of this shift from the long form to the ACS.

RESEARCH DIRECTIONS

In his presentation, Jay Breidt pointed out that as the long form is the recognized standard for various economic and demographic analyses from 2000 and the plan is to use the ACS to replace the long form after 2000, it is critical to understand ACS/long-form differences that are due to methodological changes. He noted parallels between this calibration problem and that of remote sensing. In remote sensing, one is comparing the differences between a map based on remote sensing and ground truth, both of which are subject to measurement error, features on the map are also subject to substantial sampling variance, and there are some covariates to assist in understanding differences between ground truth and the map. There is also stratification, clustering, nonresponse, mode and instrument differences, a temporal displacement problem, and spatial displacement, all of which make the analogy relatively useful.

The calibration problem has several complexities. Who are the potential users of calibrated data? At the local level, the quality of the comparison is not that crucial, since one is looking for major changes. Certainly, those trying to draw inferences at larger levels of aggregation, possibly as input to models, could make use of a calibration of this sort. The data items most likely to be of interest are the long-form variables, both in a univariate and a multivariate sense. The interesting geographic domains are likely to include various moderate levels of aggregation of interest that have correspondence with census geography, since different users need data aggregated in different ways.

The idea of making use of a form of the model proposed for use in 1990 with the postenumeration survey is a leap, since that methodology is essentially univariate. This is a concern, because that methodology might not extend in a natural way to a multivariate setting. (A more natural analogy is with missing data models suggested in Clogg et al. [1991] and in Schafer [1997].) Furthermore, any sort of regression-type methodology, in which one uses point estimates, could lead to trouble, since point predictions are too smooth.

A possible alternative is to view this as a missing data problem in which one has the missing data structure illustrated in Table 7-1.

As this table indicates, the long form, with about 17 million housing units, overlaps with (roughly) 117,000 ACS housing units in 2001 and 2002

TABLE 7-1 Missing Data Structure (numbers in thousands)

	Housing Units			
Sample Included in	Long Form	ACS 2000	ACS 2001	ACS 2002
Long form alone	16,433			
ACS 2000 alone		700		
Long form and ACS 2001	117		117	
ACS 2001 alone			583	
Long form and ACS 2002	117			117
ACS 2002 alone				583
Total housing units	16,667	700	700	700

and with an additional 583,000 unmatched ACS housing units; in addition there are 700,000 unmatched ACS housing units in 2000. Development of the calibration model could be considered as a long-form missing data problem in which the long-form data for the 700,000 housing units are all missing, or one could think of it as an ACS missing data problem, in which one has 700,000 of the 3 million ACS questionnaires for 2000, and all of the associated long-form responses are also missing.[2] Considering this as a missing data problem leads one to consider some sort of imputation approach at some level of geographic aggregation, attempting to mimic the joint distributional properties of the long-form and ACS responses. It would be useful to attempt this at the lowest possible level of geographic aggregation, so that users could have flexibility in aggregating the estimates. This approach can get complicated very quickly, but one might consider constructing a semiformal model through the formation of imputation classes, possibly guided by the results of the study of the 31 comparison sites. As is typical, one might use hot-deck or some distance-function matching algorithm to form a complete data set. One might also use multiple imputation to provide an assessment of the uncertainty.

[2]The inability to carry out individual-level matching is unfortunate, though individual matching could be done with a 1- or 2-year lag given the current ACS design for 2000. This type of comparison forces one to model not just the differences between the data sources, but also the temporal differences. Ignoring the ability to link across years, which is problematic, one is restricted to models that only use marginal information. Carrying out such an imputation at the level of the individual household would not capture changes in the long-form and ACS sampling frames.

For the development of a formal model, one might begin with the following real or hypothesized data inputs (and accompanying notation): the partial ACS implementation, denoted by a, for 2000-2002, the long-form data, L_{2000}, for 2000, the hypothesized ACS under a full implementation in 2000, A_{2000}, the hypothesized ACS responded to by the entire country, α, and the hypothesized long form responded to by the entire country, λ_{2000}. The calibration problem is to predict A_{2000} given L_{2000} and a. This could be addressed, at least in theory, using Bayes' theorem, by computing the posterior distribution, (i.e., the probability distribution of the full ACS implementation in 2000 conditional on the long form and the partial ACS implementation). That conditional probability can be broken down into five factors: (1) the probability density associated with the long-form sampling model, $p(L_{2000}/\lambda_{2000})$; (2) the probability density associated with long-form responses given ACS responses, $p(\lambda_{2000}/\alpha)$, which is a crucial, problematic element of the calibration model; (3) and (4) the probability densities associated with the ACS sampling models, $p(A_{2000}/a, \alpha)$ and $p(a/\alpha)$; and (5) the ACS spatiotemporal model, $p(\alpha_{2000}, \alpha_{2001}, \alpha_{2002})$, which describes how the ACS would measure various quantities in various regions across time. (This model is analogous to that used by the Panel on Estimates of Poverty for Small Geographic Areas described above.)

