This PDF is available from The National Academies Press at http://www.nap.edu/catalog.php?record_id=18259

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

ISBN
978-0-309-26797-7

189 pages
6 x 9
PAPERBACK (2013)

Daniel L. Cork, Rapporteur; Committee on National Statistics; Division on Behavioral and Social Sciences and Education; National Research Council

Add book to cart | Find similar titles | Share this PDF

Visit the National Academies Press online and register for...

- Instant access to free PDF downloads of titles from the
  - NATIONAL ACADEMY OF SCIENCES
  - NATIONAL ACADEMY OF ENGINEERING
  - INSTITUTE OF MEDICINE
  - NATIONAL RESEARCH COUNCIL
- 10% off print titles
- Custom notification of new releases in your field of interest
- Special offers and discounts

Distribution, posting, or copying of this PDF is strictly prohibited without written permission of the National Academies Press.
Unless otherwise indicated, all materials in this PDF are copyrighted by the National Academy of Sciences.
Request reprint permission for this book

Copyright © National Academy of Sciences. All rights reserved.

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

Benefits, Burdens, and Prospects of THE American Community Survey

Summary of a Workshop

Daniel L. Cork, *Rapporteur*

Committee on National Statistics

Division of Behavioral and Social Sciences and Education

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

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

Copyright © National Academy of Sciences. All rights reserved.

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

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

The project that is the subject of this report was supported by contract no. YA1323-11-CN-0033 between the U.S. Census Bureau and the National Academy of Sciences. Support of the work of the Committee on National Statistics is provided by a consortium of federal agencies through a grant from the National Science Foundation (No. SES-1024012). 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-13: 978-0-309-26797-7
International Standard Book Number-10: 0-309-26797-8

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

Copyright 2013 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America

Suggested citation: National Research Council. (2013). *Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop.* Daniel L. Cork, rapporteur. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Copyright © National Academy of Sciences. All rights reserved.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

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

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

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

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

www.national-academies.org

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

STEERING COMMITTEE FOR THE WORKSHOP ON THE BENEFITS (AND BURDENS) OF THE AMERICAN COMMUNITY SURVEY

LINDA GAGE (*Co-Chair*), Independent Consultant and California Department of Finance (retired)
KEN HODGES (*Co-Chair*), Nielsen, Ithaca, New York
LINDA JACOBSEN, Population Reference Bureau
PATRICK JANKOWSKI, Greater Houston Partnership
JOAN NAYMARK, Independent Consultant and Target Corporation (retired), Minneapolis, Minnesota
RICHARD RATHGE, Departments of Agribusiness and Applied Economics and Sociology and Anthropology, North Dakota State University

DANIEL L. CORK, *Study Director*
CONSTANCE F. CITRO, *Senior Program Officer*
AGNES GASKIN, *Administrative Assistant*

Copyright © National Academy of Sciences. All rights reserved.

COMMITTEE ON NATIONAL STATISTICS
2012–2013

LAWRENCE D. BROWN (*Chair*), Department of Statistics, The Wharton School, University of Pennsylvania
JOHN M. ABOWD, School of Industrial and Labor Relations, Cornell University
DAVID CARD, Department of Economics, University of California, Berkeley, and Labor Studies Program, National Bureau of Economic Research
ALICIA CARRIQUIRY, Department of Statistics, Iowa State University
CONSTANTINE GATSONIS, Department of Biostatistics and Center for Statistical Sciences, Brown University
JAMES S. HOUSE, Survey Research Center, Institute for Social Research, and Gerald R. Ford School of Public Policy, University of Michigan
MICHAEL HOUT, Department of Sociology, University of California, Berkeley
SALLIE KELLER, Science and Technology Policy Institute, Washington, DC
LISA LYNCH, Heller School for Social Policy and Management, Brandeis University
SALLY MORTON, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh
EDWARD H. SHORTLIFFE, New York Academy of Medicine, New York City
HAL STERN, Department of Statistics, University of California, Irvine
JOHN H. THOMPSON, NORC at the University of Chicago
ROGER TOURANGEAU, Westat, Rockville, Maryland

CONSTANCE F. CITRO, *Director*

Copyright © National Academy of Sciences. All rights reserved.

Acknowledgment of Reviewers

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the Report Review Committee of the National Research Council. The purpose of this independent review is to provide candid and critical comments that assist the institution in making its 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 process.

We thank the following individuals for their review of this report: Kathleen Thiede Call, State Health Access Data Assistance Center and Health Policy and Management, University of Minnesota; John Iceland, Department of Sociology, Pennsylvania State University; Dan Kasprzyk, Center for Excellence in Survey Research, NORC at the University of Chicago; Paul Overberg, Database Editor, *USA TODAY*; Richard Rathge, Departments of Agribusiness and Applied Economics and Sociology and Anthropology, North Dakota State University; and Joanna Turner, State Health Access Data Assistance Center and Health Policy and Management, University of Minnesota.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the content of the report, nor did they see the final draft of the workshop summary before its release. The review of this report was overseen by Joseph Salvo, Population Division, New York City Department of City Planning. Appointed by the National Research Council, he was responsible for making certain that the independent examination of this summary was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of the summary rests entirely with the author and the institution.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

Contents

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

List of Figures

Copyright © National Academy of Sciences. All rights reserved.

List of Tables

Copyright © National Academy of Sciences. All rights reserved.

List of Boxes

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

– 1 –

Introduction

A surprising degree of uncertainty surrounds one seemingly simple point of fact about the U.S. Census Bureau's American Community Survey (ACS): When did the ACS begin? Some statements cast the ACS as a very new survey—largely through comparison with the 2010 decennial census, in which the ACS replaced the "long-form" questionnaire administered in previous decades to a sample of respondents. Others claim a longer lifetime for the concepts and questions of the ACS, alternately describing the ACS as dating back to the first decennial census in 1790, which started the precedent of asking for more detail than a simple headcount; to 1940 and the first major use of statistical sampling for some content in a decennial census; or to 1960, the first census in which a full-fledged long-form questionnaire was administered.

In fact, a case can be made for all of the years listed above—and others—because some of the ideas behind the ACS are indeed as old as American census-taking itself. In its content, the 1790 census stuck to the basic information suggested in the Constitution's enumeration clause, but its execution established a basic, subtle precedent that underlies today's ACS and census: data on *aggregates* such as communities, states, or the nation as a whole are acquired by asking questions of *individuals* (in the 1790 case, asking heads of household for a name-by-name accounting of family members). The range of those questions began to widen over the course of the 19th century—more detailed age categories in 1800, rough indicators of occupation in 1810, queries on disability (deaf or blind) in 1830, and later expansion into topics ranging from educational status to amounts of wages earned and taxes paid. The 1940 census began the practice of covering some of the now-sprawling content of the census on a sample basis, asking certain supplemental questions only of those individuals appear-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

ing on certain lines of the census schedule. The 1960 census formalized the process, creating a separate long-form questionnaire; most households received the short-form questionnaire containing only basic data items while a roughly one-sixth sample received the more detailed long form.

As the 2000 census approached, a notion originally advanced in the early 1940s as a possible replacement for the census began to gain traction, albeit as a replacement for the long-form sample rather than the entire census. Sociologist and demographer Philip Hauser, working at the Census Bureau in 1941, is generally credited with first advancing the idea of an "annual sample census" to generate small-area estimates. Increasing calls for more timely estimates led, in part, to Congress amending Title 13 of the U.S. Code in 1976 to authorize a mid-decade census in years ending in 5, but no action was taken under the new law. In 1981, statistician Leslie Kish proposed mounting a continuous sample survey, creating "rolling" samples by accumulating one or more years of collected data and using those samples to create estimates; the Decade Census Program suggested by Roger Herriot and others in the late 1980s advanced similar notions of using rolling samples for state-level estimates (Herriot et al., 1989). Drawing from these ideas, the Census Bureau's Charles Alexander took the lead in assessing options to replace the long-form sample, which had become a greater drain on census resources (and whose quality had arguably slipped, conducted as an adjunct to the main census). Alexander's proposed Continuous Measurement Survey would pick up a new moniker—the American Community Survey—as it entered very early pilot testing in four counties in 1996. The number of test sites (counties or groups of counties) was slowly ramped up until 2000. In that year, one of the formal experiments associated with the 2000 census was to dramatically scale up ACS coverage to roughly one-third of all counties—mainly as a test of feasibility of conducting both the ACS and a decennial census simultaneously. Bolstered by favorable results, the ACS continued collection at these so-called Census 2000 Supplementary Survey levels (after the name of the 2000 census experiment) until 2005, when the ACS began operations in all counties nationwide.

To secure approval for the full-scale ACS, the ACS was made an integral part of plans for the 2010 decennial census—the basic bargain being that shifting the historical long-form content to the ACS would free the main 2010 census to be conducted as a "short-form only" count. In this spirit, the content and questions of the ACS were closely patterned after the long-form sample questionnaire used in the 2000 census. As years have passed, new questions have gradually been added to the survey. For instance, questions on health insurance coverage (discussed more fully in Chapter 2) and marital history were added after testing in 2006, and the Census Bureau plans to add a question on Inter-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

net access beginning in 2013.[1] As implemented in 2012, the ACS questionnaire contains 75 numbered questions:

- A lead respondent for the household is guided through the process of answering six roster questions about each member of the household; these questions include each person's name, relationship to the lead respondent, age, and race and Hispanic origin.[2]
- The lead respondent is then asked 21 "housing" questions; these include attributes of the housing stock (e.g., age of construction of the building, presence or absence of plumbing and cooking facilities, and the amount spent in the previous month for electricity) and general characteristics of the household (e.g., monthly rent, whether anyone in the household holds the mortgage on the property, and whether anyone in the household received Supplemental Nutrition Assistance Program/Food Stamp benefits in the previous year).
- Finally, up to 48 numbered "person" questions are asked of *each* person in the household (some questions may be skipped depending on earlier responses); these questions include U.S. citizenship, language other than English spoken at home, place of residence 1 year ago, fertility (whether the person gave birth in the past year), employment status, and various types of income in the past year.

In brief, the basic parameters of the ACS are that on the order of 295,000 housing units—spanning all counties in the United States—are selected for the ACS sample every month (or roughly 3.5 million housing units per year).[3] Those sample units are contacted by mail during their first month in the sample; in the second month, those who failed to respond by mail are selected for contact/interview by telephone, if a phone number can be determined; in the third month, a subsample of nonrespondents by either mail or phone are then selected for personal interview by Census Bureau field staff. In this way, the Census Bureau is continuously collecting data across all three modes: the data actually collected in month *M* include mail returns from the new housing units added to the sample in month *M*, phone interviews with some people from the

[1]The Census Bureau archives the questionnaires used in various years of ACS implementation on its website at http://www.census.gov/acs/www/methodology/questionnaire_archive/. Throughout this summary, when appropriate, question text is excerpted from the 2012 version of the questionnaire, the direct link to which is http://www.census.gov/acs/www/Downloads/questionnaires/2012/Quest12.pdf.

[2]This full detail is only asked about five people in the household; only name, sex, and age are asked about Persons 7–12, and the additional detail may be collected about these people in a follow-up phone interview.

[3]The full-scale ACS operated on a monthly sample of about 250,000 households until the increase to 295,000 with the June 2011 mailing. A separate sample of roughly 200,000 residents of group quarters—places like college dormitories, nursing homes, military barracks—is also covered in the ACS each year. The unique challenges presented by group quarters are largely beyond the scope of this workshop, but are addressed in detail by National Research Council (2012).

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

month $M-1$ sample (plus late mail returns from that group), and personal visit interviews with some people from the month $M-2$ sample (also including late mail returns from that group). The Census Bureau intends to implement an Internet response option to the ACS beginning in January 2013, rather than relying solely on return of the mailed paper questionnaire.

An important nuance of the ACS, as currently implemented, is that response to the survey is mandatory under law—inheriting this attribute from the ACS's role as replacement for the former (mandatory) long-form sample and as a supplement or complement to the decennial count.[4] Nonrespondents are subject to a fine,[5] and messages emphasizing the mandatory nature of the survey are prominently displayed on ACS mailing materials and noted on the questionnaires. More will be said of this requirement—and proposed changes to it—below in Section 1–B.

The data from the ACS's monthly samples are cumulated and averaged over 1 or more years—with a single year's collection of data able to provide information for large population subgroups and several years of data needed to yield reliable estimates for small groups. Specifically, the estimates currently produced by the ACS use three time windows:

- 1-year period estimates, for geographic and demographic groups of population 65,000 or greater;
- 3-year period estimates, for groups with populations between 20,000 and 65,000; and
- 5-year period estimates, for groups with populations less than 20,000—a category including the vast majority of cities and counties in the United States, as well as smaller geographies such as census tracts.

Being period estimates, and pooling several months of data, ACS estimates differ from the estimates produced by the former long-form sample, which could be interpreted as point estimates for the decennial census reference date (April 1). In addition to the conduct of the decennial census itself, the year 2010 was

[4] 13 USC § 193 states that: "In advance of, in conjunction with, or after the taking of each census provided for by this chapter, the Secretary [of Commerce—and by extension the Census Bureau—] may make surveys and collect such preliminary and supplementary statistics related to the main topic of the census as are necessary to the initiation, taking, or completion thereof."

[5] Like the decennial census, the defined penalty for "refus[ing] or willfully neglect[ing] . . . to answer, to the best of his knowledge, any of the questions on any schedule submitted to him in connection with any census or survey provided for" elsewhere in the law (including § 193, discussed in the previous footnote) is a fine of not more than $100 (13 USC § 221(a)). The penalty for "willfully giv[ing] any answer that is false" is a fine of not more than $500 (13 USC § 221(b)). However, under the terms of the Sentencing Reform Act of 1984, a person guilty of either of these infractions could be subject to a fine of "not more than $5,000" (18 USC § 3571). Prior to enactment of P.L. 94-521 in 1976, violators were also subject to imprisonment: up to 60 days for nonresponse and up to 1 year for false answers. That said, the Census Bureau's longstanding position is to emphasize the reasons for responding to the ACS rather than the possible legal penalties, and has refrained from charging ACS nonrespondents.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

pivotal for the Census Bureau in that it was the first year in which the full suite of ACS products was available for the entire nation—including many localities' first real glimpse at ACS data, through the 5-year 2005–2009 estimates. The Census Bureau releases ACS estimates primarily through its American FactFinder website at http://factfinder.census.gov.

The unique challenges and opportunities presented by the ACS data were explored by a previous National Research Council (2007b) panel. However, given its dates of operation, that panel's work was necessarily abstract and hypothetical. So too was the Census Bureau's own work to prepare data users, including the production of a series of "Compass" handbooks for different types of data users.[6] Only with the release of actual data, and following the reaction to the data and to their accompanying and prominently featured margins of error, did the ACS become real and concrete to its potential users. Now in its second decade of large-scale implementation, the release of the first sets of full data also marks an opportune time to assess the ACS, its uses, and its demands.

1-A THE WORKSHOP ON THE BENEFITS (AND BURDENS) OF THE AMERICAN COMMUNITY SURVEY

The U.S. Census Bureau requested that the Committee on National Statistics (CNSTAT) convene a:

> workshop on the benefits to a broad array of non-federal users of the data products from the U.S. Census Bureau's American Community Survey (ACS). The workshop will also address the burden on the American public of responding to the ACS questions. A workshop on the benefits of the ACS is timely because the survey has just completed its first full cycle of releasing 1-, 3-, and 5-year period estimates (for 2009, 2007–2009, and 2005–2009, respectively), and there is need to take stock of user experience with the data and to identify priority uses for the future. The workshop is part of a larger review the Census Bureau is conducting to assess the overall mission, vision, goals, and objectives of the ACS. The Bureau's review also includes examining the data products along with the methods for conducting and managing the survey.

Pursuant to this charge, CNSTAT convened a steering committee for the workshop, which met in March 2012 to discuss the task and craft the structure of the workshop.

The steering committee made two major decisions that governed the final shape of the workshop. The first was to make the notion of ACS "burden" an integral part of the proceedings and to extend it beyond the single issue of time

[6]These "Compass" guides included, for instance, volumes targeted at state and local government users (U.S. Census Bureau, 2009), high school teachers (U.S. Census Bureau, 2008b), members of Congress and congressional staff (U.S. Census Bureau, 2008a), and business users (U.S. Census Bureau, 2008c).

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

spent by respondents in completing the survey. The charge of the workshop emphasizes the benefits of the ACS and its title retained the wording used at the project's inception—literally expressing "burdens" as a parenthetical. But the steering committee sought to ensure that the challenges presented by the ACS were given an open and prominent airing, and decided to seek presentations and discussions concerning more general types of "burden" in the ACS. For instance, from the perspective of respondents, the ACS and its detailed questions raise privacy and confidentiality concerns that might serve as an impediment to response altogether. For data users, the ACS presents a significant communication burden: explaining "new" concepts (the interpretation of estimates based on multiple years of data) and old ones (margins of error) to potentially skeptical audiences. A final example of a broader burden created by the ACS follows directly from its design: Large, populous geographic units have access to a wealth of ACS estimates of 1-, 3-, and 5-year vintages while rural areas and smaller population groups face a relative scarcity of ACS estimates, necessarily waiting for 5-year accumulations that may still have very high standard errors. These types of "burden" considerations shaped the steering committee's selection of presenters within sessions; to give the issues a dedicated airing, the committee also sought to carve out a specific section of the program for a structured discussion of various types of "burden," recognizing that a single presentation or paper was unlikely to be sufficient.

The second was to cast a net as wide as possible to collect perspectives on ACS uses, and to make the workshop's agenda book (background briefing materials) a virtual poster session in order to spotlight more voices than the limited number of speaking slots could accommodate. The steering committee drafted a short "feeler" notice, asking ACS users to write back with a short description of how (and how often) they use ACS data products, and for what purposes. This feeler notice was distributed through a variety of channels, email lists, and contact networks, and yielded dozens of expressions of interest. The steering committee filled most of the slots in its working agenda using these submissions. Later, this group of feeler-notice respondents was contacted again and asked to contribute short written thoughts for the meeting agenda book—whether a particularly interesting "case study" of ACS data usage or an expanded "user profile" describing how individual users work with the ACS (and work around any of its shortcomings). In all, this background briefing book included about 30 submissions.

The background book (called "case study book" for short) and the presentations from the workshop are both available on the CNSTAT website (http://www.nas.edu/cnstat).[7]

[7]The specific page link to the workshop materials is http://sites.nationalacademies.org/DBASSE/CNSTAT/ACS_Benefits_Burdens.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

The workshop itself was held on the afternoon of June 14 and the full day of June 15, 2012, at the National Academies' Keck Center in Washington, DC. The workshop drew roughly 80 attendees across both days, with the first afternoon's proceedings drawing a particularly strong crowd. The final agenda for the workshop and a list of the contributors to the accompanying case study agenda book are reprinted in Appendix A.

1–B "THE ELEPHANT IN THE ROOM": LEGISLATIVE CONTEXT FOR THE WORKSHOP

While the workshop agenda was under development—and independent of the workshop and its charge—the ACS drew sudden legislative attention in spring 2012. Ultimately, this attention would set a most unusual context for the workshop: This workshop on the uses of the ACS took place slightly more than one month after one chamber of Congress voted to effectively terminate the survey.

Multiple bills filed in the 111th Congress sought to eliminate or weaken the requirement that response to the ACS be mandatory under law. H.R. 3131 (introduced by Rep. Ted Poe of Texas) would remove any penalty "for refusing or willfully neglecting to answer questions" on the ACS other than respondent name, contact information, "the number of people living or staying at the same address," and the response date. H.R. 5046 (introduced by Rep. Todd Akin of Missouri) would take a stronger line, making the only mandatory question on the ACS (and the decennial census) to be "the number of individuals living in such individual's residence."[8] Though the bills acquired some cosponsors (35 for the Poe bill, 8 for the Akin bill), neither advanced beyond referral to subcommittee during the 111th Congress.

In the 112th Congress, Rep. Poe reintroduced his legislation in March 2011 (now numbered H.R. 931), and it was referred to two subcommittees of the House Judiciary Committee as well as to the Oversight and Government Reform Committee (which has primary jurisdiction over the Census Bureau). After no action for 1 year, the House Oversight subcommittee with census jurisdiction[9] held a hearing on the *topic* of a voluntary (rather than mandatory) ACS on March 6, 2012. In addition to Rep. Poe, Census Bureau Director Robert Groves and three ACS data users from the business and economic development

[8]Filed on April 15, 2010, the Akin bill—even under the speediest of circumstances—could not have applied to the 2010 census, which was then well under way. It would have required placing "on the front of each [questionnaire or survey] in a conspicuous manner" an advisory statement of the form: "Constitutionally, in responding to this survey, you are only required to provide the number of individuals living in your residence. Answers to all other questions contained within this survey are optional."

[9]Formally, the Subcommittee on Health Care, District of Columbia, Census and the National Archives.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

sectors were invited to testify. Because the hearing was merely on the subject matter, and not a hearing or meeting on the bill itself, no further action was taken on the bill in the immediate wake of the hearing.

Several weeks later, on May 9, two amendments were offered during House floor consideration of H.R. 5326, the fiscal year (FY) 2013 Commerce, Justice, Science, and Related Agencies appropriations bill including funding for the Census Bureau. Both amendments took the form of policy "riders"—not explicitly striking or replacing a specific funding number, but simply prohibiting the use of any funds in the bill for certain purposes. The first amendment (introduced by Rep. Poe) would bar the use of funds in the bill to enforce penalties for nonresponse to the ACS, effectively making the survey voluntary rather than mandatory.[10] The second, offered by Rep. Daniel Webster of Florida, would prohibit use of the funds to "conduct the survey" altogether. Under the rules governing debate on the bill, consideration of each amendment was limited to five minutes on each side; the central arguments made in favor of the amendments were concern for the intrusive nature of some of the ACS questions and the argument that the detail of the ACS queries exceeds the constitutional mandate for the decennial census. After those short debates, the Poe amendment was judged to have passed by voice vote and the Webster amendment was ultimately passed by a 232–190 vote.[11] The next day, the House wrapped up consideration of the whole bill and passed it, as amended.

In the wake of the House vote, a version of the standalone Poe bill (to make the ACS voluntary) was introduced in the Senate by Sen. Rand Paul of Kentucky. Neither the Paul bill nor the Senate's Commerce, Justice, and Science appropriations bill for FY 2013 have progressed to further consideration at this writing in summer 2012; the prospects for completing work on the FY 2013 bills could reasonably be described as uncertain heading into a presidential election and the closing sessions of the 112th Congress.[12]

Introducing the workshop on behalf of the steering committee, CNSTAT Director Constance Citro noted that these recent legislative developments were an "elephant in the room" for the workshop. She noted that the workshop was

[10]Both the Poe amendment and the full bill, H.R. 931, speak only to the penalty defined in 13 USC § 221(a)—for "willfully refus[ing] or neglect[ing]" to provide a response. They let stand the penalty applying under 13 USC § 221(b), which penalizes giving willfully *false* answers.