This is an extremely complicated modeling exercise, although parts of it are relatively straightforward. For example, the sampling probabilities for the ACS and the long form are known. The last factor, the ACS multivariate response structure, would be very difficult to address, particularly since there is no information about how responses at the level of individual housing units change over time. Therefore, one would have to accomplish the modeling of this factor at some higher level of aggregation. The fitting of such a model would likely require computational methods, such as Markov chain Monte Carlo sampling, to estimate posterior means, posterior variances, and posterior quantiles and to replicate posterior predictions: that is, make multiple imputations.

There are at least two possibilities for the structure of the resulting imputed dataset. The first is to impute an ACS record for each long-form record, which would permit direct comparison of estimates at low levels of aggregation. There would also be the opportunity of making direct comparisons at the level of individual households. A second possibility would be to create a pseudo-ACS 2000 data set by augmenting the 700,000 ACS household records to get to the full ACS implementation sample size. This approach might require some weights to reflect information on the long form, and one might need replicates to capture variability, but it would support longitudinal analysis.

It is clear that long-form/ACS comparisons are extremely important. In making these comparisons, one wants to avoid confounding methodological change with true change. Response to the ACS and other factors might be quite dynamic, and if the basic mechanism underlying differences between the ACS and long form is not well understood, the calibration model might not be measuring what one wants in assessing differences, for example, between 2008 and 2000. Finally, whatever is done, uncertainty measures are needed, and the entire process needs to be thoroughly documented.

FINAL POINTS

Some proposed calibration models demonstrated the complexities faced by the Census Bureau in developing such a model to link the long form and the ACS. There is a need to understand small-area time dynamics in various ACS responses and to understand the causes of ACS and long-form discrepancies. Both of these suffer from a lack of information at the level of the individual household. The modeling of ACS and long-form discrepancies might need to be performed at a somewhat higher level of geographic aggregation, and some simple synthetic-type assumptions would then be used to "bring down" these estimates to lower geographic levels.

8

Conclusion

The thought pieces, discussant papers, and floor discussions all contributed to a productive interchange of ideas, with creative suggestions, not so much on how to resolve but rather on how to frame and provide perspectives to the many issues raised. Many of the questions raised have an empirical aspect, therefore answers will have to wait until more data are collected. The presenters provided key benefits in addressing the questions raised about alternative approaches and the underlying assumptions, illustrative methodologies used in comparable situations, and the advantages and disadvantages of these different methods.

One area of particular interest involved fund allocation programs, for which the ACS will provide increased opportunities for timely allocation of public funds at low levels of geographic aggregation. Participants argued for the importance of educating legislators about the unintended negative aspects of interactions of funding formulas and the distributional attributes of the estimates used in these formulas.

Participants generally agreed on the importance of the ACS, which under current plans (and assuming that the proposed budgets are realized) will become the main vehicle for collection of long-form-type data on a continuing basis. This potential role underscored the important issue as to who sets the policies for access to the ACS. That is, who will decide on the content of the ACS questionnaire, either as supplements, special subject modules, or as a series of screener questions to target specific groups, and at what cost? Will the ACS be treated similarly to other census surveys, that is, subject to Census Bureau constraints on data sharing, especially addresses of target groups of interest to other agencies, due to concerns about privacy and confidentiality? In effect, will the ACS have all of the characteristics of "census confidential" and the attendant problems of restricted access to microdata by outside agencies?

The question of how often to redesign household surveys on the basis of the updated ACS address list is also a policy matter, though it will have methodological features as well. Quality assurance and timing of the associated master address file updates are also important elements and therefore will require further attention.

In his concluding remarks as chair, John Rolph urged that this workshop be the first step in a continuing process of addressing the many statistical issues raised relative to undertaking the ACS. He said that the Committee on National Statistics would like to be as helpful as possible as the process of designing and fielding the ACS unfolds. Such activities might include committee-sponsored workshops focused on specific issues that need to be addressed at particular points in the design process. He invited the workshop participants to send in their ideas and suggestions for further activities.

Although the workshop was successful in generating valuable initial ideas and discussion on how to address some interesting and difficult methodological problems raised by the ACS, it is important to note that not all important methodological problems were raised. Several issues that were not raised or mentioned only briefly were: (1) additional approaches to combining information from the ACS, household surveys, and administrative records that could also be examined, especially variance component models; (2) methods to treat undercoverage in the ACS, particularly methods for using demographic analysis to address undercoverage in the ACS, and also to use the ACS to improve demographic analysis; (3) further examination of the issues raised through use of incompatible definitions in the ACS, the decennial census, and household surveys; (4) the development of estimates that (a) sum to estimates at higher levels of geographic aggregation and (b) more closely approximate direct estimates at higher levels of aggregation, along with the release of direct estimates at higher levels of aggregation—in the event that aggregate estimates are not constrained to (approximately) equal direct estimates (and also the release of direct estimates at lower levels of aggregation for analysis purposes); (5) the evaluation of the quality of the estimates from the ACS, especially given that no long form is planned after the 2000 census and therefore external evaluation opportunities will be very limited; (6) the need for specific ideas for developing models for borrowing strength from household surveys and administrative records to assist the ACS in the estimation of various outputs; (7) weighting ACS output to produce estimates that are consistent with the recognized estimates from existing household surveys; (8) using information from the ACS to develop models to effectively reduce the sample size requirements of existing household surveys and still produce estimates of the same quality; and (9) coordinating estimates between the ACS and the decennial census short-form in 2010, 2020, and so on for statistics that can be generated from the short-form items alone. These issues were at most touched on at the workshop.

In addition, as noted above, some of the issues that were discussed cannot be fully addressed until the full ACS is collected, and they may require more research. Examples include (1) the development of time-series combination-of-information models; (2) problems raised through the planned addition of questions to the ACS questionnaire, especially the effect on the quality of the information collected for the other questions; and (3) the formation of the model needed to calibrate the long form to the ACS.

This workshop helped to identify a number of interesting and important problems, many of which will have much broader application than the ACS. The workshop succeeded in making the attendees aware of these problems, and raising for discussion promising avenues for their solution.

Appendix

An Example of Combining Information

The Panel on Estimates of Poverty for Small Geographic Areas is providing assistance to the Census Bureau in its development of model-based small-area estimates of the number of children living in poverty, which is needed for input to formulas allocating substantial funds to counties and school districts to address the needs of disadvantaged children under Title I of the Elementary and Secondary Education Act. Prior to this recent work, Title I had used the most recent census long-form (sample-based) counts to allocate funds, which produced estimates that were as much as 12 years out of date. Model-based estimates at the county level (for 1993 and every two years into the future) and at the school district level (for 1995 and every two years into the future) are now being used in place of the census long-form estimates. (Contemporaneous direct estimates cannot be supported with current survey or administrative data.)

The county-level model (used for both 1993 and 1995 estimates) is an excellent example of how current best practice permits one to combine data from various sources. These model-based estimates make use of a county-level regression model, which used as the dependent variable a logarithmic transformation of the current number of children in poverty, measured by a 3-year average (to reduce variance) from the Current Population Survey (CPS).[1] This regression model makes use of (logarithmic transformations of)

[1]Since the CPS does not have samples in all counties, the regression model was fit using only about 1,300 counties.

the following covariates for a given county: the number of child exemptions reported by families in poverty on tax returns, the number of people receiving food stamps, the estimated population under age 18, the number of child exemptions on tax returns, and the number of poor school-age children in the county from the previous census. For counties with CPS sample households and with poor children in the sample, a linear combination (formally, empirical Bayes' shrinkage) of the direct estimate from the CPS and the model prediction from the regression model is computed; otherwise, the model prediction alone is used. After being transformed back to the original scale (assisted by an adjustment for transformation bias), the final county-level estimates of the number of poor school-age children are then ratio adjusted so that within each state the county-level estimates sum to a separately modeled state-level estimate.