[11]There was no request to hold a recorded vote after the chair ruled that the yeas had prevailed on the voice vote on the Poe amendment and—initially—the Webster amendment appeared to carry by voice vote as well. Rather than immediately call for a recorded vote on the Webster amendment, floor managers instead asked special dispensation to let a late-arriving speaker comment on the amendment—after which the chair moved on to other amendments. Shortly thereafter, opponents of the Webster amendment prevailed on the bill managers to vacate the voice vote and instead hold a recorded vote, which produced the 232–190 result.

[12]The Senate appropriations subcommittee's mark for the Census Bureau gave the Bureau its full request for FY 2013, including some funds for both the 2010 and 2020 decennial censuses and the 2012 economic census that were trimmed in the House version. As of early December 2012, no further congressional action had been taken on the relevant FY 2013 appropriations bills.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

not developed, and should not be interpreted, as a response to the legislation or an advocacy, one way or the other, for the ACS. Rather, the workshop is intended to document the way the ACS estimates are being used and to construct a portrait of the nonfederal ACS user base. She noted that the steering committee put no limits on what the workshop presenters (or case study contributors) say, and she encouraged the workshop participants to approach the session with the same candor.

The legislative developments that would either kill or impair the ACS—rolling out so closely before a workshop on the ACS—did not materially affect the content or the structure of the workshop, but they undeniably shaped its context and climate. The continued *existence* of the survey became a high-level "benefit" of the ACS that presenters wanted to address in their presentations. Likewise, the discussion sessions following topic blocks at the workshop that might—in different times—have involved more probing of the challenges or burdens of ACS data in specific applications instead reflected the underlying concerns: What would you do if the ACS had a lower sample size? What, if any, other data sources could you use if the ACS were to go away? And, could individual state agencies or private businesses generate similar data to meet basic needs? Hopefully, a full accounting of the workshop confirms that it did not become an advocacy platform in any direction—both benefits and deficiencies with the data were given ample time—but it also reflected the underlying tension over the basic prospects for the ACS in 2013 and beyond.

1-C REPORT OVERVIEW

This report has been prepared as a factual summary of what occurred at the Workshop on the Benefits (and Burdens) of the American Community Survey. The workshop steering committee's role was limited to planning and convening the workshop. Accordingly, the views contained in this report are those of individual workshop participants and do not necessarily represent the views of all workshop participants, the steering committee, or the National Research Council.

This workshop summary largely follows the topic blocks that were used in scheduling the workshop, though some rearrangement has been made when that seemed logical. The most prominent such rearrangement is the chapter describing the workshop's dedicated panel discussion on various aspects of burden associated with the ACS. The workshop steering committee deliberately scheduled that session for the rough midpoint of the workshop—the morning of the second day—so that it might draw from themes raised from some of the user presentations and infuse discussions of others. However, for purposes of this summary, it makes logical sense to defer the summary of this session to the end, in Chapter 8, so as not to disrupt the flow of the previous chapters. There, it

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

is also combined with the closing remarks by former Census Bureau Director Steve Murdock and the final threads of discussion at the workshop.

Chapter 2 summarizes presentations from the kickoff session of the workshop, focused on the applications of the ACS in planning for health care and transportation services. Based primarily on the experiences of nonprofit organizations and their use of the ACS, Chapter 3 describes the workshop block on ACS uses in planning for social services and preparing for (and responding to) natural and manmade disasters. To close the first afternoon of the workshop, the steering committee sought perspectives from three users in the print and online media; that discussion is in Chapter 4. The core of the second day of the workshop was oriented more around different "sectors" of ACS data users than specific topics; the experiences of state, local, and tribal government users and of business, economic development, and "data aggregator" users are discussed in Chapters 5 and 6, respectively. Chapter 7 turns to one important use of the ACS that is explicitly spelled out in federal law and regulation—to ensure compliance by state and local governments with the terms of the Voting Rights Act—as well as to related legal, political, and social equity uses of the ACS. The previously mentioned panel discussion on various aspects of burden associated with the ACS and the capstone discussion of the workshop are summarized in Chapter 8.

Copyright © National Academy of Sciences. All rights reserved.

– 2 –

Planning Health Care and Transportation Using the ACS

To kick off the workshop, the steering committee chose to profile two high-visibility policy areas—ones involving decisions and decision makers at all levels of government and actors in the public, private, and nonprofit sectors. These chosen topic areas were the rapidly changing health care environment and development and planning of critical transportation infrastructure. As a single session in a short-duration workshop, the presentations in this session only scratch the surface of the applications of the American Community Survey (ACS) data in these topic areas; in particular, the workshop treatment necessarily understates the great attention that has been given to the utility of ACS products in the transportation arena, in which the survey's information on journeys between work and home is vital to infrastructure planning.[1] But the presentations in the session combine some specific applications of the data with some more general "framework" discussions outlining the analytic process in which the data may be brought to bear to solve important problems.

The workshop session included five presentations, three in the health care

[1]See, for instance, a National Cooperative Highway Research Program (2011) report examining the technical issues in producing ACS-based products for detailed transportation analysis that still conform to data disclosure rules. The National Cooperative Highway Research Program (2007) previously issued a technical manual for transportation planners—prior to the release of multiyear data products but anticipating their use and appropriateness to replace the census long-form sample. The most recent revision of *Commuting in America* (Pisarski, 2006) relied principally on the 2000 census long-form data but also included some tabulations from the first waves of large-scale ACS deployment.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

planning arena and two involving transportation. Section 2–A describes the work of a policy resource center established to craft analyses and data products and provide technical assistance—using the ACS as a primary source—for health care decision makers, while Section 2–B examines the uses of the ACS in the public health department of the nation's largest city, including linkages to the city's own survey and data resources. Section 2–C steps back and describes the framework through which data-driven analysis can influence the siting of specific health care facilities or modification of services. With Section 2–D, the chapter switches to the transportation area, beginning with an overview of the ways in which the ACS is used by metropolitan planning organizations to model future transportation trends and infrastructure needs. Section 2–E closes by describing specific legal and regulatory requirements under which the ACS is used to document transportation agencies' compliance with social equity laws. (This specific example foreshadows some applications described more fully in Chapter 7.) The session included brief time for questions, the answers to some of which (clarifying an individual speaker's point) are folded into the earlier questions; discussion of broader questions asked of multiple speakers is summarized in Section 2–F.

2–A PROVIDING DATA AND ANALYSIS TO STATE HEALTH CARE DECISION MAKERS

Kathleen Thiede Call (School of Public Health, University of Minnesota) described the functions of the State Health Access Data Assistance Center (SHADAC), a health policy resource center for which she serves as an investigator. Funded primarily by the Robert Wood Johnson Foundation and housed at the University of Minnesota, SHADAC began operations in 2000 with the goal of making health-related data more accessible to state policy makers. (Additional detail on SHADAC's early history is summarized by State Health Access Data Assistance Center, 2007.) To this end, SHADAC supplies technical assistance to state government agencies to either analyze existing data resources or, in some cases, to collect their own data.

SHADAC projects typically involve assessing health care coverage—both access to health care services and health insurance coverage. Call suggested that states need good data on health care coverage because policy decisions concerning health care have become major (if not dominant) in state-level budgeting. Consequently, the requirements for health insurance coverage data and estimates are considerable:

- Estimates need to be valid and consistent, and need to facilitate comparisons across states;
- Estimates need to support analysis of trends and patterns over time, in part to be able to judge the effectiveness of new policies;

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

- Estimates need to support disaggregation into fine subpopulations—demographic splits by race, ethnicity, age, and poverty status, along with geographic splits by county (or at least some substate areas); and
- Access to microdata, in order to achieve this fine-grained analysis, is critical.

Call suggested that state policy makers are most interested in data on the characteristics of the uninsured—what they look like demographically and where they may be concentrated geographically. Uninsured children are of particular interest: how many children in each county are eligible for Children's Medicaid (CHIP) or State Children's Health Insurance Program (SCHIP) assistance? And, though questions of eligibility for Medicaid have been of interest for years, interest has certainly been heightened among states looking at the effects of the federal Patient Protection and Affordable Care Act (PPACA).[2] Call noted that these kinds of analyses have been done for years and that the states have relied heavily (or exclusively) on the federal government as a source of information. However, budget pressures and constraints are particularly acute at the state level, increasing the demand for reliable information about uninsurance and public program eligibility (and the effects of policy changes on that eligibility).

For their analyses, Call noted that SHADAC can draw from a variety of federal survey data sources, including three specialized health surveys: The National Health Interview Survey (NHIS) and the Behavioral Risk Factor Surveillance System (BRFSS; both conducted by the National Center for Health Statistics[3] [NCHS]) and the Household Component of the Medical Expenditure Panel Survey (sponsored by NCHS and the Agency for Healthcare Research and Quality). However, the requirements listed above are such that the principal sources for analysis are the ACS and the Current Population Survey (CPS), both surveys conducted by the Census Bureau (with the CPS sponsored jointly by the Census Bureau and the Bureau of Labor Statistics).

Prior to 2008, SHADAC relied extensively on the CPS—and the CPS retains some solid advantages. Chief among these are the consistency of the CPS data: the CPS's longer-term inclusion of relevant questions permits trends to be analyzed back to 1986, and its data releases are generally very timely. CPS control variables are such that the data are also amenable to limited disaggregation to substate levels. But there are also major drawbacks, chief among these the relatively small sample size (and corresponding sample sizes for substate pop-

[2]The Patient Protection and Affordable Care Act, P.L. 111-148, was signed into law on March 23, 2010. At the time of the workshop, key provisions of the law were under review by the U.S. Supreme Court. On June 28, 2012, the Court ruled in *National Federation of Independent Business v. Sebelius* that the core mandates under the act were constitutional as a valid exercise of the power of Congress to impose taxes.

[3]The Census Bureau is the data collection agent for the National Health Interview Survey (as well as the ACS), though the survey is sponsored and organized by the National Center for Health Statistics.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

ulations). The CPS questions of key pertinence to SHADAC are asked in the Annual Social and Economic Supplement (ASEC) portion of the CPS, which can suffer from nonresponse; Call suggested that ASEC responses have to be imputed in their entirety for roughly 10 percent of respondents each year. A thornier problem with the CPS is that the form of the key questions is intended to yield calendar-year estimates of health insurance coverage, but not necessarily contemporaneous estimates. A facsimile of the questions on the 2011 ASEC[4] indicates that "these next questions are about health insurance coverage during the calendar year 2010. The questions apply to ALL persons of ALL ages." The lead question is then: "At any time in 2010, (was/were) (you/anyone in this household) covered by a health insurance plan provided through (their/your) current or former employer or union?" Hence, the question is not quite as precise as a measure of *current* coverage and does not capture lapses in coverage.

What changed in 2008 was the addition of a health insurance question to the ACS, and that has had tremendous benefit for analyzing health coverage. A major benefit of the ACS question is asking about coverage at the time of the survey, in contrast to a calendar year reference period and long look-back requirement for coverage in the CPS questions. The ACS version of the question—"Is this person CURRENTLY covered by any of the following types of health insurance or health coverage plans?"—emphasizes current coverage; it permits yes/no answers to seven types of insurance coverage, plus a write-in category.[5] However, there is a new interpretation challenge presented by ACS estimates—explaining, for instance, what an estimate of *current* health insurance coverage means in an average computed over 1, 3, or 5 years of data. Call suggested another challenge inherent in the ACS data, stemming from its development as a general survey and not as a dedicated barometer of health and health insurance trends. Specifically, the unit embodied in each ACS questionnaire—a household, or a "census family" unit—is not necessarily the same thing as a health insurance unit. By its nature, the ACS does not probe to identify relationships within the household/family that would allow access to an individual's health plan (a policy holder and their dependents), and so that relationship cannot be directly recovered.

But, Call argued, the drawbacks of the ACS for examining health coverage are outweighed by the most profound benefit of the ACS relative to the CPS: its larger sample size, roughly 15 times that of the CPS in a given year, and its representativeness for smaller geographic and demographic units within states. Combined with the full range of covariate information available in the ACS, the larger sample size of the ACS has enabled analysis at finer, substate levels that

[4]See http://www.census.gov/apsd/techdoc/cps/cpsmar11.pdf [July 2012], pp. D-77–D-78. The questions are asked through computer-assisted interviewing, hence the syntax choices in the phrasing of the question; the question shown on the CPS interviewer's screen reflects previously collected information.

[5]This question is numbered Person Question 16 in the 2012 version of the questionnaire.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

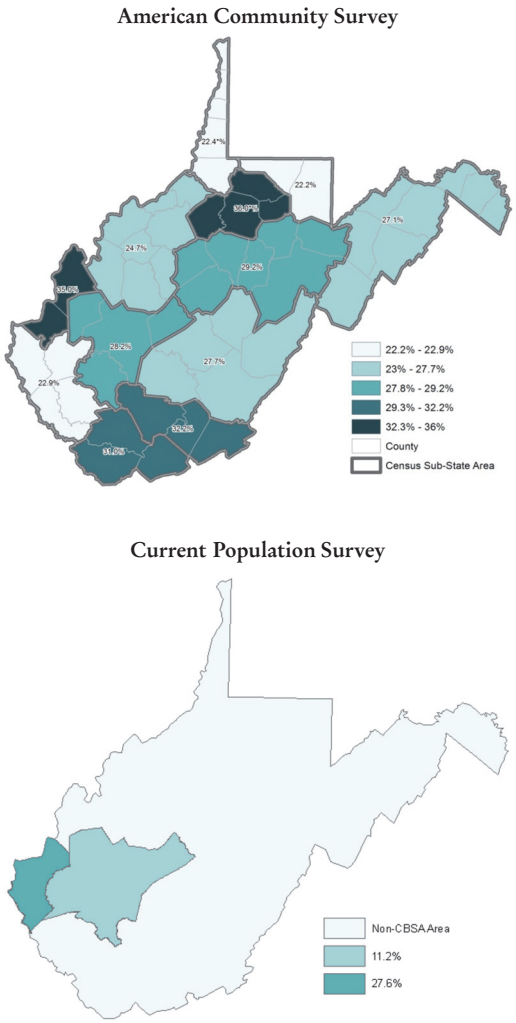


Figure 2-1 Estimated percent of uninsured persons, age 0–64 and at or below 200 percent of the poverty level, in West Virginia, derived from the American Community Survey and the Current Population Survey

SOURCE: Workshop presentation by Kathleen Thiede Call, based on data from 2010 American Community Survey and 2011 Current Population Survey (which covers calendar year 2010).

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

was previously out of bounds. Call presented several examples of data-based maps using the ACS to inform health policy questions, but one that particularly drew a point of comparison with the CPS is shown in Figure 2-1. Asking the general question "where should we allocate funds for community clinics, serving the uninsured?," Call showed the type of analysis that SHADAC conducted for officials in the state of West Virginia, showing the percent of persons in the vulnerable population of interest: low income (defined as being at or under 200 percent of the federal poverty level), nonelderly, and uninsured. Trying to answer the question using 2011 CPS data (which covers calendar year 2010), Call said that the best estimates they could generate were for three regions of the state, based on core-based statistical areas—one of which (the Huntington-Ashland metropolitan area) includes areas of Kentucky and Ohio and so is not specific to West Virginia. Call said that SHADAC would not feel confident even giving the CPS-based analysis over to the state officials. By comparison, the ACS data for 2010 permit good estimates for 12 regions—the Public Use Microdata Areas (PUMAs) within the state, a considerably fuller picture of need within the state.

Call presented and discussed two further examples of ACS-based analysis, illustrating even finer detail and one intended to capture more macro-level trends. In one, 3-year estimates from the ACS were used to profile children not included in Colorado's CHIP program at the county level. The same map was generated at the county level for children not covered under Medicaid, and Call said that the work helped propel policy debates about expanding insurance coverage for children. Call also presented a basic state-level map, shaded to indicate the percentage of persons who would be eligible for Medicaid under the PPACA.

With respect to access to the data—both by SHADAC and its clients—Call said that she had encountered a range of responses among state-level users concerning the Census Bureau's American FactFinder interface. Some have loved the interface and appreciate that one does not have to be a computer programmer to use it; others find it overly cumbersome and not user friendly. The default tabulations available in American FactFinder reflect federal poverty thresholds but not (directly) the federal poverty *guidelines* promulgated by the U.S. Department of Health and Human Services.[6] The FactFinder tabulations also do not provide direct results at the poverty cuts of particular policy interest—e.g., at or below 138 percent or 200 percent of the federal poverty threshold.[7]

In part to compensate for deficiencies in American FactFinder, SHADAC established its own online data center 2 years ago. The data center site—accessible at http://www.shadac.org/datacenter—acts as a table and chart generator using both ACS and CPS data, enhanced to describe health insurance units (as

[6]See http://aspe.hhs.gov/poverty/12poverty.shtml for further description of the distinction between the thresholds and the guidelines, and their 2012 versions.

[7]As of September 2012, the American FactFinder interface was modified to directly add a 138-percent-of-poverty-threshold tabulation.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

distinct from family/household units) and to automate data cuts based on federal poverty guidelines. Like the Census Bureau's American FactFinder interface, SHADAC believes it important to present both estimates and standard errors overtly, so that users can assess differences across population and demographic groups of interest.

Call closed by noting her concerns about the possibility of reductions in effective ACS sample size and generalizability that she said would occur if the survey were made voluntary; she said that such an outcome could greatly impair the representativeness of the data and the states' ability to benchmark and to look directly at some subpopulations of interest within state boundaries. And—in terms of a "wish list"—she argued that the great value of ACS data is its timeliness, yet availability of coverage estimates for half-years or even quarters would be ideal for time-sensitive policy debates. Given SHADAC's health coverage and access focus, Call's "wish list" included addition of two questions: a self-reported indication of general health status and a question on access to health care services. During the discussion period at the end of the session, Call was asked what specific form of question on access might be most useful; she replied that some typical ones from other standalone surveys include questions of the rough forms "Do you have a usual source that you go to for care?" or "At any point in the last year, have you gone without health care because you couldn't afford it?" Questions of this type are part of the NHIS and BRFSS, and could be the model for a more general question on the ACS. On the health status question, even something as basic as "Would you say your health is generally excellent, very good, good, fair, or poor?"—combined with other covariates available in the ACS—could spur important and interesting research.

2–B PUBLIC HEALTH SURVEILLANCE AND ADDING VALUE TO OTHER HEALTH DATA RESOURCES IN NEW YORK CITY

Established in its present form in 2002 through the merger of the existing Departments of Health (itself dating back to 1805) and Mental Hygiene, the New York City Department of Health and Mental Hygiene (DOHMH) is the chief public health agency for the nation's largest city. James Stark, an epidemiologist from the DOHMH Bureau of Epidemiology Services, described DOHMH's use of ACS data in a presentation developed in collaboration with methodology unit director Kevin Konty.

Stark commented that all seven of the "content divisions" within DOHMH[8] use ACS-based analysis in some form, directly or indirectly. The Epidemiology Services bureau is the principal support arm for this analysis, generating population descriptions or deriving population estimates and profiles that are used

[8]The Epidemiology Division, headed by a deputy commissioner, is one of these divisions; Stark's Bureau of Epidemiology Services is housed within that division.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

throughout the department. Besides the Epidemiology Division itself, Stark noted that his comments were based on the ACS uses by the Disease Control and Emergency Response divisions of DOHMH, as well as the office of the Commissioner of Health.

Similar to Call's remarks, Stark said that the most common demand for ACS data is to construct basic demographic profiles. Much more than basic age-and-sex characteristics, Stark said that DOHMH bureaus had needed (and requested) data on very precise groups drawing from many ACS variables, including:

- Household composition for public employees (distinct from private-sector employees);
- Recent immigrants, including language spoken; and
- Enrollment in private school (for comparison with public school data).

One perennial demand—discussed in more detail in the context of implementing the Voting Rights Act in Section 7–A—is to understand the primary languages spoken by New York City residents, broken down by neighborhood. Stark displayed a map derived from 2000 census long-form-sample data showing the range of non-English languages spoken in the city, plotted so that individual dots represent about 150 households that use a particular language. This analysis demonstrates, for example, the wide range of language diversity in the borough of Brooklyn, and the dominance of Spanish as the primary non-English language spoken in Manhattan and the Bronx. At the request of the commissioner's office, and the communications officer in particular, the Epidemiology bureau has replicated this analysis using the most recent 5-year small-area estimates, with the intent of continuing to update the map over time. Stark said that this will help DOHMH produce and provide health-related material for city residents that is both neighborhood- and language-specific.

Given New York City's large population, counts and analysis by neighborhoods or other small areas within the city are of particular interest to DOHMH. However, their analyses also require work with the ACS at higher levels of geographic region as well: counties (the boroughs of the city), the surrounding counties around New York City, and the nation as a whole. Again, given New York City's size, Stark noted that DOHMH is frequently called upon to put analyses within the city in context, through contrast with the rest of the nation; the ACS has proved particularly useful in this regard.

Though ACS-based estimates are interesting in their own right, one primary use of the ACS by DOHMH is to generate denominators, to compute rates based on the department's own health surveys. Chief among these are the New York City Community Health Survey (CHS), a telephone-only cross-sectional survey of approximately 10,000 adults within the city that is administered by the DOHMH Bureau of Epidemiology Services each year. In its administration and content, the CHS is patterned after the Behavioral Risk Factor Surveillance System (BRFSS) conducted by the U.S. Centers for Disease Control and Pre-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

vention (the annual sample size for which is roughly 350,000 adults).[9] The epidemiology bureau uses the ACS to generate control totals for post-stratification weighting to produce CHS estimates; this includes splitting the city by borough and each combination of education status, marital status, and household size. Since 2009, the CHS has included a sample from a list of households having only cell phones, not landline phones; Stark indicated that his bureau has used ACS data—and resulting glimpses in change over time—to try to ensure that this cell-phone-only sample is working properly.

Besides the CHS, Stark said that his bureau also helps other parts of DOHMH link to, and use ACS data as supplement to, a variety of other data collections run by the city. These other data collections include both interview-based surveys as well as registry/records data, such as special surveys of physical activity and use of public transit options; a specialized analogue of the BRFSS aimed at youth rather than adults (age 18 and older); and a periodic New York version of the National Health and Nutrition Examination Survey (NHANES; the federal version is maintained by the National Center for Health Statistics).

In working with public health surveillance data, Stark noted that DOHMH fashions its approach after the Public Health Disparities Geocoding Project headed by Nancy Krieger of the Harvard School of Public Health. Krieger's project was discussed and summarized at an earlier National Research Council workshop; see National Research Council (2009:§ 2–A.1). Recognizing that socioeconomic status can be an important predictor of disease—linked to neighborhood contextual effects that could be associated with disease—the work estimates area-based poverty measures for geographic pockets throughout the city, using the percentage of residents who live below the federal poverty line. For some of their work, DOHMH uses United Hospital Fund (UHF) areas to approximate neighborhoods; these groupings combine multiple ZIP Code tabulation areas to create (for New York City) about 40 districts that are finer than whole boroughs but larger than individual ZIP Codes. DOHMH anticipates updating these analyses using ACS data, with interest in comparison with similarly defined areas/districts throughout the nation; technically, one question with which they are grappling is the appropriate variety of ACS estimates to use (3- or 5-year estimates).