The state-level model was developed in a similar manner to the county-level model. The state-level regression model uses as the dependent variable the estimated proportion of poor school-age children as measured by the CPS (using only a single year, given the larger sample size at the state level). The covariates used in this regression model are the proportion of child exemptions reported by families in poverty on tax returns, essentially the proportion of people receiving food stamps; the proportion of persons under 65 years of age who did not file a tax return; and the residual from the analogous census regression of the proportion of poor school-age children from the most recent census on the other three covariates contemporaneous with that time period. As in the county-level model, a linear combination (again based on empirical Bayes' methods) of the direct CPS estimate and the model prediction is used (though in practice, the estimated model error variance has been so low that the regression prediction has usually received the full weight).

For income year 1995, the requirement was to provide poverty estimates at the level of school districts. At this low level of geographic aggregation, the above approach based on regression modeling cannot be used, since corresponding data, especially for the covariates, does not now exist on a uniform basis. Therefore, the Census Bureau adopted a simple shares approach, distributing 1995 county-level estimates of the number of poor school-age children to school districts according to the school district to county poverty shares, measured using the 1990 census long form (ignoring some minor complexities).

In the future, the ACS is expected to play an important role in the estimation of the number of school-age children in poverty at the school district level, either by direct estimation based on aggregation of data over several years, or by combination in one of several ways, with other data series that are and might become available at the school district level (e.g., data on food stamp participation, data on school lunch participants, and poverty rates estimated from tax filers.) It is quite likely that even with the large sample size of the ACS, small-area estimation techniques will be required to combine information over time and geography to develop high-quality estimates. Issues of comparability of the decennial census, the CPS, and the ACS will need to be addressed, as will any changes in tax or welfare programs that affect data comparability over either time or geography.

References

Alexander, Charles H., Scot Dahl, and Lynn Weidman
1998 Making estimates from the American Community Survey. In *Proceedings of the Social Statistics Section*. Alexandria, VA: American Statistical Association.

Clogg, C., D. Rubin, N. Schenker, B. Schultz, and L. Weidman
1991 Multiple imputation of industry and occupation codes in census public-use samples using Bayesian logistic regression. *Journal of the American Statistical Association* 86:68-78.

Kish, L.
1981 Population counts from cumulated samples. Pp. 5-50 in Congressional Research Service, *Using Cumulated Rolling Samples to Integrate Census and Survey Operations of the Census Bureau*. Prepared for the Subcommittee on Census and Population, Committee on Post Office and Civil Service, U.S. House of Representatives. Washington, DC: U.S. Government Printing Office.

National Research Council
1992 *Combining Information: Statistical Issues and Opportunities for Research*. Panel on Statistical Issues and Opportunities for Research in the Combination of Information, Committee on Applied and Theoretical Statistics, Board on Mathematical Sciences, Commission on Physical Sciences, Mathematics, and Applications, National Research Council. Washington, DC: National Academy Press.

1998 *Small-Area Estimates of School-Age Children in Poverty*. Panel on Estimates of Poverty for Small Geographic Areas, Committee on National Statistics, Commission on Behaviorial and Social Sciences and Education, National Research Council. Washington, DC: National Academy Press.

1999 *Small-Area Estimates of School-Age Children in Poverty*. Panel on Estimates of Poverty for Small Geographic Areas, Committee on National Statistics, Commission on Behaviorial and Social Sciences and Education, National Research Council. Washington, DC: National Academy Press.

2000 *Small-Area Income and Poverty Estimates: Priorities for 2000 and Beyond.* Panel on Estimates of Poverty for Small Geographic Areas, C.F. Citro and G. Kalton, eds., Committee on National Statistics. Washington, DC: National Academy Press.

Office of Federal Statistical Policy and Standards

1978 *Statistical Policy Working Paper 1: Report on Statistics for Allocation of Funds.* Prepared by the Subcommittee for Allocation of Funds, Federal Committee on Statistical Methodology. Washington, DC: U.S. Department of Commerce.

Rosenbaum, Paul R.

1987 Model-based direct adjustment. *Journal of the American Statistical Association* 82:387-394.

Schafer, J.L.

1997 *Analysis of Incomplete Multivariate Data.* London: Chapman and Hall.

Scott, A.J., and T.M.F. Smith

1974 Analysis of repeated surveys using time series methods. *Journal of the American Statistical Association* 69:674-678.