Stark also outlined an epidemiological study of Legionnaire's Disease, a form of severe pneumonia that is believed to be transmitted through contaminated water, in which ACS data on occupation proved very useful. DOHMH's Bureau of Communicable Diseases sought to use its surveillance data on reported cases to examine the hypothesis that occupations potentially associated with contaminated water (e.g., plumbing or air cooling system repair) may re-

[9]Additional information about the CHS is available from the DOHMH website at http://www.nyc.gov/html/doh/html/survey/survey.shtml [July 2012], while the federal BRFSS is described fully in links from http://www.cdc.gov/brfss/index.htm [July 2012].

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

sult in an elevated risk of Legionnaire's Disease. Public health officials followed up on every reported case in the city within a 10-year period, classifying respondents by the occupational categories defined in the ACS. These results were then used to calculate relative risks of incidence of Legionnaire's Disease, suggesting, for instance, that categories like cleaning and janitorial and machinery have higher rates of incidence than others (e.g., office or health care).[10]

DOHMH couples its direct estimation of disease prevalence—based on surveillance of reported cases—with model-based estimates of disease risk. The CDC's NHANES includes numerous laboratory samples for specific pathogens, and the array of population covariates available in the ACS permits models to be constructed using the ACS variables as predictors. Stark briefly discussed work done by a DOHMH colleague, constructing a prediction model for hepatitis C in New York State (and City), down to the county level using the ACS Public Use Microdata Sample (PUMS) data.

Consistent with the themes discussed in Chapter 3, Stark said that DOHMH has also found opportunities for use of the ACS in preparing for and responding to emergencies. They use ACS data to study the spatial distribution of particularly vulnerable populations, for purposes of updating preparedness plans. In particular, DOHMH staff from the Bureau of Emergency Management used the functional disability questions from the ACS (as well as other variables such as poor English literacy) to construct an index of risk. These scores were then weighted based on ACS PUMS data to estimate the population within different degree-of-vulnerability groups within the city. The result of this work is identification of areas of potential need for additional services during an emergency, for not just New York City but the 30-county metropolitan area.

2–C FRAMEWORK FOR ACS- AND DATA-BASED HEALTH CARE PLANNING

Speaking from decades of experience as an independent consultant, instructor, and author—as he put it, mostly in the area of health care and always with a heavy infusion of census data—Rick Thomas (Center for Population Studies, University of Mississippi and Health and Performance Resources, Memphis, Tennessee) framed his presentation as an outline of general approaches to using the ACS in the health care area.

Thomas suggested that the major opportunities, and needs, for use of ACS data in health care planning may fall into seven general (and mainly sequential) categories:

[10]Per Stark's request in his presentation materials not to directly cite the draft results of DOHMH work, this summary is deliberately vague on the details of this example and some of Stark's other specific examples.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

- *Community profiles:* basic documentation of the potential service population in question;
- *Health status assessment:* a more detailed examination (than the community profile) of the population's health characteristics;
- *Health services demand estimates:* an estimate of the need for health services in the population, based on the levels of risk and demand found in the health status assessment;
- *Determination of need:* the practical, technical estimate of potential demand for new services, such as would be needed in a certificate of need application for a new facility;
- *Site selection:* the actual physical siting of a new facility, service location, or deployment of personnel;
- *Business development:* work to grow the client base for services, given the business-like nature of modern health care; and
- *Other health planning uses:* a catch-all category for various other applications.

Actors of various sorts are involved in these types of analyses: individual institutions or businesses (e.g., hospital or medical service groups), state agencies, local governments, and (though not the focus of this workshop) federal agencies.

As a consultant, Thomas said that he looks to three categories when he accesses and uses ACS data. First are the basic demographic and socioeconomic characteristic data that are crucial to building the early profiles and generating the snapshot of the population of interest; in this regard, ACS data are a natural starting point for many analyses. Second are project-related data—more refined queries depending on the application, such as whether the topic is providing services for children, or seniors, or a specific health problem (e.g., chronic disease or reproductive health), and pulling the ACS information that is most relevant to the specific subpopulation. Finally, as noted earlier by Call and Stark, there are a few questions on the ACS that speak directly to health issues—for instance, disability and health insurance coverage, as well as insights that can be drawn related to fertility. These health-specific data in the ACS are few in number but can be very helpful, in Thomas's experience.

This type of analysis is demanding in that it requires data on past, present, and future timeframes. Census and ACS data can provide the past and present views, and serve as the platform for projecting forward into the future. In health care planning, there is usually a desire to profile the population and project needs 5–10 years into the future. Specific to the ACS, Thomas said that he typically uses the 5-year average data products—particularly if it is a smaller target population that is of interest (e.g., a single county) or if it is desirable to drill down to the census tract level. He took care to note that he uses the 5-year averages "with verification"—because the geographic area or demographic

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

group may have been undergoing dramatic changes that might be masked in the wide-window 5-year data, he will look at alternative data sources to triangulate in those cases and to ensure that the 5-year-average glimpses are appropriate.

With that as prelude, Thomas walked through the analysis steps that he would take if presented with a solid query from a client—the (fictional) ABC Health Network, which wants to add a new clinic in an area that it perceives is currently underserved. Though he cast this example as one of physically siting a new facility, he noted that the same process could be used to study the feasibility of adding some new service to existing facilities, to modifying services, or even to eliminating redundant sites or services.

Reiterating and walking through the basic steps, Thomas suggested that he would proceed in the following way:

- *Create a demographic profile:* This would include basic characteristics on the size of the population and how they are distributed (whether the population is concentrated or dispersed). The basic demographic information from the ACS would be most useful here: distributions by age, sex, race, and Hispanic origin.
- *Create a socioeconomic profile:* Specific ACS variables of interest here would be household structure, marital status, educational attainment, employment status, type of occupation, and income level. Consistent with other speakers, Thomas noted that the proportion of the population of interest living in poverty is of critical importance for much of the health care consulting that he does. In some situations, Thomas said that he might incorporate other variables, to get a sense of transportation access, geographic mobility, or primary language spoken at home.
- *Screen population for eligibility:* Federally qualified health centers have to demonstrate certain criteria within their target population in order to qualify for funding from the U.S. Department of Health and Human Services, and similar criteria may apply within states. These kinds of criteria may include proportions of minority population, low-income population, unemployment, and educational attainment—all of which may be derived from the ACS. For many health programs, the number and proportion of single-person households would also be a factor of interest, as would fertility measures.
- *Develop a health profile for the population:* Assuming that the population meets the eligibility criteria, the next step would be to develop a health profile for the area. This, in turn, focuses first on estimating *morbidity*—the prevalence of disease in the population—and projecting the possible *demand for health services*. The problem with both of these concepts is that there are no direct data on them for the whole, general population; hospital admissions and other data give some insight to reported cases, but do not speak directly to the prevalence of underlying disease or risk

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

factors that have not (yet) resulted in treatment. So, in both cases, the quantities of interest must be modeled. Thomas suggested the following basic approach:

1. Obtain age/sex distributions (i.e., age in 5-year intervals) from the ACS;
2. Determine which of four (geographic) regions covers the population of interest;
3. Apply calculated rates of either morbidity or health service utilization, by age, sex, and region, from data resources (derived from actual incidents) maintained by the National Center for Health Statistics (NCHS) to the target area's population;[11] and
4. Compute the number of disease cases or utilization estimates, for relevant health conditions (in the case of morbidity) or for a variety of services (in the case of estimating potential service demand; for instance, estimating number of general physician visits, number of potential well-child exams, or number of surgical follow-up episodes).

Although this is a modeling approach—and one predicated on a very crude geographic measure, the census regions used by NCHS—Thomas reported that he is constantly surprised at (and finds it "scary") how accurate the model can be in estimating the number of cases. For this type of consulting, he said that the model can consider hundreds of different conditions, but that he focuses on the most frequent (say, the top 50) and those that are most relevant for the program he is working with.

From his time using the data, Thomas concludes that the ACS data are very useful to health care planning analyses. Indeed, in many cases, they are essential; there are not very many alternatives for some of the data items, and the ACS is able to provide information at a good and adequate level of detail. In the health care arena, Thomas said, people are used to working with health outcome data that are at least 2–5 years old, so the relative timeliness of the ACS data are a considerable benefit. It is also helpful that, for a lot of purposes, the ACS has moved easily into the position of being considered the "standard"—and so is acceptable to and trusted by the federal grant-making agencies such as the Health Resources and Services Administration or other Health and Human Services agencies.

In terms of accessibility and usability of the data, Thomas said that the data are easily accessible "now that I understand what I am doing, but it didn't start

[11]The National Health Interview Survey, administered by NCHS, is designed to be representative at the national level and for four broad census regions: West, Midwest, South, and Northeast. Other NCHS data such as the National Ambulatory Medical Care Survey, the new National Hospital Care Survey, and its predecessor National Hospital Discharge Survey are coded to the same breakdown of four census regions; these surveys would be used to estimate utilization levels, as discussed below.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

out that way." But once one gets used to the routines of deriving data from the Census Bureau's American FactFinder interface or other files, it is fairly straightforward. Particularly at the tract level, there are sometimes uncomfortably large margins of error to deal with, but Thomas sees those as "something that we have to live with for now."

In terms of suggested improvements to the ACS, Thomas echoed the main ones suggested by Call—the "obvious," ideal suggestions to increase the sample size and speed up the turnaround time (e.g., making the estimates even more timely). He also suggested that this point in the ACS's history may be an opportune one to reassess the ACS questions, based on user input—adding some and deleting others. On the much broader horizon—looking 20–30 years ahead and envisioning a continuous and ongoing need for more and better health data—Thomas suggested that there needs to be continued assessment of the need for data for a variety of uses, and health care planning is a particularly important one. As much money and resources as the health care system consumes and creates, the data needed for effective planning are not always there. The legal reporting requirements under the PPACA of not-for-profit hospitals are one example: Thomas suggested that on the order of 3,000 hospitals are now going to have to do community health need assessments every 3 years, under the terms of the new law, and most of those facilities had not been in the routine habit of doing those assessments in the past. Similarly, there is money in the PPACA for expanding the federally qualified health clinics, each of which is going to need data and analysis of the type outlined in this presentation—a very big pool of applications to keep tabs on. Thomas further suggested that health care providers are increasingly going to have to focus on marketing their services—putting still further demands on the data stream.

As a final note, Thomas suggested that an important path for the future—using the ACS and other sources—will be meeting the need to understand the health care consumer. There are already many health insurance companies and—under the PPACA's health care exchanges—there will now be tens of millions of individuals potentially choosing a health care insurance carrier. Accordingly, the insurance providers are going to be increasingly interested in profiling their potential customers, and developing strategies to identify and engage them.

2-D REGIONAL TRANSPORTATION MODELING IN GREATER SAN DIEGO

Switching from applications in health care to uses of ACS data in the similarly high-stakes area of planning for transportation services, Beth Jarosz (senior demographer) described uses of the ACS in transportation modeling by the San Diego Association of Governments (SANDAG). Comprised of the 18 incorporated cities and towns in San Diego County as well as the San Diego

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

County government, SANDAG is the metropolitan planning organization and council of governments for the San Diego, California, region. In addition to the regular members of SANDAG's board, advisory members from other entities in and around San Diego also participate in SANDAG activities; these include neighboring governments (Imperial County, California, and the Mexican government), transportation authorities (Caltrans, the San Diego Metropolitan Transit System, and the North County Transit District), and the U.S. Department of Defense (given the large presence of Naval Base San Diego and other military facilities in the region). In addition to serving as the San Diego region's designated census data center, SANDAG focuses on issues of major regional impact such as air quality/environmental planning, housing development—and transportation planning.

Jarosz described region-level transportation planning as a multistep process, beginning with identifying stakeholders and assessing their needs and goals. Her particular focus, transportation modeling, is the basis of the next few steps: developing a set of alternatives from available data and information, using them to predict outcomes, and testing and evaluating them in order to select a final alternative. The analytic work in transportation modeling helps with the next immediate step—budgeting for the work—and then implementing. Planning agencies then have the responsibility to evaluate whether the chosen alternative has done what it was intended and expected to, and amend the plans as necessary. In theory, she noted, the final step is achievement of a finished, "perfect" transportation system; in practice, as she noted, anyone who has ever been stuck in traffic knows that the perfect transportation system is a constantly moving target. Consequently, planning agencies like SANDAG repeat this process every several years (often a 4-year cycle), trying to map out transportation infrastructure needs 20–40 years into the future.

Focusing on the transportation modeling steps in this general process, Jarosz presented a simplified outline of the transportation modeling process, with specific reference to the points where ACS data and products enter the mix; this general structure is shown in Figure 2-2. Similar to the general health care planning process described by Thomas in Section 2–C, the first step is preparation of demographic and economic profiles for the area and populations of interest, and then trying to forecast future trends in the population. In transportation, there is interest in the population, housing, and jobs in the region in the next few decades, and the ACS is an important source of those data. Particular variables of interest include household structure (headship), group quarters characteristics, school enrollment, and housing structure type preferences (single versus multifamily or mobile home). In a region like San Diego, with its military presence, information and forecasts on the active-duty military personnel (and their dependents) are particularly important. These demographic and economic characteristics—the data that enter the forecast models—are obtainable from the ACS summary tables or from analysis of the ACS PUMS files.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

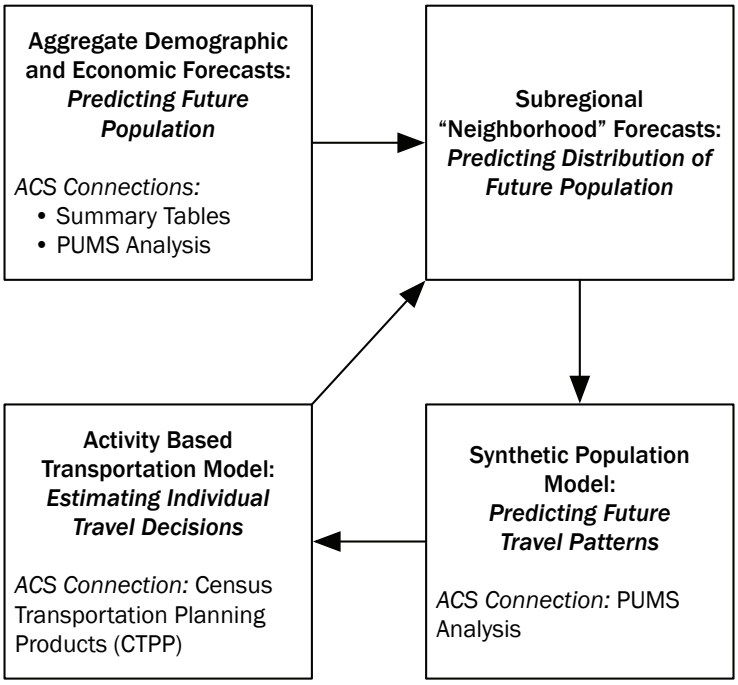


Figure 2-2 Simplified outline of transportation modeling process, showing connections to American Community Survey products

SOURCE: Adapted from workshop presentation by Beth Jarosz.

These underlying demographic and economic forecasts of the basic composition of the population are further scrutinized in what Jarosz described as subregional or "neighborhood"-level forecasts—drilling into finer geographic detail than the region as a whole, and attempting to predict how the population and characteristics will be distributed spatially over time.

The next step is application of a synthetic population model—the basic idea of which is to simulate actions by all actors in the transportation network and model how they will travel through the system. This work is done in tandem with a separate activity-based transportation model (also shown in Figure 2-2) that makes forecasts to simulate individual trips within the system (e.g., when commuters leave for work in the morning, whether they make stops en route to work or home, and when they arrive at work or home).[12] For the synthetic pop-

[12]In the case study/agenda book for the workshop, Guy Rousseau describes the synthetic pop-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

ulation model, ACS variables such as automobile ownership, family structure (presence or absence of children, which would affect the odds of one or more trips to school in a day), work status (and number of workers in the household), and income (as a correlate of the mode of transportation a person/household might select) can play important roles. The synthetic population model typically draws from analysis of ACS PUMS data, while the activity based transportation model draws from special tabulations of data that the Census Bureau distributes as part of the Census Transportation Planning Package (CTPP).[13] CTPP tabulations are coded to the special level of geography typically used by transportation planners: Traffic Analysis Zones (TAZs), collections of census blocks that are defined by the Census Bureau in partnership with local transportation officials and that may be finer grained than census tracts or block groups.

Completing the explanation of the basic outline shown in Figure 2-2, the activity-based model and population synthesizer are part of a feedback loop to the subregional, "neighborhood" forecast. This is because transportation activities can affect the geographic distribution of the future population: Increased traffic congestion in one place might make a certain neighborhood less attractive for future development or, conversely, it might flag an area where the transportation infrastructure must be built up (making the neighborhood more attractive in the long term).

Jarosz stressed that the ACS is essential to transportation planning for the simple reason that it is the only source of small-area trip data, through the detailed CTPP custom tabulation. Through the CTPP, the ACS is the only systematic source of data on flows—giving an indication of where commutes begin and where they end, so that planners can predict how commuters travel between the two. The question of when respondents leave their homes in the morning is sometimes challenged by critics as invasive; Jarosz observed that collection of this information raises privacy concerns, but that both the Census Bureau and the downstream data users are deeply cognizant of those concerns. Like other ACS products, the custom CTPP tabulation is subject to review by the Bureau's Disclosure Review Board and complies with the privacy protection requirements in Title 13 of the U.S. Code; it is also subject to statistical techniques to curb the disclosure of personally identifiable information.

---

ulation model and activity-based model used by the Atlanta Regional Commission in more detail, speaking particularly about the challenges of converting their existing models from census long-form sample inputs to ACS inputs.

[13] As Jarosz noted in her talk, the CTPP data tabulation is funded by pooled funds provided by state transportation departments and metropolitan planning organizations around the country; it was originally compiled from the long-form sample, and is now being converted to the ACS. Though CTPP is often used as shorthand specifically for the data product, the program itself includes training, technical assistance, and research for the transportation community. At the time of the workshop, CTPP tabulations based on 3-year ACS data were available, with 5-year data scheduled for release later in summer 2012.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Because the CTPP, and the ACS, is the lifeblood of transportation planning and modeling, Jarosz said that the transportation planning community is particularly sensitive to the potential drop in response that might accompany a voluntary ACS. From the travel modeling standpoint, the calculus is stark: Smaller sample sizes (as could occur under a voluntary ACS) would make the data unreliable, particularly at the relatively fine TAZ level of geographic aggregation. Smaller sample size would necessitate more data suppression (to protect privacy) and unreliable results. It is as simple as "nothing in, nothing out"—even if there existed a perfect model for predicting travel flows and modes, using unreliable data in a perfect model would still produce undesirable results. She said that these models, and these data, are being used to plan billions of dollars in transportation infrastructure nationwide, which argues for obtaining the best data available to spend those funds wisely. (In the discussion following the presentations, Jarosz noted her strong approval with a comment that a voluntary ACS might compromise the representativeness of the ACS sample, and that this was at least as harmful as a straight reduction in sample size that might result from a switch to voluntary methods.)

Through other work for SANDAG, Jarosz said that data users have come to expect the level of small-area detail that the ACS has been able to provide. The basic socioeconomic variables on the ACS—income and poverty, race and Hispanic origin, and age—are strengths of the data; the ACS questions that permit estimation of disability status have also been valuable for regional planning purposes. Foreshadowing a theme that would be addressed in more detail in the next presentation, Jarosz said that data users have come to expect quality small-area data on questions like the primary language spoken at home; for instance, she has been asked by transit authorities to detail the languages spoken within a half-mile radius of a particular transit stop because they need to produce documents and signage for people who might be affected by a service change. Potentially smaller sample sizes would make it harder to identify very tiny language "clusters" and address community needs. High-quality analysis and planning depends on high-quality data as the input, and that would argue for (if anything) an expansion of the sample rather than a contraction.

2-E MEETING LANGUAGE IMPLEMENTATION REQUIREMENTS IN PUBLIC TRANSIT IN HARRIS COUNTY, TEXAS

In her presentation, Jarosz briefly mentioned the important use of ACS data in establishing compliance with federal, state, and local law and guidelines on environmental justice, social equity, and public access to services. This theme was carried forward by Vincent Sanders, lead transportation systems planner for the Metropolitan Transit Authority of Harris County, Texas (hereafter, METRO),

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Box 2-1 Legal Mandates: Fairness and Access to Federally Funded Services

Title VI of the Civil Rights Act of 1964 (codified as 48 USC § 2000d):

> No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance.

Executive Order 12898 (59 FR 5517; February 7, 1994):

> To the greatest extent practicable and permitted by law, . . . each Federal agency shall make achieving environmental justice part of its mission by identifying and addressing, as appropriate, disproportionately high and adverse human health or environmental effects of its programs, policies, and activities on minority populations and low-income populations in the United States [and its territories and possessions].

who described his agency's moves to use the ACS to certify compliance with foreign language assistance regulations.

As stated in Box 2-1, the enactment of Title VI of the Civil Rights Act of 1964 codifies the principle that government programs must provide equivalent benefits to all segments of the population; Executive Order 12898 in 1994 created similar language for environmental justice, dictating that federal agencies must avoid creating disparate negative impacts (in terms of health or environmental effects) among low-income or minority communities. As part of the enforcement of these provisions, federal grantmaking agencies may impose reporting requirements on their state and local agency recipients, documenting their compliance with antidiscrimination rules; Jarosz and Sanders noted in their presentations that the ACS is gaining increased use in providing quantitative evidence of compliance. The role of the ACS in the allocation of billions of dollars of federal funds is well known, but Jarosz and Sanders suggested that the ACS is an important part of evaluating how those allocated funds are spent.

In his presentation, Sanders described METRO's work in adapting ACS data for demonstrating compliance with foreign language assistance requirements. In 2007, the Federal Transit Administration (FTA)—an important funding source for transit agencies like METRO—slightly adjusted its requirements for documenting Title VI compliance. Specifically, as summarized in Box 2-2, the new guidance to FTA fund recipients called for creation and maintenance of a language implementation plan for the relevant service population's limited English proficiency (LEP) constituency. The language specifically called for these plans to include a strong quantitative assessment of the proportion of LEP persons likely to be served by the recipient program and estimation of the frequency with which LEP persons "come into contact with the program" in any form (e.g., signage, printed materials like timetables, promotional materials). Plans

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

are also meant to include some qualitative components, including an assessment of the "importance of the program [or service] to people's lives."

Obtaining this information is difficult because METRO is a complex intermodal transportation system serving a growing and diverse population. One of the largest transit agencies in the country, METRO covers almost 1,300 square miles in its service area with bus (about 1,250 buses on 130 routes), light rail (a 7.5-mile network), and high occupancy vehicle/carpool service.[14] In 2011, METRO logged almost 77 million boardings on its bus and rail components—alone, a massive number of points of contact with its service population—and its light rail boarding ratio (per mile of track) is second only to Boston in the nation. From a regional planning perspective, Sanders said that METRO is looking at expanding its commuter rail service and investigating possibilities of high-speed capacity in the region—all of which, he agreed with Jarosz, will use the ACS as the basis for modeling and planning.

Sanders commented that the revised requirements came into sharp focus for METRO in 2009, when schedules were such that the agency was subject not only to a triennial review (including the revised Title VI compliance provisions), but also a general FTA audit. In 2009, prior to the rollout of the full set of ACS products, METRO made the natural first step of turning to the most comprehensive extant data concerning the LEP population, which was the 2000 census long-form sample data. But the combined triennial review and audit processes had the salutary benefit of providing the agency with immediate feedback for revision: FTA appreciated the analysis using 2000 census data but noted that those data were effectively obsolete for a growing and changing region like Houston. As context in his presentation, Sanders displayed a tabulation from the first set of 5-year ACS numbers for Harris County, showing the top five languages (other than English) spoken in the county; Spanish dominates the other non-English languages, but the data suggest that the county has become home to significant linguistically isolated Asian communities (Vietnamese, Chinese, and Tagalog included in the top five).[15]

Hence, Sanders and METRO began casting for alternatives—the first of which was to follow up on a suggestion in FTA guidance to obtain data on LEP services from area school districts (in this case, prior to the rollout of the full set of ACS products in 2010). With some scrambling, Houston METRO was able to obtain information on the concentration of LEP students from the Houston Independent School District (the region's largest) and five smaller school districts serving part of the METRO service population. The school district data permitted Houston to create a map of its transit hubs and network structure,

[14] As implied by the agency's full name, METRO's service population is principally in Harris County; however, Sanders noted that the agency's service population spreads over slightly into neighboring Fort Bend, Montgomery, and Waller Counties.

[15] The fifth language group in Harris County, roughly as large as the Tagalog-speaking population, is French.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Box 2-2 Legal Mandates: Language Implementation for Limited English Proficient Users of Transportation Systems

A May 2007 Federal Transit Administration (FTA) Circular for recipients of FTA financial assistance (FTA C 4702.1A) updated previous guidelines issued in 1988. The new circular reinforced a triennial Title VI compliance reporting process for most FTA recipients (except for some metropolitan planning organizations, which are on a 4-year reporting process; p. II-2). Specifically, FTA recipients are asked to document their progress on "tak[ing] responsible steps to ensure meaningful access to the benefits, services, information, and other important portions of their programs and activities for individuals who are Limited English Proficient (LEP)" (p. IV-1). The suggested vehicle for meeting the "meaningful access" requirement is the development and implementation of a formal language implementation plan.

The core elements of such a language implementation plan are specified in regulations promulgated by the U.S. Department of Transportation (USDOT) and applicable to all USDOT programs, including FTA (beginning at 70 FR 74087 [2005]). Section V of the regulation summarizes the basic factors that must be assessed in order for FTA recipient agencies to "determine the extent of [their] obligation to provide LEP services" (70 FR 74091; list reformatted for emphasis):

> Recipients are required to take reasonable steps to ensure meaningful access to their programs and activities by LEP persons. While designed to be a flexible and fact-dependent standard, the starting point is an individualized assessment that balances the following four factors:
>
> (1) The number or proportion of LEP persons eligible to be served or likely to be encountered by a program, activity, or service of the recipient or grantee;
>
> (2) the frequency with which LEP individuals come in contact with the program;
>
> (3) the nature and importance of the program, activity, or service provided by the recipient to people's lives; and
>
> (4) the resources available to the recipient and costs.
>
> As indicated above, the intent of this policy guidance is to suggest a balance that ensures meaningful access by LEP persons to critical services while not imposing undue burdens on small businesses, small local governments, or small nonprofit organizations.

(Section VII of the regulation goes on to specify elements of the actual implementation plan, e.g., staff training and providing notice of LEP access.)

A separate "handbook for public transportation providers" on implementing the USDOT policies, issued by the FTA Office of Civil Rights (April 2007), describes the "four-factor framework" in greater detail, with specific reference to 2000 census (long-form sample) and ACS tables. The handbook also walks through the steps for downloading language use tables from the then-current version of the Census Bureau's American FactFinder site.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

overlaid on a choropleth map of LEP student concentration, as one indicator of the LEP population. Sanders complemented this map with a similar analysis of (self-reported) LEP households by census block group, derived from the 2009 version of Nielsen Claritas' Census Data Update package—an analysis that was complicated by tight budget resources and the pressing deadline for submissions.

Facing another triennial review and update of METRO's plans in 2012, Sanders said that the ACS "is coming to our rescue." He noted that METRO is now trying to consistently use census tracts as the level of aggregation rather than block groups, but that this level of resolution has been proving useful. In terms of understanding the LEP population, the most recent ACS data suggest that 12.69 percent of the METRO service population is LEP and the data have lent themselves to more refined mapping and planning; they have learned that half of that LEP population (in areas with a LEP percentage higher than a certain threshold) are located within a fairly tight one-quarter-mile catchment area of route service on the METRO network. Working with the ACS data and keeping with the more general Title VI mandate, METRO has begun overlaying its transportation network map layer over ACS-derived maps of racial composition and poverty levels to document how and whether the current network is satisfying user needs.

In the discussion following the presentations, Sanders was asked whether METRO planned to make further use of the school district data on LEP student data. Sanders answered that METRO was generally pleased with the coverage of the data it assembled from school districts; he has done some direct comparison of the school district data with ACS tabulations and noted that the agency's early work with the school district numbers had yielded surprises and directions for future work. However, METRO is sufficiently comfortable with and confident about the ACS as a consistent source of data for areas and groups within its whole service population that it does not plan to repeat the school district collection.

Sanders closed by noting that the Houston region is continuously changing; he said that the area is becoming known for its sociodemographic diversity, and METRO is one of several Houston planning agencies that feel a strong need for the ACS to continually update their plans and projections.

2–F DISCUSSION

In the questions and discussion following the presentations, several speakers were asked to clarify whether they calibrate or check the ACS data on basic demographic groups (race and ethnicity) against the decennial census. Sanders answered that he and METRO use the 2010 decennial census counts for some of their count tabulations, but also use the ACS 2006–2010 estimates with some comparison. Jarosz clarified that she and SANDAG tend to use the ACS data as

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

a source for *rates* or percentages rather than raw counts; for counts, they use the decennial census totals as well as their in-house population estimates program. Thomas agreed, using the ACS for rates rather than counts.

Lester Tsosie (Navajo Nation and a workshop presenter; Section 5–D) asked a similar question, whether the presenters had developed any order of precedence for the decennial census and the 1-, 3-, and 5-year ACS products. Jarosz answered that it depends on the specific question at hand. For benchmarking population counts or counts for age or sex groups, the natural preference is for the decennial census counts. And, while some region-level questions might be answerable with 1- or 3-year ACS numbers, much of her work involved analysis at the census tract level, giving the 5-year products particular prominence. Sanders agreed, reiterating their use of the decennial census for counts but the ACS for rates (e.g., above or below threshold percentages of LEP population for Title VI compliance documentation).

Andrew Beveridge (Queens College and Social Explorer, Inc., and a workshop presenter; Section 7–C) followed this general thread of calibration between the ACS and the decennial census, asking the Census Bureau staff in the room about the shock in the 1- and 3-year estimates caused by the changing population base used to weight the user survey. To construct ACS estimates, the survey data must be weighted based on population controls that—until a new decennial census takes place—are based on the Census Bureau's population estimates program, essentially updating population counts every year between the censuses. For large-population areas covered by the 1- and 3-year ACS data, this meant that ACS products released in 2010 were controlled to 2009-vintage population estimates; the products released the following year could make use of the completed 2010 census counts. In cases where the 2010 census counts were discrepant from the existing population estimates—including, as Beveridge noted, large cities like New York and Detroit—these differences and their effects on the ACS estimates made the ACS products less useful for tracking change over time. Beveridge asked why the Bureau did not go back and revise the earlier estimates, or otherwise work to bridge this shock in the estimates; James Treat (ACS Office Division Chief, Census Bureau) answered that this is certainly an issue the Bureau is aware of, but is basically a resource constraint. Generating the current estimates is intensive enough, with activities and requirements stacked on each other, that the resources do not exist to revisit earlier products.

Scott Boggess (U.S. Census Bureau) challenged the presenters with a question that would recur throughout the workshop. Though several of the speakers had commended the timeliness (and relative timeliness to other information) of ACS releases, he noted that the Census Bureau gets continued feedback saying that it would be nice if the releases came out even faster; Call had included more timely releases, even on a half-year basis, in her "wish list." The question was whether there are specific examples where a difference of weeks or months in release would really affect the usefulness of the data—whether there are specific

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

instances where ACS estimates being available in June would be much better, in some sense, than estimates coming out in September. Jarosz said that the call for increased timeliness reflects the comments and feedback that people like the presenters in this session hear from their downstream clients and users; she said that in the age of information technology, elected officials, the public, and the media all expect "instant gratification" and the most current information. That said, she empathized with the Bureau's concerns about the production cycle as well; indeed, she said that she regularly spends time teaching people about how long it takes to process census and ACS data. She noted her own personal comfort with the current release time, echoing Thomas's comment during his presentation that analysts in health care are used to data being of considerably older vintage (4–5 years old). Stark agreed, noting that some of the other data used by DOHMH use ZIP Code as their sole geographic key—as a result, there are some analyses they do where DOHMH has to wait (sometimes many months) for results to be released at the ZIP Code tabulation area (ZCTA) level.[16] When asked to work on a new public health initiative, DOHMH has to account for the possibility that they might need to switch to a different level of geography—where a quicker analysis might or might not be as good as one using the ZCTA-keyed data.

[16] As indicators of postal delivery areas, ZIP Codes do not necessarily form a coherent geographic layer; they may overlap or may technically relate to a specific geographic point rather than an area, as in a ZIP Code for a place without home mail delivery (all P.O. Box delivery). Hence, the Census Bureau periodically constructs ZIP Code tabulation areas to approximate the delivery areas of 5-digit ZIP Codes.

Copyright © National Academy of Sciences. All rights reserved.

– 3 –

Planning Social Services and Responding to Disasters

The workshop's second session maintained a focus on particular themes—in this case, the applications of the American Community Survey (ACS) in planning and administering social services (not just the allocation of funds but the way federal and state funds are administered at subnational levels) and in preparing for and responding to disasters. But, by its construction, it also served to focus on a particular sector of ACS users: nonprofit organizations.

Section 3–A summarizes a presentation on the use of the ACS for studying welfare "safety net" policies and its strength relative to other data sources, while Section 3–B outlines the way in which a research organization serves as an "interpreter" of ACS estimates for state and local policy makers. The disaster planning and recovery portion of the session included both a specific example—work to assess the impacts of Hurricanes Katrina and Rita and New Orleans' recovery from those natural disasters (Section 3–C)—and a more general description of the framework for using the ACS and other data for disaster preparedness (Section 3–D). The workshop would revisit the theme of disaster preparedness in one of its featured uses of ACS data by private business; see Section 6–E. (The only questions in the closing minutes of discussion were to clarify individual remarks, so that material is woven into the other sections of the chapter rather than a final standalone section.)

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

3–A CONTRAST WITH THE CURRENT POPULATION SURVEY FOR STUDYING LOW-INCOME "SAFETY NET" POLICIES

Linda Giannarelli (Urban Institute) began her discussion by noting that her remarks could not be exhaustive of the hundreds, if not thousands, of ways that the ACS data have been brought to bear on examining the low-income population and social welfare. Rather, she said that she wanted to focus specifically on examples from work at the Urban Institute and elsewhere. Touched on briefly in Kathleen Thiede Call's discussion (Section 2–A), the comparison of ACS data with those from the Current Population Survey (CPS) was a principal focus of Giannarelli's presentation.

Giannarelli began with a brief overview of the relevant piece of the broader CPS—the module that longtime data users think of as the "March supplement" but which is formally known as the Annual Social and Economic Supplement (ASEC) to the CPS. A detailed battery of questions on income, work experience, family structure, and other topics, the ASEC is administered to roughly 100,000 households per year; this includes the households already in the CPS sample in March of a particular year as well as additional Hispanic households (since 1976) and additional households with children ages 18 or younger (since 2001; this latter category is also known as the "CHIP sample" because it was designed to improve estimates of participation in state Children's Health Insurance Programs) (U.S. Census Bureau, 2006:11-5–11-6).[1] This major supplement to the CPS has particular prominence because its information on income in the preceding year is the government's source of its official poverty statistics; echoing Call's description in the previous session, Giannarelli observed that the CPS ASEC has been "the workhorse of federal surveys regarding population issues" for decades, given its inclusion of questions on family structure and demographic issues as well as (dozens of questions on) types of income. The ASEC sample is designed to be representative of the nation as a whole, of broad census regions, and of individual states. However, Giannarelli added the significant caveat that the sample size of the CPS is not big enough to support analysis of low-income households in most states.

There are sharp trade-offs in the detail and scope of analysis that can be done using the ACS rather than the ASEC. In terms of income data, all types of welfare income are collected in one variable while they are split across different questions, and programs, in the ASEC; with the ASEC, for instance,

[1]The CPS differs from the ACS in that the CPS universe is intended to be the civilian noninstitutional population (completely excluding military personnel), while the ACS includes a sample of persons living in nonhousehold group quarters. The ASEC differs from its parent CPS sample in that it "includes military personnel who live in households with at least one other civilian adult" (U.S. Census Bureau, 2006:11-6). As would be borne out in questions later in the workshop, the CPS (and ASEC) also differ from the ACS in that response to the CPS is voluntary, not required by law.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

one can look specifically at benefits received under the Temporary Assistance for Needy Families (TANF) Program. Likewise, some policy-relevant types of income such as child support, unemployment compensation, and workers compensation are lumped together into one "all other income" variable in the ACS. On employment, the ACS asks only for the duration of a person's employment in ranges of weeks and not a more precise measure;[2] it asks only a yes/no question on whether a person has "been ACTIVELY looking for work" in the previous 4 weeks, and not a question on how many weeks a person has been seeking work.[3] Finally, ACS questions on benefits received—arguably most crucial to studying welfare "safety net" policies—are limited to specific amount breakdowns for "Social Security or Railroad Retirement," "Supplemental Security Income (SSI)," "any public assistance or welfare payments from the state or local welfare offices," and "any other sources of income" such as veterans' benefits or unemployment compensation.[4] Clearly, Giannarelli said, this lack of detail and confounding of different income types is not ideal for safety net analysis—among other things, the ACS provides no direct insight on income received in Supplemental Nutrition Assistance Program (SNAP) benefits (formerly called Food Stamps),[5] Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) benefits, or public housing assistance or vouchers.

That said, Giannarelli echoed Call's bottom-line conclusion: the great advantage of the ACS over the CPS ASEC, and the thing that makes the ACS "irresistible" to researchers, is sheer sample size. To illustrate the point that the ACS supports state and substate analysis that the CPS ASEC simply cannot, Giannarelli displayed the 2009 ASEC and 2008 ACS sample sizes (expressed as number of people, not households) for the states of Georgia, Illinois, Massachusetts, and Wisconsin; she also displayed the numbers of those persons in each state sample with low income (less than 200 percent of the poverty line).[6] For Illinois, for example, ASEC sampled roughly 8,200 people, 2,600 of whom are low income; the ACS's coverage for Illinois was about 123,000 people, just over 30,000 of whom would quality as low income—and, Giannarelli said, there is simply no comparison in the degree to which that low-income population in the ACS sample could be scrutinized without having to either combine years of ASEC data or look only at very high-level geography (state or region).

[2]In the 2012 version of the questionnaire, Person Question 39a asks "During the PAST 12 MONTHS (52 weeks), did this person work 50 or more weeks?" (yes/no) and 39b asks "How many weeks DID this person work, even for a few hours . . .?," with responses being 50–52 weeks, 48–49 weeks, 40–47 weeks, 27–39 weeks, 14–26 weeks, and 13 weeks or less.

[3]This latter question is Person Question 36 on the 2012 version of the ACS questionnaire.

[4]These are Person Questions 47d, e, f, and h on the 2012 ACS questionnaire.

[5]As discussed in Section 5–B, the ACS does include a yes/no indicator of whether anyone in the household received SNAP/Food Stamp benefits during the past year, but respondents are not asked about specific levels of assistance received.

[6]Persons in group quarters were removed from the ACS totals, to match the coverage of the CPS ASEC.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Giannarelli outlined four types of safety net analyses that can be done using ACS data, two of which she described briefly and two of which she sketched out with specific examples. The two mentioned in brief are highly similar to the types of profile analyses described by previous speakers in making use of the wide range of characteristic information available in the ACS. First, demographic profiles (numbers and characteristics) can be constructed for persons and families in poverty—either using the official measure or working through the definitions in the expanded Supplemental Poverty Measure (SPM)—at the state and, often, substate level. The advantage with the ACS is that these profiles can be derived from a single product; in the past, many states would have to combine averages across multiple years of CPS data even to get a sense of state-level poverty. And, second, for those benefit types that are clearly delineated and captured in the ACS (e.g., Social Security or SSI), the ACS permits profiles and characteristics of families or persons receiving those benefits using measures not available in administrative data.

The third type of safety net analysis made possible by the ACS makes use of the additional covariate information in the ACS to compute person and household eligibility for benefits under state-level requirements, and then to compare the eligible population with those who actually receive the benefits. As specific examples—and further contrast between the CPS and ACS as sources—Giannarelli mentioned two examples of work done by the Urban Institute for different clients. First, the U.S. Department of Health and Human Services (HHS) periodically asks the Urban Institute to generate state-level estimates of eligibility for federally funded child care subsidies under the Child Care and Development Fund. In the past, this work has required using 2 years of CPS ASEC data to construct the state-level estimates but—though this work does generate some useful insights—Giannarelli conceded that the resulting standard errors on the estimates are sufficiently large as to make one question the utility of the estimates for the states. For instance, a state-level point estimate of children eligible for these subsidies might be 25,000 but, given the dollars involved, one has to wonder whether the natural next statement—that a 95 percent confidence interval for the number of eligible children suggests that the count is between 15,000 and 35,000—is really useful or informative. By comparison, the Food and Nutrition Service of the U.S. Department of Agriculture contracts with the Urban Institute to estimate state-level eligibility for WIC benefits, and the institute has been doing this work using a combination of CPS and ACS—a combination, in part, because data on WIC benefits are not directly collected by the ACS questionnaire. Simply put, Giannarelli said, there is no way that the Urban Institute could have even attempted to compute state-level eligibility estimates for a program focused on such a precise population (women, infants, and children under age 4) using the CPS ASEC alone.

As further examples of work of this form being done by other organizations, Giannarelli noted New York City's Center for Economic Opportunity (CEO),

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

which is tracking trends in poverty over time, adopting a revised measure of poverty initially suggested by a National Research Council (1995) panel. Their work—including efforts to adapt the New York City methods to estimates of poverty for New York State as a whole—has relied on the ACS. Likewise, she said that the ACS has been used extensively by the Institute for Research on Poverty at the University of Wisconsin–Madison; that group has looked at Wisconsin's eligibility estimates at both the statewide and substate levels.

The New York City and Wisconsin groups have also done work in what Giannarelli called the fourth type of ACS-based safety net analysis—"what if" analysis, to simulate outcomes had a certain policy not been implemented or had a specific mix of eligibility requirements been altered. She specifically cited New York City work making use of the ACS Public Use Microdata Sample (PUMS) data, comparing the city's actual and hypothetical poverty rates (computed under the city's adopted formula) had the changes in SNAP eligibility included in the 2009 American Recovery and Reinvestment Act not been implemented.[7] Absent the 2009 act's changes, the city's study found that the percent of population in poverty might have ticked over the course of 3 years by about 3 percent (New York City Center for Economic Opportunity, 2012).

Wrapping up, Giannarelli described a few safety net analyses conducted by the Urban Institute with the ACS data. In one study, funded by the Annie E. Casey Foundation, the motivating question was how states' current safety net policies affect child and non-elderly adult poverty. She and her Urban Institute colleagues focused on three states—Georgia, Illinois, and Massachusetts—classifying those states' existing policies as narrow, medium, and broad safety nets, respectively. The study calculated state-level estimates using SPM-type poverty measures and made use of 2008 ACS data supplemented by data from the institute's TRIM3 microsimulation model.[8] Key results from that work are evident in the graphs reproduced in Figure 3-1. In the top graph, comparing the absence of safety net policies (the leftmost "no safety net" bars) with the full set of safety net policies (the rightmost "all benefits" bars) suggests that the safety net policies serve to cut child poverty rates by half; the bottom graph shows that two specific safety net programs (TANF and SNAP) both serve to reduce the child poverty rate to varying degrees by state. Giannarelli commented that the Urban Institute could never have conducted this type of analysis using the

[7]The American Recovery and Reinvestment Act, commonly referred to as the 2009 stimulus bill, increased SNAP benefits by 13.6 percent. The hypothetical estimates were based on the normal expansion that would have applied to Food Stamps without the 2009 act; it also made assumptions about the growth rate of the SNAP caseload in the city. See New York City Center for Economic Opportunity (2012:32–34).

[8]Formally the Transfer Income Model, Version 3, TRIM3 simulates tax and health programs, generating estimates at the individual and family levels as well as for geographic entities (state and nation). TRIM3 is developed by the Urban Institute with primary funding from the Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. Additional information on TRIM3 is available at http://trim3.urban.org.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

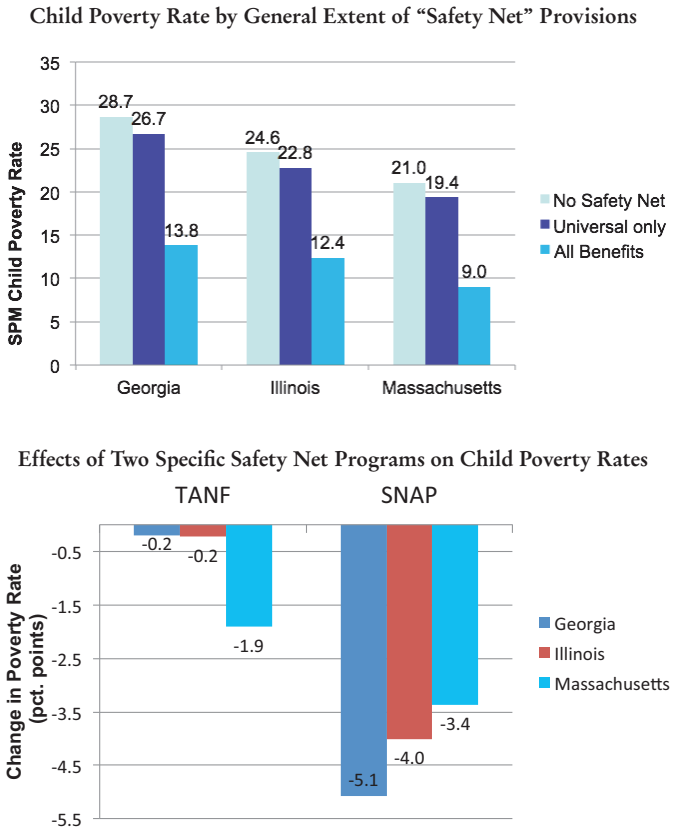


Figure 3-1 Effect of welfare "safety net" provisions on child poverty rate, Georgia, Illinois, and Massachusetts, 2008

NOTES: SNAP, Supplemental Nutrition Assistance Program; SPM, Supplemental Poverty Measure; TANF, Temporary Assistance for Needy Families Program.

SOURCE: Workshop presentation by Linda Giannarelli, based on data from the 2008 American Community Survey.

CPS alone; even if they had combined years of CPS data to get a sufficiently large sample of low-income people in each of the states, the safety net programs change sufficiently from year to year that the results would not be credible.

She added that the Urban Institute had developed a body of work doing "what if" state poverty analysis depending primarily on CPS (in combination with TRIM3); for instance, this type of analysis was done for poverty commissions in Connecticut and Minnesota in 2008–2009. In 2009–2010, the Urban

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Institute used funding from the Annie E. Casey Foundation to adapt TRIM3 to use and work with ACS data for these "what if" studies. To date, the institute has completed work on two projects making use of the new ACS-based methods, studying poverty analyses and the effects of specific packages of policy changes in Illinois (for Heartland Alliance) and Wisconsin (for Community Advocates).

3–B INTERPRETING ACS RESULTS TO INFORM SOCIAL SERVICE PROVIDERS

Established in 1989 as the Heartland Alliance Mid-America Institute on Poverty and bearing its current name since 2010, the Social IMPACT Research Center (IMPACT) is a research and evaluation arm of the nonprofit Heartland Alliance for Human Needs and Human Rights in Chicago.[9] For more than 12 years, IMPACT has been involved in the generation of an annual *Report on Illinois Poverty* (the most current version of which is Social IMPACT Research Center, 2011); as IMPACT's associate director Amy Terpstra noted in her workshop presentation, IMPACT's work focuses on populations or issues impacting populations who are economically vulnerable or experiencing economic hardship. Its primary role is to convey important issues and trends affecting quality of life for low-income individuals to local social service agencies, policy makers, and the general public. In this work, Terpstra suggested that IMPACT and agencies like it play an important role as "interpreters" of ACS data for the broader public—and that, for purposes of informing policy debates, the availability of ACS data has been a "game changer."

Terpstra began her remarks by recalling when she started getting into this work; she said that she was shocked at the degree with which policy decisions related to low-income individuals were made based on gut feelings—or whims and fancies—rather than empirical evidence. IMPACT takes as its goal equipping decision makers with good information, or the best information available, and to make it accessible and easily digestible.

Terpstra structured her presentation around a few major ways in which IMPACT uses ACS data, the first of which is to educate and promote policy change—in particular, policies to help people experiencing economic hardship. IMPACT's annual report on Illinois poverty is a primary resource in this regard. The annual report itself grew out of concerns in the late 1990s when the national economy seemed flush yet Heartland was still seeing people "coming through our doors needing jobs, needing extra support to make ends meet." The annual report grew out of a desire to educate elected officials, primarily, on continuing struggles with poverty. Terpstra said that the structure of the annual report has

[9] All of the studies and products mentioned in Terpstra's presentation are generally available at IMPACT's website, http://www.heartlandalliance.org/research.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

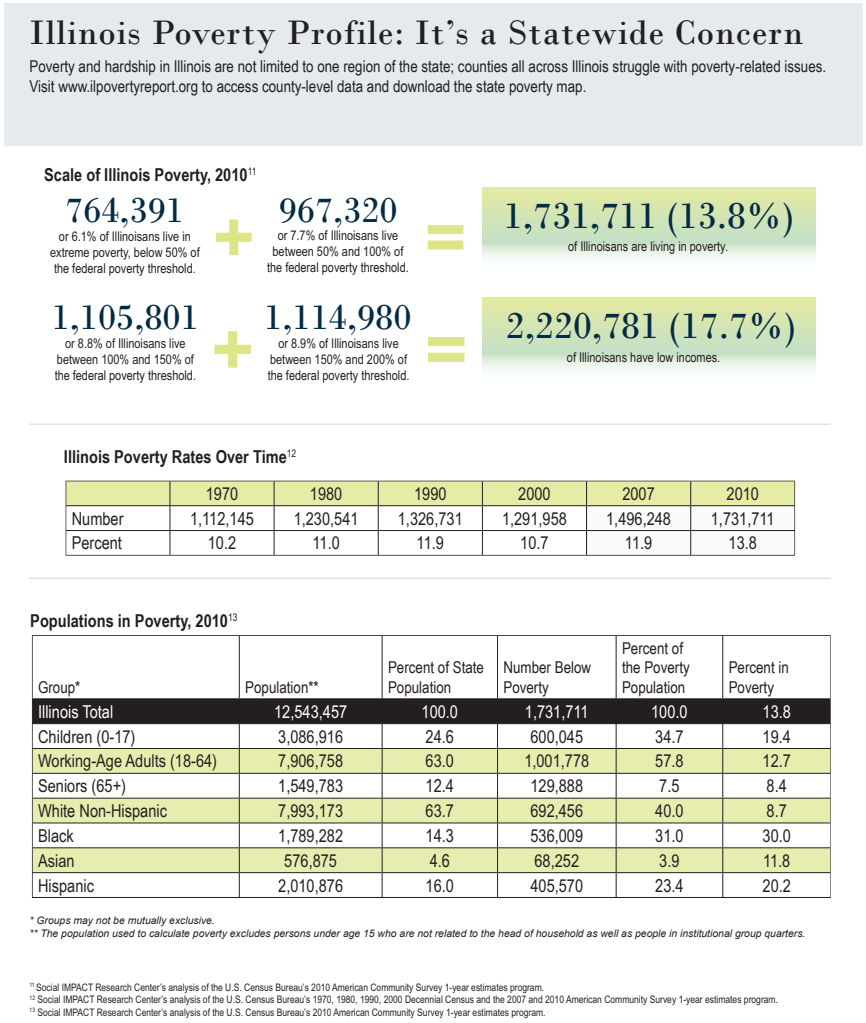
evolved over the years, but that IMPACT still endeavors to make it a very visual data book, using information graphics and graphic design to make the data visually appealing and accessible. She displayed a summary page from the 2011 report (reproduced in Figure 3-2); as she observed, the page is not particularly infographic-laden, but it does cleanly break down high-level figures on poverty in the state. For people well steeped in the data, the figures are nothing new, but for most people—and many decision makers—the concept that there are about 1.6 million people in Illinois who are poor is "mind-blowing." In addition to presenting the basic facts of poverty, IMPACT's annual report also focuses on what it calls "pathways out of poverty"—summarizing statistical indicators of employment, health and nutrition, assets, and housing.

Consistent with Call and Giannarelli's experiences, Terpstra said that IMPACT used to rely heavily on the CPS for the facts and figures in its reports, but has now transitioned to using the ACS. She said that this has provided greater flexibility in analysis and greater confidence in the results. In 2012, with the availability of 5-year ACS numbers, IMPACT revisited the way that it presented information for all 102 counties in Illinois. In the past, the available information on the counties was basically put into an appendix listing in the print report; this year, IMPACT created a web-based portal at http://www.ilpovertyreport.org to allow users to directly access ACS results and other data for the counties in an interactive manner. In addition to screen display, the site provides users with the capacity to download data tables and to access readymade "fact sheets" for easy reference.

Terpstra said that IMPACT and Heartland Alliance are finding that county social service agencies are using this web portal quite extensively. This is encouraging because those local agencies are probably IMPACT's primary user constituency; through IMPACT's analyses, the local agencies are able to better understand needs in their community and decide what information they can take to their elected officials, to inform debates on budgets and priorities. Live since December 2011, the Illinois Poverty Report portal has had on the order of 6,000–7,000 unique users.

The online data and report portal is one increasingly important tool but, for purposes of IMPACT's second principal use of ACS data—documenting local trends in poverty and related phenomena—some more old-fashioned approaches remain effective. Each fall, when new ACS (and CPS) data generally become available, Terpstra said that IMPACT does a very big push to pull in that information and then to generate updated series of custom "fact sheets" for a wide assortment of geographies (counties and cities) in Illinois. She displayed some pages of the standard fact sheet, providing data for Winnebago County in northern Illinois; one of those pages is reprinted in Figure 3-3. Like the annual poverty report, the intent is to present the information in an easy- and attractive-to-read format. IMPACT then disseminates these fact sheets to its list of interested users and agency partners, and also launches a fairly extensive me-

Copyright © National Academy of Sciences. All rights reserved.



Illinois Poverty Profile: It's a Statewide Concern

Poverty and hardship in Illinois are not limited to one region of the state; counties all across Illinois struggle with poverty-related issues. Visit www.ilpovertyreport.org to access county-level data and download the state poverty map.

Scale of Illinois Poverty, 2010[11]

764,391 or 6.1% of Illinoisans live in extreme poverty, below 50% of the federal poverty threshold. + 967,320 or 7.7% of Illinoisans live between 50% and 100% of the federal poverty threshold. = 1,731,711 (13.8%) of Illinoisans are living in poverty.

1,105,801 or 8.8% of Illinoisans live between 100% and 150% of the federal poverty threshold. + 1,114,980 or 8.9% of Illinoisans live between 150% and 200% of the federal poverty threshold. = 2,220,781 (17.7%) of Illinoisans have low incomes.

Illinois Poverty Rates Over Time[12]

	1970	1980	1990	2000	2007	2010
Number	1,112,145	1,230,541	1,326,731	1,291,958	1,496,248	1,731,711
Percent	10.2	11.0	11.9	10.7	11.9	13.8

Populations in Poverty, 2010[13]

Group*	Population**	Percent of State Population	Number Below Poverty	Percent of the Poverty Population	Percent in Poverty
Illinois Total	12,543,457	100.0	1,731,711	100.0	13.8
Children (0-17)	3,086,916	24.6	600,045	34.7	19.4
Working-Age Adults (18-64)	7,906,758	63.0	1,001,778	57.8	12.7
Seniors (65+)	1,549,783	12.4	129,888	7.5	8.4
White Non-Hispanic	7,993,173	63.7	692,456	40.0	8.7
Black	1,789,282	14.3	536,009	31.0	30.0
Asian	576,875	4.6	68,252	3.9	11.8
Hispanic	2,010,876	16.0	405,570	23.4	20.2

** Groups may not be mutually exclusive.*
*** The population used to calculate poverty excludes persons under age 15 who are not related to the head of household as well as people in institutional group quarters.*

[11] Social IMPACT Research Center's analysis of the U.S. Census Bureau's 2010 American Community Survey 1-year estimates program.
[12] Social IMPACT Research Center's analysis of the U.S. Census Bureau's 1970, 1980, 1990, 2000 Decennial Census and the 2007 and 2010 American Community Survey 1-year estimates program.
[13] Social IMPACT Research Center's analysis of the U.S. Census Bureau's 2010 American Community Survey 1-year estimates program.

Figure 3-2 Summary page, Social IMPACT Research Center Illinois Poverty Report, 2011

SOURCE: Extract (p. 4) from Social IMPACT Research Center (2011) (with minor cropping for size), as displayed in workshop presentation by Amy Terpstra.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

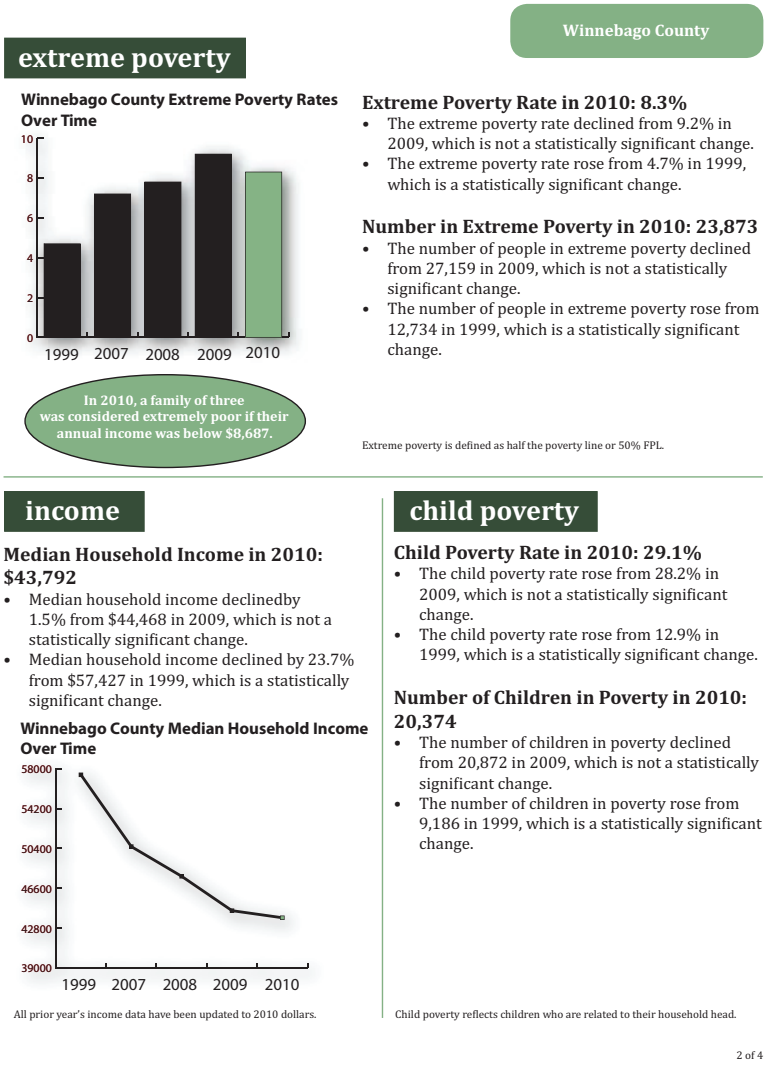


Figure 3-3 Excerpt of poverty, income, and health insurance coverage profile, Winnebago County, Illinois, 2010

SOURCE: Extract (p. 2) from Winnebago County factsheet, downloaded from http://www.scribd.com/doc/65972489/Winnebago-County-Fact-Sheet, as displayed in workshop presentation by Amy Terpstra.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

dia push to local media; this in turn involves drafting custom press releases that local editors can most easily adapt and publish.

On this note, Terpstra said that she wanted to express one soapbox position in the presentation—concerning the Census Bureau's embargo policy on ACS releases. In 2011, the Bureau's policy changed—tightened—so that IMPACT and policy groups like it had their early-access rights (under an embargo period) removed, which severely handicaps their ability to serve as "interpreters" for local media reporters. IMPACT had to switch to doing all (or the bulk) of its analysis and trying to update dozens of fact sheets on the day of release. She argued that the Bureau should revise its embargo policy or implement a secondary policy covering agencies that provide media support services. Otherwise, the tighter embargo policy serves to blunt the media attention that might otherwise accompany the new data. IMPACT has found a niche in serving reporters and media sources without the means to conduct analysis on their own—in many cases, even lacking access to Microsoft Excel or the most basic of spreadsheet programs to manipulate data. Those small media outlets, technically, are the ones that *could* have access to the data under the embargo period but they completely lack the wherewithal to do anything with embargoed data; meanwhile, organizations that could serve to parse the data for those reporters are locked out. (The topic of the ACS embargo policy would naturally arise again in the workshop session on media perspectives; see Chapter 4.)

Another use that IMPACT tries to make of the ACS is to actually *have* an impact on programs and policies concerning people experiencing economic hardship—to drive informed, solutions-based change. The example that Terpstra walked through in this area is work that IMPACT has done for the Greater Chicago Food Depository, one of the numerous social service agencies that make data requests of IMPACT. IMPACT's relationship with the Food Depository goes back many years, and IMPACT has performed many different iterations of the analysis of the need for nutrition support among seniors in Cook County and the city of Chicago. In the work for the Food Depository, the basic objective is to estimate hunger among seniors in small areas in Chicago and then to compare those findings with actual disbursements (of funds from federal nutrition programs and of food from the Depository itself). The most recent iterations of the analysis make use of 5-year estimates from the ACS, building census tracts into the 77 long-established community areas that are commonly used as "neighborhoods" in studies of Chicago. The community-area ACS estimates were then compared to administrative data from the nutrition programs; maps were then drawn up to clearly portray areas where need was greatest (and areas where services were most lacking). The Food Depository has taken this analysis and started to draw up plans for new service in some areas of the city; similar action has been taken based on a similar analysis IMPACT did on children's nutrition programs.

Completing this train of thought, Terpstra quickly displayed a screenshot

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

of an ongoing data book (and data product) that IMPACT continually updates, which aggregates ACS data to the previously mentioned community areas that are well known and understood in the city. Again, she said, IMPACT serves as an interpreter and conduit for the data; most of their downstream data users do not have the capacity to directly manipulate ACS files into the familiar community areas, so IMPACT provides a service by assembling the tract-level ACS data into a more usable format.

Finally, Terpstra summarized two examples of IMPACT's use of the ACS to inform targeted poverty reduction strategies. First, IMPACT does analysis of extreme poverty conditions for the state's Commission on the Elimination of Poverty, beyond the production of the annual poverty report described above. The commission itself was established by the state based on Heartland Alliance's work, noticing an increasing trend in extreme poverty (defined as people with income below 50 percent of the federal poverty level); IMPACT was then engaged to conduct some specialized tables and graphics on this population and its characteristics. The first cut at this work was done early in IMPACT's work with ACS data. If it is repeated, Terpstra suggested that it would likely be more sophisticated given the experience that they have acquired in working with ACS PUMS files. But the work was sufficient to demonstrate three clear subgroups of interest in the extreme poverty population. One segment consists of vulnerable populations who are prone to extreme poverty because they either cannot work or are not expected to work; this segment includes young children, seniors, and persons with disabilities. A second segment consists of people who *are* working in the labor force, but who do not have enough work (e.g., part-time) or simply do not make enough at work to lift themselves out of extreme poverty. But the third segment raised questions and concerns because they are unemployed but—on paper, at least—look like they should be able to work; they are of working age, they are not disabled, and they are not in school or college. This research was useful to the commission in structuring its work and framing its recommendations to the Illinois legislature; the commission is still in existence and using the analytic framework developed by IMPACT, and Terpstra said they intend to update the analysis for the commission.

Second, IMPACT has also been engaged to perform analyses in support of the state's Human Services Commission. Working primarily with 1- and 3-year ACS estimates, IMPACT has generated extensive analyses of areas of need for youth services, disability services, and housing and homeless services in the state. For instance, ACS data have been used to study extremely rent-burdened households—those paying over half of their income on rent—and then projected potential demand and need for assistance from state funds targeted to relieve such households.

Terpstra concluded by reiterating that the ACS has been a game changer for IMPACT's work; very little of the work they have performed in recent years would be possible with any other data set. Accordingly, she noted that

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

IMPACT is worried about the prospect of the ACS being either defunded or hobbled. Such a development would radically change the way that social service organizations, advocacy groups, local media outlets, and state legislators have come to expect reliable information.

3–C TRACKING DISASTER IMPACT AND RECOVERY IN POST-KATRINA NEW ORLEANS

Allison Plyer began her remarks by noting a similarity of mission between her organization and Terpstra's. The Greater New Orleans Community Data Center (GNOCDC)—at which Plyer serves as deputy director—is a member of the National Neighborhood Indicators Partnership (NNIP)[10] organized by the Urban Institute. More colloquially, Plyer described the member organizations of the partnership as "a group of data geeks from around the country trying to help local communities work with data"; like Terpstra's organization, a major part of this effort is interpretive, making those data more understandable to a variety of users. In Plyer's case, a great deal of GNOCDC's work in recent years has been marshaling data resources to explain the devastating impact of Hurricane Katrina in late August 2005 and New Orleans' recovery. As she explained, data from the ACS have proved instrumental and, without ACS data, many constituencies would be "flying blind."

Though the Katrina example looms large, Plyer displayed a county-level map, shaded based on the number of times each county had been included in a presidential disaster declaration between 1964 and 2010.[11] The red shading on the map—particularly dense in areas like eastern North Dakota (flooding), southern California and the Pacific Northwest (flooding and wildfires), the tornado belt in the plains, and the Gulf Coast and Florida (hurricanes and severe storms)—underscores the basic notion that large parts of the United States are at risk of catastrophic disasters, natural and man-made. It is also well known that many heavily populated areas face significant disaster risk—the San Francisco Bay area and the large cities of Florida—but other risks are not quite as obvious. To illustrate this point, Plyer showed two graphics; one, showing New York City's defined hurricane evacuation zones, is sufficiently abstract to provide some "emotional distance" when viewed. But the other—a rendered aerial map of what central Boston could look like in the year 2100, under the combined brunt of rising sea levels, natural subsidence, and a category 2 hurricane storm surge—is considerably more arresting.

[10] Additional information on the partnership and its member organizations in 36 cities is available at http://www.neighborhoodindicators.org.

[11] An expanded version of this map is available from the Federal Emergency Management Agency at http://gis.fema.gov/maps/FEMA_Presidential_Disaster_Declarations_1964_2011.pdf; the higher-level page at http://gis.fema.gov/DataFeeds.html (cited by Plyer as her source for the rendered map in her presentation) contains links to the underlying data in a variety of formats.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Plyer observed that a previous National Research Council (2007a:72) panel had commented that "people who are responding to disasters complain that they are often operating in a data vacuum" and that, in this vacuum, "responders have difficulty setting short-term priorities, allocating scarce resources efficiently, or establishing strategic plans for longer-term recovery efforts." Plyer said that when a disaster strikes, the first "data" that are commonly available are images—photos from the media—that can have an effect on their viewers but that do not really provide a sound basis for characterizing who has been affected by the disaster, how much assistance is needed, and where those resources should be steered. GNOCDC originated in 1997 to help civic and nonprofit leaders in New Orleans to use data to work and plan strategically. By 2001, GNOCDC had developed its website (http://www.gnocdc.org) with easy-to-use data profiles (using data from the 2000 census and its long-form sample), broken down by the 73 New Orleans neighborhoods defined by the city's planning commission. GNOCDC also emphasized a willingness to answer questions from users unable to find information on the site; Plyer briefly displayed an "Ask Allison" page from the site, asking interested (or perplexed) site visitors contact information and offering New Orleans-area nonprofit organizations some free consulting time to address concerns. Plyer displayed graphs showing that the GNOCDC website averaged about 5,000 visits per month between August 2003 through July 2005—many visitors involved in planning and advocacy activities, as Terpstra described, and many local groups seeking data to support grant applications.

That steady state of visitors to the website (and the operations of GNOCDC itself) was upended as Hurricane Katrina gained strength over the Gulf of Mexico and tracked toward southeast Louisiana between August 26–28, 2005, prompting an unprecedented mandatory evacuation order from the New Orleans city government on the morning of August 28. As is now—tragically—well known, the storm made landfall southeast of New Orleans early on August 29; storm surge waters and catastrophic failure of the city's protective levees and floodwalls caused flooding in roughly 80 percent of the city's land area; the hurricane maintained strength as it hit the Gulf Shore of Mississippi before finally starting to weaken inland. The scramble for information in the buildup and aftermath of the storm is evident from the website access graphs; from a historic steady clip of 5,000 visits per month, GNOCDC's site experienced 40,000 visits in August 2005 and 80,000 in September 2005. Starting in October 2005, the visits to the GNOCDC site began to stabilize around the new steady average it has since maintained, about 15,000 visits per month. This massive spike in web traffic came as GNOCDC's small staff was—like the rest of the city's populace—physically displaced, to Georgia, Texas, and elsewhere. (Fortunately, the center's computer server was physically located in Kentucky.)

Plyer recounted some of the queries received by GNOCDC in the days and weeks following Katrina, and the first period of recovery, all of which she de-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

scribed as "tough questions"—often concerning detailed, fine-grained segments of the population—with "little good data" then available for guidance:

- From a national charitable organization, a request for information on the number of low-income seniors in areas across Louisiana, to steer their aid;
- From a state emergency preparedness agency, a request for solid counts of non-English-speaking persons in small areas of southeast Louisiana, in order to best print and circulate evacuation guides and other materials in Spanish, Vietnamese, and French;
- From the local public defenders' office, a request for a comprehensive demographic profile of the post-Katrina city—without which they would have no basis for determining whether court juries are actually representative of the population;
- From a state health agency, a request for detailed demographic statistics by small area, in order to make sure that state HIV/AIDS outreach efforts were being appropriately planned during the area's economic recovery; and
- From a large real estate group, an updated demographic profile for a specific high-ground New Orleans neighborhood, to make the case to a major potential client that the specific neighborhood was booming.

Faced with these questions, Plyer said that GNOCDC had to make use of the best data then available—flagging census blocks by their extent of flood damage (as determined by the U.S. Geological Survey and others) and aggregating small-area population statistics from the 2000 census (and its long-form sample) for flooded and nonflooded areas. This provided some useful insight on the number of people who could return to relatively undamaged (and high ground) parts of New Orleans when the city reopened. But the data—already 5 years old—were static, and so were not ideal for chronicling the city's recovery. As the city began to repopulate, it remained an open question of how the demographics of the city were changing, and which pre-Katrina residents were returning and which were not. Accordingly, GNOCDC eagerly welcomed the ACS as it entered full-scale collection—and was greatly relieved that the region would not have to wait until the 2010 census for a good reading on New Orleans demographics.

As 2006 and 2007 ACS data became available, GNOCDC began to generate series of analyses that it has since updated on an annual basis. For example, the ACS data showed that the populace of Orleans Parish had changed strikingly along some key variables: significantly fewer people who completed a high school degree and fewer households lacking access to a vehicle, a drop in the percentage of population living in poverty, and an uptick in the percentage of foreign-born population. Plyer conceded that their analyses lack a clear base for pre- and post-Katrina comparison because the ACS data for 2004 were still being

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

collected at the reduced, Census 2000 Supplementary Survey level; GNOCDC used the 2000 census long-form sample data as the basis for comparison with 2006 and 2007 ACS data. Later, in response to a question, Plyer noted that use of the ACS data was not struggle-free because they rely on the Census Bureau's population estimates breakdowns to weight the sample responses; if those estimates are off—as might reasonably happen in an area undergoing drastic population shifts—then the ACS estimates might be problematic. Plyer answered that this was a problem that GNOCDC and the Census Bureau struggled with; there is a mechanism for localities to challenge the Bureau's population estimates if they can submit alternative data and arguments, and Orleans, Parish had its 2007 population estimates revised upward after GNOCDC and local officials argued that the Census Bureau estimate seemed low.

ACS data on economic conditions have proven particularly useful in studying the rebuilding area in recent years because they are actually responding to multiple shocks. Post-Katrina, the number of people in poverty in New Orleans dropped significantly because, as Plyer said, "those folks had a hard time returning." But then the poverty rate ticked upward with the national recession. On related lines, GNOCDC began partnering with the Urban Institute to produce a series of housing reports, examining trends in greater New Orleans (principally using ACS numbers) and comparing them with other cities or the nation as a whole. These analyses document noticeably higher housing costs in Orleans Parish, post-Katrina and regardless of whether the housing is rented or owned/mortgaged. They also show that, in 2010, about 35 percent of New Orleans homeowners are housing-cost-burdened, in that they pay at least 30 percent of their pre-tax income on housing—a figure greater than the national average (but less than levels in cities like Las Vegas and New York). Similarly, GNOCDC has partnered with researchers from the Brookings Institution on a major project to track New Orleans' recovery and to put it into long-term perspective: comparisons with on the order of 30 years of trend data.

Plyer acknowledged that other sources can provide specific glimpses—for instance, Louisiana Department of Education data, compiled from local districts, were the source of one displayed graph tracking public school enrollment before and after Katrina. And, arguably, other sources might provide more detail (for instance, wage data from the Bureau of Economic Analysis). But she emphasized that, "to look at community well-being, we really needed the ACS." She said the ACS was pivotal in answering the myriad questions that came into GNOCDC, and without it all of those requesters—"the business sector, the legal community, policymakers, emergency preparedness folks, public health, nonprofits, and the media"—"really would have been flying blind for five years."[12] When disasters strike, "everything is uncertain"—and uncertainty in informa-

[12]In mentioning the media, Plyer endorsed Terpstra's comments about serving an interpretive role for media outlets; many of the reporters "definitely count on us" at GNOCDC because they

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

tion can paralyze public policy decisions and business investments; Plyer concluded that the ACS is critical to reducing this uncertainty.

Plyer used her closing minutes to discuss some work that GNOCDC is doing to work with one major challenge associated with ACS data, which is the presentation of the uncertainty inherent in the estimates. As Terpstra and IMPACT do for neighborhoods in Chicago and places within neighborhoods, GNOCDC is planning on rolling out additional ACS data tables by neighborhood for New Orleans. Plyer displayed a couple of screenshots of the tabular interface, including columns for the standard error of each estimate. She said that other NNIP partners had just posted the standard errors there, without much more context than that. GNOCDC is currently completing work on a small online widget as a companion piece for the new ACS tables; for instance, users can enter the percentages and the associated standard errors and have displayed a plain-English, yes/no statement as to whether there is a statistically significant difference or not. Users will be able to access a similar widget after they use other features of the site, such as combining income categories. Through these easy-to-use features, they hope to make the margins of error less mysterious (or frightening) to their downstream data users.[13]

3-D FRAMEWORK FOR USING DATA IN DISASTER PREPARATION

Closing the session, Russ Paulsen (executive director for community preparedness and resilience, American Red Cross) conceded that his remarks would be unlike other presentations in that they would not be ACS-centric; he said that he would be unable to sort out exactly what information Red Cross derives from the ACS versus the CPS versus any other data source, and that analysts at the American Red Cross national headquarters tend to use data products prepared by outside vendors. What the workshop steering committee asked him to do was to talk through a general framework through which data like those from the ACS are used in disaster response, recovery, and preparedness planning—the field in which Paulsen said he has some 20 years of experience, including the major response to Hurricane Katrina discussed by Plyer.

simply "don't have spreadsheets" and lack the ability to directly manipulate data. This thread would be revisited in the media perspectives session; see Chapter 4.

[13]The ACS margins of error at the census tract level, and some oddities in the data, are "the bane of our existence with ACS data," Plyer said—ending her talk by showing a map of tract-level median household income derived from 2006–2010 ACS data. Much of the picture that results makes sense; relatively wealthy and relatively poor areas of the city stand out from each other—except for a one-tract pocket of the otherwise low-income Lower Ninth Ward that shows a "higher than average" median income. As is well known, the Lower Ninth Ward suffered Katrina's worst devastation, and still struggles to recover. Put bluntly by Plyer, "we don't know what to do with this"—the anomalous item has baffled GNOCDC, local housing planners, and other officials.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Paulsen said that the American Red Cross uses data (and the ACS) throughout the entire "disaster cycle": preparedness for the event, response when it happens, recovery from the effects, and back to preparedness. Similar to the point Plyer made by showing the map of presidentially declared disaster incidences, Paulsen displayed statistics for American Red Cross disaster response in 2011: mobilizing about 28,000 disaster workers and 2.6 million relief items and opening just over 1,000 emergency shelters. The organization fielded responses costing at least $10,000 in nearly every state as well as Puerto Rico. Though the major, large-scale disasters—among them the 29 tornadoes, 27 floods, and 15 hurricanes American Red Cross responded to in 2011[14]—are most prominent in media coverage, Paulsen noted that the organization responds to thousands of smaller disasters—on the order of 70,000 house fires alone.

That demographic information from a collection like the ACS can be useful in responding to a large-scale disaster is fairly clear and was made vivid by Plyer's remarks. Granted, Paulsen continued, it might not be immediately obvious how that demographic information might be useful in responding to a house fire. But it is crucial for an organization like the Red Cross to have sound data on which to base its projections and its decisions on allocating resources and staff. The American Red Cross has developed formulas to project how many temporary shelters might be needed in an area when a disaster strikes, and this formula is based heavily on the demographics of the affected area. As Paulsen said, it is important that those demographic data be up-to-date for the model to work; projections based on 10-year-old data would not be terribly helpful. Similar projections and formulas are used in estimation and planning—predicting how many meals might need to be served during a disaster response, how many responders or vehicles are needed, and the potential cost of the response. In addition to these projections of the scope of the disaster and the requisite magnitude of response, data play a role in a variety of ad hoc reports while the response is in progress. The Red Cross is a broad national organization building on the work of individual chapters, so data from the disaster response stage inform the reports from local chapters and are used by the national headquarters to process local reports and to assess how resources are being allocated throughout the organization.

For a particular disaster—a fairly localized one, like a tornado—Paulsen sketched the basic way in which data like ACS estimates are used in disaster response. The first step harkens back to the beginnings of the frameworks described for using ACS data in health care and transportation planning (Sections 2–C and 2–D): quick assessment of the demographic profile of the affected area (say, a county), and at a more granular level as appropriate, to try to get a sense of where impacts are likely to be worst and where Red Cross ser-

[14]Per Paulsen's slides, other major disasters covered by American Red Cross in 2011 were 45 multifamily fires, 10 wildfires, 4 blizzards, and the August 2011 Virginia earthquake.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

vices might be needed most urgently. Specific data variables considered in this stage include population density, age breakdown (with particular emphasis on the elderly, who are more likely to need—and use—emergency shelter service), language barriers, and combinations of age and ethnicity that might correlate with dietary and nutritional requirements. A variety of economic variables—percentage living below the poverty level and housing tenure (renter/owner)—are also important to assessing outcome—so, too, is the variable on the extent of housing vacancies (including seasonal homes) in the area. Mapping these data is often the most effective way to focus services and resources to the subareas of most acute need.

In disaster response, Paulsen said, the "name of the game is getting ahead of the curve as fast as you can." In its planning, the Red Cross works to have people on hand with the specialized skills needed to address particular problems—but it still takes time to get them into position where they can do the most good. A community like Pascagoula, Mississippi, might ordinarily have a local Red Cross staff of two people; during the response to Katrina, the Red Cross needed to deploy on the order of 25,000 people to Pascagoula, many (if not most) from outside the local area. So, he stressed, the data-driven assessments of impact and need in the wake of a disaster must be completed quickly, and the value of data in making effective resource allocations decreases sharply with the oldness of the data. Put bluntly, he said, "we are going to make bad decisions" if all that is available are 10-year-old data. Paulsen diverged from his presentation to comment on the challenge raised by Census Bureau staff in the earlier discussion session (Section 2–F) on whether a shift of a few weeks or months in data release time would really affect results. As an end user and a manager doing disaster response, he said, "I am really hungry for speed because I feel like I am going to make better decisions the more current the data is"; the actual effective difference in response of a few weeks or months of increased data recency would depend on the magnitude of the change caused by the disaster. For longer-term planning and for disaster preparedness—as he would discuss next—change is slower and so lags might not matter. But fresh, recent data are invaluable in the immediate response and recovery phases of the disaster cycle.

Paulsen echoed Plyer's comment about the usefulness of timely, accurate data in the next step in the cycle—recovery from the disaster. However, given that Plyer had discussed recovery in great depth, Paulsen moved on in the cycle to discuss the role of data in disaster preparedness. He said that preparedness is his new major focus in work at the Red Cross; having led recovery from Katrina and having worked in disaster response for a long time, he said that he wants to really get ahead of the curve and try to reach a point where people do not have to suffer so much when disasters occur. As he put it, Red Cross services cannot be instantaneous at every disaster everywhere, no matter how big the Red Cross is; "people have to do some stuff on their own, and they can reduce their own risk."

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

The Community Resilience Strategy that he is now working to instill in individual local communities is a four-step process, and each of the four steps relies on good quantitative information:

- The first step is *community assessment*—just as in disaster response, thinking through the first steps of a preparedness plan is to take stock of available existing resources.
- Integral to that overall assessment is understanding the *demographics* of the community.
- These steps combine to allow a community to document its assets and to assess gaps—in other words, to *map the community's vulnerabilities* and be aware of them.
- Finally, with the vulnerabilities known, an *action plan* to address them is developed.

Paulsen said that this kind of strategy is difficult to accomplish at high levels of geography; it is tough to do in a focused way in whole states, counties/parishes, or cities. Rather, he said, this is work that needs to be done at the neighborhood level to be most effective—and so data from the ACS, to drill down to very small areas, are critical.

Specific demographic variables covered by the ACS that Paulsen suggested are most useful in a pareparedness strategy include total population, age breakdowns, race and ethnicity breakdowns, foreign-born population, language other than English spoken at home, educational attainment, household structure and median household income, percentage living in nonpermanent housing stock (i.e., mobile homes), percentage lacking phone service, percentage lacking an available vehicle, and percentage unemployed (persons over 16 not in labor force). These data are important to getting the "big picture" of the community and honing in on vulnerabilities—and they are vital to Red Cross staff when disasters do strike, to focus attention on areas of greatest need.

In summary, Paulsen said that he cannot imagine having implemented response and recovery to Katrina without the work done by Plyer and GNOCDC—or working through longer-term recovery issues in the area without data more current than 2000 vintage. As a professional in disaster response, he cannot imagine an effective response with really old data; likewise, it is inconceivable that either effective disaster planning or preparedness can be done well without fine-grained small-area data like those provided in the ACS.

Paulsen closed by making brief note of one specific application in which Red Cross staff use census, and now ACS, data regularly and extensively: blood services. The major blood types in the commonly used ABO classification system (A, B, AB, and O) and their + and − variants (depending on the presence or absence of the Rhesus [Rh] factor) occur in different proportions among race and ethnicity groups. Beyond those main types, some rare blood types are ef-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

fectively unique to specific demographic groups.[15] As a result, the Red Cross uses ACS and census data to understand the potential blood donor markets and the potential recipient markets for local hospitals. By effectively modeling population by blood type, Red Cross can also mount special collection efforts in particular small areas and target collection sites.

[15] The American Red Cross provides overview pages on blood types and the rare blood types that vary by race and ethnic origin at http://www.redcrossblood.org/learn-about-blood/blood-types and http://www.redcrossblood.org/learn-about-blood/blood-and-diversity, respectively.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

– 4 –

ACS and the Media

In earlier workshop presentations, speakers from two nonprofit organizations spoke about one important role they play with respect to American Community Survey (ACS) data: serving as an "interpreter" of sorts for the complex data, both to the general public and to media outlets. The media are an important consumer and user of ACS data (and products derived from those data) because they serve to communicate important findings and trends from the data; many more people are exposed to the data through the window of the media than will ever access and work with the data themselves. The media are an important front line in the communication of results of the survey and an important and diverse data user constituency themselves—ranging from the small media outlets described by other speakers as lacking any capacity for original data analysis to outlets capable of working a new ACS release into a complete package of stories on changes in American life.

The workshop presentations in the block dedicated to media perspectives included an important and considerable range of views. The workshop steering committee asked Haya El Nasser from *USA TODAY* to speak generally about the way in which news stories are carved from new data releases and general challenges of writing and communicating about the ACS (Section 4–A). From this general discussion, the session pivoted to a profile of a specific, intensive, data-driven exploration and how it came to be—a multipart profile of the impact of immigration (and illegal immigration) on the California economy developed by Ronald Campbell of *The Orange County Register* (Section 4–B). The final speaker—graphics editor Ford Fessenden of the *The New York Times* (Section 4–C)—took El Nasser's comments a bit further by talking about the unique opportunities and challenges created by ACS data in the graphical presentation

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

of data, typically through the *Times*' online platform.[1] The discussion period following the speakers' presentation (Section 4–D) revisited the thorny problem of the presentation (or lack thereof) of standard errors associated with ACS estimates.

4–A FINDING STORIES IN ACS DATA

Though she began her remarks by calling herself "probably the least technically savvy" of the speakers, Haya El Nasser took care to note that her newspaper—*USA TODAY*—has a longstanding commitment to working with and presenting census and ACS data, and that she personally has been writing on related issues for almost 16 years. Noting the comments made in the previous session, she commented that she does have Excel and enjoys working with extracts of the data herself; more generally, she said that *USA TODAY* is fortunate to have the tools and the capacity to do extensive work with the raw data. In that regard, she credited *USA TODAY* database editor Paul Overberg as their "secret weapon" in working with full data releases and getting things into the shape to form finished stories, and many of her workshop comments described the interactive process between her and Overberg in mining the stories from ACS data.

Walking through the process by which a "data dump"—what reporters like her tend to call a new release of ACS data (or a new set of results from the decennial census)—she began by noting that some directions for data-driven stories come from basic hunches and instincts. Having covered stories in this area full-time for a number of years, El Nasser said that she talks to a lot of people about demographic trends; she reads a lot of material and travels to work on stories. From all of that—as well as her experiences within her own community—she said that she makes a habit of thinking about these "hunches of changes," whether they are on things like changes in household size or moving patterns (families moving or not moving in greater numbers) or a sense of shifts in demographic structure brought on by economic conditions (e.g., older children moving back to the parental home). El Nasser said that some of these hunches make their way into written notes (or just Post-It® notes) while she keeps others in her head, but that she compiles them for discussion with Overberg when a new census or ACS release is about to come out. In addition to being another sounding board on the possible veracity of a hunch, Overberg—as the person responsible for wrangling the raw data—also serves as a "reality check" on what

[1] In addition to the three journalistic perspectives presented at the workshop, database reporter Phillip Reese from *The Sacramento Bee* contributed a short description of stories that he and his newspaper have specifically derived from the ACS for the workshop's case study/agenda book. That summary, in turn, is a concise recap of a presentation he delivered in ACS use by the media at the Population Association of America annual meetings shortly before the workshop, in San Francisco in May 2012.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

types of analysis can or cannot be done on deadline for a compelling "first-day story" to accompany the new data release. Through these conversations, they narrow their focus down to a few areas or hypotheses where they might have anecdotal evidence but also feel that the data analysis might corroborate what they see on the ground.

With those topics (and seeds for possible stories) in mind, they then wait for the numbers to go up, and start to work through possible stories while they have access to the data under an embargo period. After Overberg does initial work with the large data files and distills them into more workable form, El Nasser said that she starts looking at the newly created spreadsheets, playing around with the rankings and rate of change. Quite deliberately, they put a time limit on this initial exploring and eyeballing—typically within a couple of hours—and talk through what they see. In addition to more refined analysis, this initial exploration leads to the next step, which is tapping the very broad pool of sources—demographers and other social scientists—who can weigh in on what they are noticing in the data. El Nasser candidly conceded that this early discussion and corroboration of data trends with outside sources might be a technical violation of the ACS embargo policy lamented by Terpstra (Section 3–B), but it is essential to getting any compelling story done on time (and having confidence in the trends they observe).

Sometimes, she said, there emerges a good, strong theme for the "first-day" story based on the new data. At other times, it takes more time than the unforgiving release deadline permits to develop such a strong theme. In those latter cases, the first-day story becomes what they colloquially call "the stew"—highlighting a variety of changes in different variables, to give readers some sense of the extent of demographic changes occurring in the country. Not meant as a pejorative, "the stew" actually serves to highlight the strength of a data source like the ACS, when it can be used to describe trends across many variables and for all varieties of geographies (region, key states, large metropolitan areas, and so forth).

Whether a single-theme story or "the stew," work on the story develops along two principal fronts (though certainly in collaboration with each other). Through her own looks at the data and discussion with sources, El Nasser works on the reporting of the story—and works through the challenge of "turning the numbers into words" and making them comprehensible to the broader public. Meanwhile, Overberg and other *USA TODAY* staffers work on the maps and graphics to accompany the story. In modern journalism, the focus is not just how to render data views from the ACS or other sources in print, but also in how they may be rendered on the newspaper's website. Naturally, she said, this raises issues like making sure that the online product is in sync with the printed version and—with the necessary design and computation—the process can become extremely elaborate for a 1-day data release (but, she conceded, also a lot of fun).

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Beyond the first-day story, they have the luxury to go back to the data and mine it for additional details and stories in the following weeks or months. Detailed drill-downs into the data can be done while reporters travel to develop "focus stories"—fuller reporting packages to illustrate what the numbers are showing. She said that first-day stories tend to emphasize hot-button topics of known reader interest—among them trends in commuting and in housing costs (e.g., the share of household income that goes toward housing costs). With specific reference to housing, new data releases also commonly have a separate housing story developed by a reporter in *USA TODAY*'s financial section, typically with insight and analysis work by the Joint Center for Housing Studies at Harvard University. As an example of a longer-term focus story, El Nasser said that they are currently working with ACS data for a story package on trends in mobility in the country; the release of data on mobility trends as detected in the Current Population Survey (CPS) is a primary spark for the story but—as noted by earlier speakers—the much smaller CPS sample size limits the degree to which it can be disaggregated by geography or demographic group. Hence, they are working with the ACS data to get at the richer detail, to round out the story. Likewise, shortly before the workshop, *USA TODAY* ran a story that dipped back into ACS data on housing tenure to discuss the degree to which the United States is becoming a "renter's nation" (increasing share of renters versus owners).

Turning from the general process to specific aspects of *USA TODAY*'s use of ACS data, El Nasser made the point that the overriding value of the ACS for news editors—and the reason why she and *USA TODAY* particularly like it—is the freshness and the timeliness of the data. Put most simply, "new data make news" in a way that older data simply cannot. She recounted particular memories from the 1990s, going to editors with stories in 1996 and 1997 making good use of the data from the decennial census long-form—good stories, of which she and fellow reporters were proud—and having them rejected out of hand, with dismissive looks essentially saying "Are you crazy?" To an editor—focused on the news of the moment—a story that looks like it is based on 6- or 7-year-old data is practically dead on arrival and the counterargument that the long-form data are really the only thing out there does not gain much traction. To be sure, she said, *USA TODAY* ran some stories of that variety—and still does—"but nobody got too excited about them." Further, they do occasionally get buy-in from editors to conduct more archival or historical stories—for instance, for a longer piece accompanying results from the 2010 census that looked at broad major national changes relative to the 1990 and 2000 censuses[2]—so the rejection of seemingly old data is not absolute. Still, newer data get the bulk of editorial

[2]El Nasser recalled that the story she had just mentioned that accompanied the 2010 census results also made use of the most-current ACS data, on items such as marital status and age at first marriage, for comparison of trends with the previous census long-form samples.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

attention and having good data to talk about how the country is changing every year, with fresh numbers every year, is "quite amazing for us"; it is that freshness and newness of results that makes the ACS an irresistible source for the media.

Regarding the geographic levels that *USA TODAY* tends to use, El Nasser said that they value the ACS for its ability to provide national-level "bird's eye views" on trends as well as profiles by state, county, and metropolitan area. Their use of finer-grained geographies than county or metropolitan area is relatively rare. They have not really pursued tract-level analysis—and have not really needed to—but that option is certainly open, for instance, to focus a story on the impact of the foreclosure crisis on a particular city neighborhood. She did recall one instance where Overberg used data from the Public Use Microdata Sample (PUMS) in support of a story on the diversity of the 2010 incoming kindergarten class and to suggest their demographic future. Overberg has also used 5-year ACS data at the Public Use Microdata Area (PUMA) level for a story on children in poverty; the text looked at poverty rates by different age groups and studied income ratios at the county level, but the full package included the glimpses at the PUMA level.

El Nasser made the first mention at the workshop of a peculiar identity crisis faced by the ACS, a topic that would recur in subsequent discussion. When estimates from the "2010 ACS" rolled out in 2011, on the heels of more detailed releases from the 2010 census, El Nasser and others began working on a first-day story covering six or so trends that were perhaps linked to the national recession's impact on individual Americans' lives—delayed marriages, fewer divorces, more crowded households, changing housing vacancy rates, and so forth. But initial response from editors was skepticism and confusion: why should we look at and use these numbers instead of the 2010 census? And so, El Nasser said, they had to go through a mini-seminar of sorts to clarify the differences and distinctions between the census and the ACS. She commented that she had been asked to talk about communicating the uncertainty inherent in ACS estimates to readers—and she would—but that the editors' reaction spoke to a potentially more fundamental concern: continued uncertainty and misunderstanding about exactly what the ACS is. One massive consolation is that the Census Bureau has an excellent reputation as a trustworthy source of information; as El Nasser put it, her editors—who act as a filter by taking on the role of the reader—"couldn't care less if it is the ACS or CPS or whatever . . . the bottom line for them is that it comes from the Census Bureau."

This combination of trust in the Census Bureau and confusion over what exactly the ACS means is the reason why media stories commonly use the generic references "census data" or "data from the Census Bureau." El Nasser said that reporters certainly know the precise source of the data in the stories, but find it difficult to win these battles; proper attribution of the ACS as a source begs (in editors' interpretations) for an explanation of what the ACS is and how it differs from the census, and that is a tough sell when space is at a premium. What

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

she and Overberg have done is to try to take care to always insert "American Community Survey" in the source line for *graphics* for the stories, so that people who are interested in following up on specifics have the clue that the actual source is the ACS. Outside of the first-day story crunch, the less time-sensitive focus stories are more likely to include ACS attribution in the text—particularly if they combine data from multiple surveys or censuses.

El Nasser was candid in saying that similar confusion or misunderstanding is the explanation for why "we don't talk about margins of error in stories"—outside the specific application of political polling, where the concept of a margin of error has come into public parlance, the concept of standard errors and uncertainty is sufficiently opaque to editors (and general readers) that the general strategy is to avoid it in story text. She was quick to point out that, in reporting the story, they are cognizant of the standard errors that are now prominently presented in ACS tabulations; in fact, she commended the Bureau for doing a good job in presenting the standard errors and thus making it fairly evident to users whether some basic comparisons are statistically significant. She relayed a comment and point of criticism from Overberg that the Bureau's online access tools are not friendly for handling some kinds of basic manipulations—for instance, combining detailed age groups into broader categories and recomputing the standard errors. To that end, adding a calculator widget (of the sort being developed by Plyer [Section 3–C]) or an improved facility for custom-building tables would be greatly beneficial.

El Nasser concluded by reiterating that the bottom line for a news organization, driven by timely information, is that it will use the data that provides the most current, timely insights at the geographic levels of greatest pertinence. The decennial census has long been a trusted source for this information, but the ACS has become indispensable for its topics and geographic coverage. Noting the recent legislative developments, she commented that she and *USA TODAY* certainly hope that the ACS remains viable; her own sense is that ACS funding is likely to weather the immediate budgetary storm but that some period of costly and time-consuming work to square the ACS with new expectations is likely to follow.

4–B DATA-BASED INVESTIGATION: IMPACT OF IMMIGRATION IN CALIFORNIA

While El Nasser described the general process of generating news stories from ACS data—and, in particular, stories on a tight first-day-of-release deadline—Ronald Campbell of *The Orange County Register* described his own personal experience with a longer-term data-based investigative report. His series of stories on the growth of the immigrant labor force in California since 1990 and its impact on the state's economy was published in the *Register* in

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

September 2010[3] (and would ultimately be recognized in April 2011 with a "Best in Business" award from the Society of American Business Editors and Writers) and synthesized data from previous decennial censuses as well as the then-available estimates from the ACS. He said that the experience could be described as a case study of data-driven reporting or—at times, and less politely—as a "project from hell"; though generally favorable about the data and their utility, he would end his remarks with an ominous prediction about ACS use in the media, and so set the theme for later discussion at the workshop. (In addition to the condensed version of the narrative in his presentation, Campbell contributed a fuller written account for the case study/agenda book; this summary draws detail from the written case study as appropriate.)

Campbell said that he has long been involved in computer-assisted journalism—the field's term for using technology and databases to tell stories—and hit on the idea of using PUMS data from recent censuses and the ACS to assess prevailing conventional wisdom about immigrant labor in California. Immigration is an impossible topic for reporters in California to avoid; as he began work on the series, he said that he knew that one bottom-line statistic—that immigrants constitute roughly one-third of the labor force in California—is an "extraordinary number." He added that California politicians typically like to speak of the state's economy on a global scale—if it were treated as a country in its own right, California's economy would be the eighth or ninth largest in the world. So the role of immigrant labor in an economy of that size was an interesting hook for a story to pitch to Campbell's editors. A clinching detail was the knowledge that the only developed economy outside the Middle East with a larger proportion of immigrants in its workforce is the vastly smaller economy of Luxembourg.

This project was a difficult one to take on, for a variety of reasons. The general topic was "wildly controversial" and had been heated for many years, particularly following debates over the state's Proposition 187 in 1994;[4] common assertions cast in debates included that immigrant workers in California were predominantly illegal/undocumented, poorly educated, and poorly paid drains on state and local resources. As a piece of journalism looking at trends over several decades, the story also had to be compelling as both a history lesson and as fresh, topical news. With such a contentious topic, Campbell knew that

[3] An overview "Series at a Glance" page at http://www.ocregister.com/articles/choices-265585-immigrants-driven.html includes links to all the stories and sidebars from the series. Supporting spreadsheets are also hyperlinked in the online story footnotes.

[4] Passed with a 59 percent vote in a 1994 statewide referendum, Proposition 187 (originally introduced in the California legislature as the "Save Our State" initiative) would limit public education, health care, and other social services to persons verified to be U.S. citizens or legal immigrants. Challenged on constitutional grounds, the proposition was essentially stayed from implementation by federal court injunction shortly after its statewide passage; in 1997, it was held to be unconstitutional in the U.S. District Court for the Central District of California and, in 1999, the state of California withdrew its appeals.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

his work and analysis of the relevant data had to be "bulletproof"; as he put it, he and his editors "knew that, no matter how careful we were, readers would give us hell." Accordingly, he set two basic parameters as he set about assembling the story: first, to establish a "high threshold of proof" up front—using 95 percent confidence intervals and describing results in text only if they carried a margin of error of plus-or-minus 5 percent—and, second, getting permission from his editors to post *all* of his numbers and spreadsheets online. Through this posting, Campbell said that he wanted readers to see not only the numbers he worked with, but also those—clearly branded in red shading—that failed his "threshold of proof" test. Echoing El Nasser's comments about the inability to describe statistical concepts when space is at a premium, he also lobbied his editors to devote space in the story layouts to a "nerd box"—text boxes describing the underlying methodology.

Campbell said that it did not take long to encounter difficulty in stitching together a story from a variety of census and ACS data sets of differing vintages—conceptual definitions (such as occupation categories) change over time and geographic boundaries are in similar flux. More fundamentally—and counter to previous speakers' highly favorable comments about the ACS sample size—the annual ACS sample size (as reflected in 1-year PUMS files) is so small relative to decennial census long-form samples that many comparisons one might want to make are simply not viable. To make more comparisons work, he turned to the most recent 3-year PUMS file (the 5-year numbers remained more than a year away from release) and specifically to the version of the file distributed by the IPUMS project (described below in Section 6–B). He acknowledged that direct comparison of a 3-year interval estimate to a point-in-time estimate like those from the 1990 or 2000 census long-form samples runs counter to the Census Bureau's usual guidance, but he said that he felt that he had little choice. Other workarounds that he implemented to facilitate his analysis included working almost exclusively with ACS percentage estimates rather than counts and to focus on larger geographies. To illustrate the point, he described his first attempt to use the 2006–2008 ACS PUMS file to derive the number of immigrant workers in each of the state's 233 PUMAs; the screenshot of the table he displayed was filled with red shading—his indicator for estimates with relative standard errors greater than 5 percent—because all 233 estimates failed the test. Switching from counts to percentages of foreign-born workers, Campbell was able to generate usable results for all but 17 PUMAs.

For purposes of mapping data, Campbell said that he used the "consistent PUMAs" developed by the IPUMS project, which are designed to be consistently identifiable units (geographic areas) over the 1980–2000 censuses as well as the ACS samples from 2005 forward. These consistent PUMAs tend to be larger in size than the most recent 2010-ACS-vintage PUMAs, which further helps with the margin-of-error problems.

For this investigative series, the key issue was not just profiling the immi-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

grant population but seeing how they are distributed across the workforce—what kinds of jobs do they hold, and what proportion of various professions or industries are held by immigrants? Campbell said that this examination of a particular variable type proved difficult. The long-form samples from the 1970–2000 censuses were able to support conclusions about the immigrant population within "hundreds" of professions, and comparing across censuses suggested those occupations where the immigrant share had increased dramatically. But the picture became murkier with the ACS data. By his reliability standards, Campbell said that he could only obtain good results from the 1-year 2008 ACS numbers for 44 out of roughly 334 job categories coded in the ACS data (this looking at the state as a whole, with a pool of some 6 million foreign-born workers). Results improved turning to the 3-year, 2006–2008 ACS data; the three numbers supported reliable estimates of the foreign-born share of 108 job categories, which he subsequently reduced to a set of 90 for comparison with the earlier censuses. In general—and, again, not challenging previous speakers' comments about using the ACS relative to the CPS but instead summarizing work in his own application—Campbell said that he found it frustrating that "I deliberately had to be less specific in what I was aiming for" (in terms of the number of years of pooled data and in geography) to get reliable results.

In the end, his series on immigration in California documented and highlighted some trends that ran well counter to popular "received wisdom." As he summarized in his written account for the case study:

> Drawing on four decades of census data, the series showed that immigrants were responsible for most of the growth of California's labor force since 1990, that hundreds of thousands of immigrants worked in high-wage jobs as doctors, scientists or engineers, and that they had little or no apparent effect on the income of well-educated [U.S.-born workers].

On the last point, Campbell raised the example of two job categories that had experienced major influxes of foreign-born workers between 1970 and 2008, registered nurses (RNs) and automobile mechanics. Specifically, the 2006–2008 ACS data suggested that the foreign-born shares of these two job categories were roughly comparable in 1970 and in 2008 (as reflected in the 2006–2008 ACS), about 37 percent of RNs and 45 percent of mechanics in 2008. True, inflation-adjusted wages between the two job categories diverged greatly over the years—mechanics made slightly more than RNs in 1970, while the mean wages of RNs in 2008 was more than twice that of mechanics. But Campbell argued that the difference beween the categories is not immigrant or foreign-born status but the degree of educational attainment that has come to be expected in the two jobs, with RNs shifting from typically having 2 years of college (or less) in 1970 to at least a bachelor's degree in the 2008 data.

Campbell wrapped up his comments (and his written case study) with what he called "issues for ink-stained wretches"—guidance for reporters thinking about undertaking similar analysis with ACS data. He reiterated the importance

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

of setting (and maintaining) appropriate thresholds for confidence intervals and margin of error, and of documenting the chosen methodology before starting analysis, revising it as need be during analysis, and (ideally) publishing it with the finished story. However, he cautioned his peers that "ACS promises more than it delivers"—the great promise of the ACS for reporters was that tract-level data would be highly reliable, and so provide reporters with neighborhood-by-neighborhood comparisons and "thousands of different stories." However—large though the ACS sample may be—he said that the standard errors on such fine-grained estimates simply do not support that level of scrutiny. He cited the example of his home census tract in Orange County; "even a casual inspection of tract-level data" from 5-year ACS data yields some estimates that he "would not hesitate to use [in] the paper" (median high school educational attainment of 97.4 percent with a 1.4 percent margin of error). But it also includes others that seem sufficiently out of whack (median household income about $119.9 thousand, plus or minus 16.6 percent). In Campbell's opinion, the large margins of error make the ACS tract-level data a less reliable source for reporters than previous decades' census long-form samples.

Commenting that he, El Nasser, and Fessenden constitute a somewhat unrepresentative sample of reporters in terms of sensitivity to data analysis issues, Campbell agreed with a basic point made by Terpstra and Plyer (Sections 3–B and 3–C): most reporters do not (or cannot) perform their own analysis of numbers. Though data-driven (or computer-aided reporting) is making some inroads, it is still a specialty—and a general fear of numbers still pervades many newsrooms. However, where Terpstra and Plyer emphasized the corresponding need for "interpreters" to make ACS data and findings more usable by the media, Campbell feared that some reporters will continue to try to use the data and yet remain reluctant to really probe the estimates' standard errors and understand the "potential pitfalls." Hence, he ended with a cautionary prediction: that sometime in the next several years, a news organization will have to retract a major story—based on ACS data—"because a reporter failed to understand the margin of error that is built into that data."

4–C GRAPHICS AND PRESENTATION OF DATA TO NEWSPAPER AND NEWS WEBSITE READERS

Taking the podium immediately after Campbell's prediction, Ford Fessenden (graphics editor, *The New York Times*) agreed with the premise that "there will be a huge problem at some point" when failure to understand ACS margins of error will force the retraction of a big story. But he stressed a more upbeat take on using the ACS data, particularly for small areas. He described his bottom line position as being that "what is most important about the ACS is exactly what is most problematic about the ACS"—its capacity

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

for allowing users to drill down to finer geographies and smaller groups. As Campbell said, the estimates can become much wobblier at those drilled-down levels but "there are massive amounts of truth there." He suggested that there is certainly a limit to the degree to which ACS data can make its way into print—no newspaper would want to print endless tables, and the final stories can only contain so many numbers before they lose their narrative. But while the *Times* might not "want all this stuff in the paper, we [still] want it all"; they want access to the data for custom analyses, they want that access to provide readers with the capacity to directly interact with the numbers, and they conclude that "pushing the envelope" by drilling down to detail in ACS data "is a good thing."

Fessenden continued that journalists too frequently write about trends at macro levels but are really guessing about the real trends. As an example, a news story in New York might assert that the complexion of the borough of Queens is changing because of an increasing number of immigrants (in contrast to other boroughs of the city)—an interesting story, and one that could be made very compelling with some anecdotal descriptions. But that macro-level story masks even more interesting changes going on within individual neighborhoods in Queens—concentrations of different Asian groups in Flushing and other neighborhoods, for instance. It is the small areas that really tell the story, Fessenden said, and it would be "a terrible waste" to either ignore the finest-grained data or to be "fearful about presenting" them. He commented that the ACS has vast potential for reporting—not only for spawning stories about what new releases of the data "say" (relative to prior releases) but as an invaluable source for background reporting in other stories.

Fessenden said that the thing that needs to be understood in order to explain the interaction between the media and the ACS (or other data sources) is very basic: that print journalism is fundamentally about narrative writing, telling the story. Data might suggest a direction for a narrative, or provide pointers to groups of people who may be interesting to profile in a story, or help corroborate individuals' experiences in the service of a story—but to newspapers and their editors, the basic truth is that data in their own right are "boring" relative to the narrative. This explains, in part, the aversion to numbers and methodological discussion described by Campbell and El Nasser. Computer-assisted or data-driven reporters have made some inroads in infusing data into the newspapers; Fessenden noted that he was such a computer-assisted reporter before becoming a graphics editor, and that reporters like Campbell, El Nasser, and himself serve as "Billy Beanes[5] of the newspaper business" in trying to weave stories from data. The unwillingness to distract from the narrative carries over to the presentation of uncertainty inherent in ACS estimates. During his pre-

[5]Billy Beane, general manager of Major League Baseball's Oakland Athletics since 1998, is known for basing player assessments on statistics and data analysis (rather than the instincts of scouts), and was the subject of Michael Lewis' 2003 book *Moneyball*.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

sentation, Fessenden directly echoed Campbell's language—by and large, "we [at the *The New York Times*] don't and we won't communicate that uncertainty" by directly presenting the standard errors, and the phrase "margin of error" "will not appear except in the graphics about polling," where it has become accepted. Instead, Fessenden said, what is incumbent on him and his peers is to study the margins of error and their implications—"to try to understand it ourselves and then tell people something we believe is true."

Extending that point, the other main point that Fessenden said that he wanted to make about data-driven reporting and the increasing role of the ACS concerned "correctible errors." An important role played by data users in news organizations is trying to catch errors (or validate anecdotal findings) before they make their way into print. As an example, he cited a recent instance at the *Times* of a reporter being tasked to do a story on a new, hot topic in the city: commuting by bicycle. The reporter started with a glance at some data sources, including ACS tabulations, but—"with no understanding of [margins of error] or outliers or anything like that"—the results were perplexing; estimates "were all over the place." The reporter then went out and conducted interviews, returning a written story that "beautifully" recounted what the reporter had observed. The lead of the story suggested that many people commute by bicycle, but the story went on to quote "a bunch of people who had never seen anybody on a bicycle," and the story wound up suggesting that these interview stories confirmed what the data had said. Fortunately, Fessenden said, the story was stopped from publication by someone in the approval cycle who recognized that making use of the newspaper's "ability to go a little deeper" into the data could improve the story. In this case, a correctible error was caught before it reached publication, and the story—and its supporting data analysis—could be reworked to give a clearer impression of what is actually occurring with bicycle commuting.

As he displayed and spoke about some graphics projects that he has worked on for the *Times*, he noted that the newspaper is not always overt in presenting the standard errors associated with the estimates. One such project allows users to click and construct households of various configurations (e.g., male/female unmarried partners with one child under 18 and one child over 18); the count and prevalence of such households based on 5-year (2005–2009) ACS data are displayed, along with a quick race/ethnicity group and household income graph comparing those households to the overall population.[6] He conceded that the calculator does not attempt to show the specific-household counts as an interval estimate, and so it might give some "sense of false precision." But he said that

[6]This family project is located at http://www.nytimes.com/interactive/2011/06/19/nyregion/how-many-households-are-like-yours.html; the credit line notes that the calculator uses the IPUMS version of the file (see Section 6–B) and that it makes derivations from the data by Social Explorer, Inc. (see Section 7–C). A third graph pulls data from the 1900–2000 decennial censuses for the user-specified household type (when that detail can be extracted) and graphs the count over time.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

they assume that people know that as the household is made more and more complex, the numbers get smaller (and more uncertain). Understanding that the calculator could drill down to counts "so low that even we get kind of freaked out" about the corresponding level of uncertainty, the page handles the problem by suppressing results for groups that fall below some threshold level (around 50 people in the 5-year sample or a weighted count of less than 1,000). For such rare groups (e.g., single male primary householder with two foster children), the calculator says only that there are "fewer than 1,000" such households and says that "there are too few households of this type to chart."

As another example—emphasizing the richness and geographic detail in the ACS data—Fessenden turned to a series of maps also generated from the relationship variable in the ACS, showing the prevalence of some nontraditional household types (e.g., "missing generation" households including a married couple and a grandchild, married couples with two adult children living at home).[7] Some of the maps show the nation as a whole, at the metropolitan-area level (as well as balance-of-the-state remainders)—or map New York City at the census tract level. Fessenden said that this set of maps was interesting to readers—giving a picture of household structure in the nation that cannot be obtained from any other source—but was particularly interesting because it spurred *The New York Times* staff to probe for explanations for some of the trends evident in the maps. The "missing generation" map—married couples with a grandchild—was an interesting one to parse, showing high levels of these types of households in the South and in rural areas; working with sociologist and demographer sources, the staff suggested that this might be "a map of drug abuse and rural poverty." Pushing down to the tract level within New York City, the maps are undoubtedly busier yet still show some strong patterns; the tract level is useful as being specific enough to see interesting things yet not so specific that "[a reader] says, 'Whoa! You are wrong; that is not right in that place.' " Among the interesting patterns evident in the tract-level data are strong concentrations of unmarried partners in some parts of Queens (like Long Island City) and increasingly upscale parts of Brooklyn; likewise, he noted that the map of single mothers with children under 18 overlaps considerably with pockets of poverty and minority-population concentration.

ACS data were also used to generate a text-based map by metropolitan areas that ran in the *The New York Times* in January 2012. For each metropolitan area, Fessenden and his colleagues derived the 99th quantile of household income—that is, the income level necessary to qualify as "the top 1 percent"—and printed the area name and estimates arrayed by their rough geographic location.[8] Fessenden recalled that this project was frightening; rushing on a dead-

[7]The map series can be viewed at http://www.nytimes.com/interactive/2011/06/17/nyregion/maps-of-family-types.html?ref=nyregion. The maps draw from the IPUMS version of the ACS data file.

[8]The original text-based map does not appear to be archived on the *The New York Times* website,

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

line, they encountered some unusual outliers. Consulting with their in-house experts as well as some outside consultants, they eventually reassured themselves by looking at other quantiles (e.g., 95th rather than 99th) and found the results to be relatively stable. One or two anomalies remained—"you don't see them there" on the map, because they omitted the areas from the final map, because they did not want to commit a "correctible error" themselves.

Summing up his basic approach to dealing with the uncertainty inherent in ACS estimates—as he put it more colorfully in his slides, "how we learned to stop worrying about sampling error and embrace the ACS"—Fessenden said the strategy is to try to emphasize trends. They try to keep the level of geographic resolution reasonable—small-area units in which differences in standard errors are not incredibly volatile. He also suggested that the very practice of mapping the data is an informal mechanism for borrowing strength from nearby errors; visually, the reader can reconcile rates in one area with those in the immediate vicinity and so have a reasonable intuitive feel for larger trends. He cited a personal favorite example that ran in January 2011—a "mash-up of immigration [foreign-born householders] and race" mapped by census tracts in New York City that would have been impossible to do without the 5-year (2005–2009) ACS data.[9] This effort was an unusual one for the *Times*—it had not generated a similar map in the past, from census long-form data—and proved "enlightening in a way that maps of simple ethnicity or simple foreign-born [could] not." Call-outs on the map flag the tracts with the largest concentrations of specific foreign-born immigrant groups—for instance, the largest single concentration of Dominican immigrants in a tract in Washington Heights, Manhattan, and the single most diverse tract in the city (at least 100 people from nine major foreign-born groups) in Queens Village, Queens. He acknowledged that the map is, visually, very busy, but that the *Times* staff concluded that "we were on pretty good ground" presenting the figures in this way. He further acknowledged suggestions from Joseph Salvo of the New York City Department of City Planning (see Section 7–A) that they might consider using a more aggregate level of geography to overcome some of the volatility. But, for now, they enjoy the insights obtained by pushing the data as far as possible.

As an example of detailed analysis contributing to background reporting (and more generalized statements in narrative stories), Fessenden pointed to a package of graphs and a data-animation video that uses ACS data to describe the

but a variant—that constructs a map after the user types an income level to compare with the rest of the nation—is located at http://www.nytimes.com/interactive/2012/01/15/business/one-percent-map.html?ref=business.

[9]The referenced page is http://www.nytimes.com/interactive/2011/01/23/nyregion/20110123-nyc-ethnic-neighborhoods-map.html. A series of submaps, under the main tract-map showing the ACS data, compares change in size for various specific immigrant groups (e.g., East Asians, Carribbean) based on the 2000 census long-form sample compared with the 2005–2009 ACS estimates.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

changing demographics along the course of the New York City Marathon.[10] It draws particular contrasts between 2009 ACS and 1980 census long-form data (the premise of the story being that the route has not changed much since it was first adjusted to trek through all five boroughs in 1976, but the demographics have). The difference between 1980 and 2009 median household income is plotted along the marathon's course, with high peaks in areas where median income has grown the most (e.g., Long Island City) and brief dips below ground level in areas that have become worse off (e.g., Mott Haven, in the Bronx). Similar above- and below-ground graphs along the marathon course are used to examine changes in foreign-born population and particular racial groups. Fessenden briefly mentioned that he also made use of ACS data in background pieces and maps for a series of stories in 2011 to commemorate the bicentennial of Manhattan's street grid layout—again reconciling the current ACS numbers with historical census data and emphasizing the massive demographic changes in parts of the city.

Fessenden closed by briefly displaying another tract-level map using 5-year (2005–2009) ACS data to describe housing stability in the city; tracts are shaded based on the median year in which householders moved into the area (emphasizing in-moves since 2000 and collapsing all arrivals before 1990 into a single shading category). In this case, a reporter was dispatched to a particular tract in East Elmhurst, Queens, because that tract showed up as the single most stable tract in the city—an average tenure of greater than 30 years. Fessenden rued that this was a case "when the desk gets a little out ahead of you" because the map suggests the tract might be a bit of an outlier; neighboring tracts seem to have considerably more recent housing arrivals, on average. But the reporting did bear out remarkable stability—"and that is our other check," he said by way of conclusion. Data like the ACS can support the reporting, and vice versa; Fessenden said that he and the *The New York Times* value the ACS because "there is so much more good in there than there is danger or negativity that we want it in the paper."

4–D DISCUSSION

Much of the closing discussion session for the first day of the workshop involved El Nasser and Fessenden being asked for additional comment on the presentation of margins of error.[11] Terri Ann Lowenthal (see Section 7–D) asked the presenters to comment on Campbell's assertion that the ACS promises more than it delivers with respect to the reliability of tract-level data—whether the journalists believe that the increased timeliness of ACS data (not being up to 10

[10]The marathon graphic is located at http://www.nytimes.com/interactive/2011/11/05/nyregion/the-evolving-neighborhoods-along-the-marathon.html.

[11]Campbell had to leave the workshop immediately after his presentation in order to be able to participate in another conference.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

years old, like the previous long-form samples) offset the increased margin of error. She followed up by asking whether the journalists thought that there was anything the Census Bureau could do to help address the point suggested by all three speakers that reporters cannot or will not present (and so might not fully understand) margins of error, and what they would do if ACS data were not available at all.

El Nasser replied first, agreeing with Lowenthal's premises; she and her colleagues understand the problems and complications associated with ACS data, but concluded that Fessenden put it best with his basic conclusion that "the good outweighs the bad." She said that she understood Campbell's concerns and criticisms—cutting the data by small area *and* by specific job types "slices it down to such a small level" that the analysis is decidedly complicated. But she said that for she and *USA TODAY*, the great value of the ACS is that "it is news"—providing a wealth of data on such a regular schedule. She conceded that *USA TODAY* might not push down to the neighborhood level as much as *The New York Times* or a local newspaper like *The Orange County Register*, but the ACS's capacity for analysis by metropolitan area and county is "quite amazing." She said that she does not necessarily agree with the argument that the ACS promised more than it delivers because "I think we always knew what the problems would be; my expectations were not that much greater." Clearly one could wish for more and more precision, but the current ACS seems to be a reasonable compromise.

Fessenden agreed with El Nasser's points, particularly the assertion that the ACS "is news." More than that, "it is the news business; it is what we like." Recalling the earlier question from a Census Bureau staffer on the practical difference that might come from ACS releases in June rather than September (Section 2–F), Fessenden said that the answer to that question is "an easy one for me"—"the fact that we have this data in much-more-like-real time is really important, and the only thing I would ask for is more." He said that the availability of the ACS has helped create a culture at *The New York Times* that data can be used to say things about what is going on in the city, and what is going on in policy, in ways that were simply unknown before. Just the fact that the ACS is releasing data on a regular schedule ensures that there will be new stories using the ACS all the time—the newer and fresher the data, the better. That said, he noted that he and his colleagues learn to live with a necessary problem: They very much like pushing the data into very small-area analysis but understand that this necessitates working with the 5-year numbers. The problem is that combining so many years means missing the ability to answer some questions (or possibly getting some misleading readings) because the newest 5-year numbers meld pre- and post-recession numbers.

Returning to the question on margins of error, James Treat (ACS Office division chief, Census Bureau) commented that the products from the 2000 census long-form sample did not publish margins of error with its estimates, yet

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

the Census Bureau decided to make the margins of error much more "in your face" in ACS products; they are included in every table. He wondered whether margins of error were an issue in working with the 2000 data or whether error has become an issue for media users simply because they are so much more prominent in the ACS products. Fessenden answered that he thought that the different style of presentation in the ACS products is a big part of the concern; he reiterated that he and his colleagues have to be mindful of the margins of error in deciding what to print and how to format finished products. Patricia Becker (APB Associates) added, and clarified, that there were margins of error for long-form-sample estimates—that the number of cases for geographic areas were "buried in the back" of the products so that the standard errors could be calculated, with a lot of digging. She commented that "there were errors in the long-form data, particularly for small population groups or small geographic areas, and they are every bit as bad as [for] the areas in the ACS. Most reporters didn't know about them; demographers did." She suggested for the record that the margins of error on the long-form estimates in Campbell's analysis might be much more favorable to the ACS.

Campbell Gibson (U.S. Census Bureau, retired) commented that it is good that the ACS makes the margins of error prominent in its products, but suggested that two points were missing from the discussion about comparing the ACS to past decennial censuses. One is that—as he understood it—the original sample size for the ACS was supposed to be about 3 percent of households per year, in which case a 5-year collection would more closely approximate the census long-form sample. But, for budgetary reasons, the ACS sample is about half that level. The other point (and, partially, a modest defense for suppressing the standard errors on the long-form products) is that the long-form samples yielded estimates only once every 10 years—a time span so long that changes in variables like percentage of foreign-born in an area could have changed so markedly that the question of whether the difference was significant from census to census was "kind of passé." But the annual availability of estimates from the ACS naturally creates the temptation to compare data from one year to the next and—particularly for small areas—changes in some variables might be much more subtle.

Andrew Beveridge (Social Explorer, Inc; see Section 7–C) said that he wanted to add to this discussion his view on the ACS products' presentation of margins of error in cases where there are very small numbers or proportions for a particular group or area. In those instances, the Census Bureau's approach has been to publish margins of error that allow for *negative* counts or proportions. He suggested that the general ACS data products might be improved if the Bureau implemented the methodology it used in generating a special tabulation of languages spoken in households to determine requirements for alternate-language voting materials under the Voting Rights Act (see Chapter 7). In that work, the Bureau produced Bayesian estimates that directly borrow strength

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

from estimates in nearby areas and tracts; the alternative of presenting hundreds or thousands of estimates akin to "0 plus-or-minus 133" seems absurd. Roderick Little (associate director for research and methodology, Census Bureau) closed the day's discussion by agreeing with Beveridge that products making fuller use of modeling is a good direction—and added that, as the son of a newspaper editor, he found the session on media perspectives fascinating and that it is the start of a conversation on the distinction between simply "producing facts" relative to "providing useful evidence" for people.

Copyright © National Academy of Sciences. All rights reserved.

– 5 –

State, Local, Tribal, and Urban/Rural Uses of ACS Data

Tasked, as the workshop was, to focus on nonfederal uses of American Community Survey (ACS) data, many of the speakers in several of the workshop sessions touched on issues involved in using ACS data to inform state and local government. Users at the state, local, and tribal levels—whether working directly for governmental units and affiliated agencies or participating in the Census Bureau's State Data Center network to disseminate data broadly—are a sufficiently large and important sector of the ACS user base to warrant a dedicated workshop session. Included in this mix are users who focus on the general issues of rural policy, an area of study for which the ACS presents special benefits and burdens: While those interested in urban policy and large populated areas face the "problem" of handling a wealth of ACS estimates of different vintages (1-, 3-, and 5-year estimates all covering large population groups), workers in rural policy must rely exclusively on the 5-year estimates.

Section 5–A summarizes work using the ACS to examine ancestry, migration, education, and—particularly—racial and ethnic diversity for the Minnesota state government. The contrast in rural and urban uses of the data is spotlighted in Sections 5–B and 5–C, the first of which describes work by the Rural Policy Research Institute and the second (complementing the applications by *The New York Times* discussed in Section 4–C) illustrating the capability to map ACS findings for small-area neighborhoods in New York City and State. Section 5–D discusses the usability of ACS data to describe the unique geographic and demographic characteristics of individual chapters within the Navajo Nation. The

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

workshop steering committee asked Jacqueline Byers of the National Association of Counties to serve as leader for the discussion of these presentations; her own comments from the perspective of county government users, and the general discussion, are summarized in Section 5–E.

5–A STATE GOVERNMENT USES: HIGHLIGHTING DIVERSITY AND INFORMING POLICY IN MINNESOTA

On a continuous, daily basis, the Minnesota State Demographic Center (SDC) fields questions and works on projects that depend critically on the ACS. These include questions from other state agencies,[1] nonprofit organizations, media outlets, or the general public—whether through the center's dedicated email "helpline" or other means. Not wanting to minimize those smaller, frequent requests for information, Susan Brower (Minnesota state demographer) began her comments by saying that she would focus on a few larger, longer-term projects to suggest the substantive range of the center's work with the ACS to inform state policy.

She started by describing recent work making use of the ACS's ability to zero in on trends in very specific population groups. Over the past four decades, Minnesota has become home to increasing numbers of Asian and Pacific Islanders, in large part reflecting arrivals of refugees fleeing repressive governments in southeast Asia. As Brower described, the mix of particular Asian Pacific ethnic groups in Minnesota is considerably different from that of the entire U.S. population, based on her analysis of 2008–2010 ACS data (via the Integrated Public Use Microdata Series [IPUMS]). In the general U.S. population, the proportions of southeast Asian groups (e.g., Cambodian, Hmong, Vietnamese) and east Asian groups (e.g., Chinese, Japanese, Korean) are roughly on par with each other, with east Asian groups holding a slight plurality and a somewhat smaller percentage of south Asian groups (e.g., Asian Indian, Tibetan, Bangladeshi). By comparison, southeast Asian groups make up a clear majority of Minnesota's Asian population, with a markedly smaller share of east Asian groups and slightly fewer from south Asian groups. In specific, the largest single Asian Pacific ethnic group in Minnesota is the Hmong population; the state's approximately 3,000 Tibetans make Minnesota rank second among U.S. jurisdictions in Tibetan-Americans; and the state's concentration of Karen (a Burmese–Thai minority group) is believed to be the largest outside of southeast Asia (Council on Asian Pacific Minnesotans, 2012:i, 4).

In spring 2012, the Minnesota SDC collaborated with the state's Council on Asian Pacific Minnesotans—a state agency tasked to report to the governor and legislature on issues important to Minnesota's Asian community—on a broad

[1]The Minnesota State Demographic Center is administratively housed in the state Department of Administration.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

report on the demographic and economic condition of Minnesota's Asian Pacific ethnic groups. The work resulted in an April 2012 report, *State of the Asian Pacific Minnesotans* (Council on Asian Pacific Minnesotans, 2012), that made extensive use of ACS data as well as counts from the 2010 census. Brower said that her center's role in this report was to perform the analysis and (examining the standard errors on the estimates) demonstrate to the Council that the results were reliable and stable.

For the report, Brower chose to perform special tabulations from the IPUMS 3-year microdata (2008–2010) from the ACS, profiling the nativity, ancestry, and socioeconomic and housing conditions of the state's major Asian subpopulations. She said that she found it necessary to build the special tabulations, rather than use the detailed tables (which she made certain to describe as "easily accessible" from the Census Bureau's American FactFinder interface), to provide reassurance that the analysis referenced *all* of the Asian population. In particular, she said that she and the Council wanted to respect the very different cultural and historical backgrounds of the different Asian Pacific subpopulations and, therefore, carefully address each subpopulation separately. She said that the American FactFinder tables tend to describe results for the Asian population overall, and she and the Council needed finer resolution than that for this project. To illustrate the point, Brower noted that the two largest segments of Minnesota's Asian population—Hmong (26 percent) and Indian (15 percent)—are "at very different ends of the socioeconomic spectrum" and in background in coming to Minnesota. The major influx of the Asian Indian population occurred recently (over the past 10 years or so) and are more likely to be fluent in English.

Analysis of the ACS data suggests great diversity by specific ethnicity group in the broader Asian-ancestry population in Minnesota. Looking at "all Asian" ancestry people in Minnesota in the 2008–2010 ACS data, 21 percent hold a bachelor's degree or higher. But that figure masks very different educational attainment by specific groups; the majority Hmong population includes only 14 percent with a bachelor's degree or higher; Asian Indians are second largest in size in Minnesota's population but tend to be better educated, leading other Asian subpopulations at 82 percent with a bachelor's degree or higher. The state's Chinese and Korean subpopulations count over 50 percent with at least an undergraduate degree; only 9 percent of the state's Cambodians and 6 percent of Laotians have a degree. Additional interesting results were generated by looking at nativity, comparing the ACS data with the 1990 and 2000 census long-form samples. Brower found that the percent of native-born Asian Minnesotans with at least a bachelor's degree has steadily decreased over the decades—to the extent that, as of the 2008–2010 ACS data, foreign-born Asians in Minnesota were more likely to have an undergraduate degree than native-born Asians.

Brower briefly described two other recent projects in which the Minnesota SDC used ACS data to inform specific policy uses. In April 2012, the Min-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

nesota SDC performed analysis of 5-year (2006–2010) ACS PUMS data at the request of the Minnesota Department of Human Rights. Specifically, the department wanted to focus on equal employment opportunity (EEO) compliance for a very precise economic sector—the construction industry—to see if state-mandated hiring goals were being met on construction projects funded by the state. Recalling the earlier question from Census Bureau staff on the practical import of a few weeks or months in the timing of ACS data releases (Section 2–F), Brower noted that the timing does matter—and that this EEO work was one example where the ACS's current release schedule gives it a unique advantage over other products. She said that the Census Bureau plans to release special tabulations for EEO compliance at the end of calendar year 2012, and that these special tabulations will facilitate much more detailed analysis than SDC's PUMS work. However, waiting until the end of 2012 was not a viable option—the state human rights commissioner faced a deadline to get the data on hiring goals during the state legislative session because 2012 is a revenue bonding year. Major capital projects were in line for the session, and the state wanted to make sure that hiring goals were in line with expectations. SDC plans to repeat the analysis when the special tabulations become available later this year, to permit the human rights department to update its goals.

Just as ACS data are used at the federal level to allocate federal funds, so too are they used by state governments to distribute funds to localities. Brower recounted that, in May 2012, SDC helped the Minnesota Department of Employment and Economic Development score applications under the Small Cities Development Grants Program. These grants—made from money allocated by the U.S. Department of Housing and Urban Development—are intended to help small, rural communities fund infrastructure and housing improvements. The grant monies are distributed according to a formula based on the percentage of persons in poverty at the city place level; SDC compiled data from the 5-year 2006–2010 ACS data to use in the formula.

Brower said that SDC is currently working on a longer-term planning project, expected to culminate in a themed version of the center's annual report to the governor and legislature, tentatively titled *The Time for Talent*. The basic goal of the report is to make use of the ACS's detail on educational attainment and occupation to document the levels of talented workers in the state, and to use its data on place of birth to get a read on the movement of talented workers in and out of the state. For this work, Brower said that the SDC took the slightly unusual step of looking at four individual-year ACS PUMS files; cognizant of the economic recession, they were concerned that some interesting trends might be masked by looking at 3-year or 5-year data as a whole. She said that the numbers that they obtained from the 1-year files were seemingly too small to be stable, but they determined that the average of those single years seemed to perform better. She emphasized that the work was still under way at the time of the workshop, and that additional analysis would attempt to use

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Table 5-1 Number of Movers to Minnesota with at Least a Bachelor's Degree by State of Origin (or Abroad), 1995–2000 and 2007–2010 (annual averages)

State	1995–2000[a]	2007–2010[b]
Wisconsin	16,207	19,764
Abroad	14,366	17,296
North Dakota	8,191	6,360
Florida	1,930	5,800
California	6,760	5,680
Illinois	7,936	5,526
Iowa	8,169	5,069
Texas	3,426	3,718
South Dakota	3,843	3,341
Colorado	2,574	2,996

[a] Person Question 15a on the 2000 census long-form questionnaire asked "Did this person live in this house or apartment 5 years ago (on April 1, 1995)?" and included "No, outside the United States" as a response. Question 15b asked "Where did this person live 5 years ago?" and permitted specification of a city, county, state, and ZIP Code.

[b] Person Question 15a on the 2012 ACS questionnaire (with similar versions on earlier questionnaires) asks "Did this person live in this house or apartment 1 year ago?" and includes "No, outside the United States and Puerto Rico" as a response. Question 15b asks "Where did this person live 1 year ago?" and permits specification of an address, city, county, state, and ZIP Code.

SOURCES: Tabulation from 2000 Census data and Integrated Public Use Microdata Series, 2007–2010; adapted from workshop presentation by Susan Brower.

the ACS data to zero in on the so-called STEM fields and occupations (science, technology, engineering, and mathematics), on entrepreneurs, and on creative talent.

Some illustrative high-level results from the work are shown in Tables 5-1 and 5-2. Table 5-1 suggests that Minnesota continues to draw large numbers of workers with a bachelor's degree or higher from the neighboring state of Wisconsin—and nearly as many from outside the United States. Taking a longer-term, life-course look by examining educational attainment and state of birth, Table 5-2 suggests that Minnesota ranks fairly high in retaining degreed workers who were born in the state; only the country-sized economies of Texas and California, as well as North Carolina, have higher "staying power."

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Table 5-2 Percentage of Degreed Workers Living in State of Birth, 2007–2010

State	Percent
Texas	69.7
California	65.7
North Carolina	61.0
Minnesota	*59.1*
Georgia	58.7
Utah	57.4
Washington	56.0
Tennessee	55.4
South Carolina	55.1
Wisconsin	54.1

SOURCES: Tabulation from single-year Integrated Public Use Microdata Series, 2007–2010; adapted from workshop presentation by Susan Brower.

She concluded by reiterating that Minnesota's state government makes heavy use of ACS data for three principal purposes: developing population profiles of Minnesota-specific groups, conducting research to support development of state policy and distribution of funds, and crafting special reports to help with long-term planning. She also commented that, shortly before the workshop, the SDC hosted a new forum for boosting public awareness of ACS data: its first "Minnesota Data Opener" (a takeoff on the opening of fishing season, which began at roughly the same time), in partnership with Minnesota Public Radio. The data opener took the form of an online chat. SDC published a "workbook" of data on education (derived principally from the ACS) in advance, and invited the public to look at the data themselves and comment on what they found most interesting. In particular, the "data opener" included a contest, asking people to produce some kind of visualization from the data. They were very encouraged by the results—it was an "engagement project using data," and the results were numerous, interesting ways of presenting data for the SDC to consider moving forward.

5–B PLANNING HUMAN SERVICES IN RURAL AMERICA

Kathleen Miller—program director of the Rural Policy Research Institute (RUPRI) at the University of Missouri—described RUPRI's recently completed Geography of Need project, an effort to compare the needs for various human

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

services in rural America compared to urban areas. RUPRI undertook the Geography of Need project in collaboration with Colleen Heflin (also from the University of Missouri) and with funding from the Kresge Foundation. The work grew out of Kresge's interest in recent years in redesigning their human services portfolio. Particularly in light of the economic recession and funding pressures created by federal and state budget deficits, Kresge tasked RUPRI to develop a national overview of how needs for services vary by geography and urbanization, and what types of needs for services occur in different combinations in different places. That regions of the country—many of them poor—are characterized by persistent poverty is well known, but the way in which needs for services vary within those areas is much less clear.

As RUPRI set about conceptualizing and planning this study, Miller said they did so by following three basic rationales:

- *Human service needs depend on the characteristics of the population in need:* For instance, military veterans have unique needs to services, and also have existing access to an array of services not open to other populations (e.g., U.S. Department of Veterans Affairs health centers). Needs for human services can also vary across the life course—the elderly have very different needs than do young children—and in combination with geography or other factors (e.g., a town whose composition changes due to either a large outmigration of young people or a rapid inmigration of retirees relative to one with very young populations).
- *Specific local conditions affect the demand for and provision of human services:* These local conditions could include cultural differences or language barriers (linguistic isolation). In rural settings, transportation barriers are of particular concern. As Miller put it, "all the services in the world are not going to help if you can't get to them"; if someone lives in a rural area an hour or more away from any service—"everything from a gallon of milk to your pediatrician"—and their car breaks down, they are stuck.
- *Human service needs are also a function of the economic structure of an area:* Again, as mentioned above, some rural areas are characterized by persistent poverty and limited employment opportunities, and that will affect the range of needed services.

Starting work, Miller and RUPRI faced the task of deciding what level of geographic resolution to pursue in mapping the landscape of need. State-level analysis was simple to rule out as being too coarse, but going down to finer levels of geography like Public Use Microdata Areas (PUMAs) or tracts raised issues of "poor data and smaller sample size." The basic task to construct a national-level portrait put additional constraints on the problem: The researchers had to think in terms of a set of indicators that could be calculated for *all* units at the chosen resolution (and not just cities in metropolitan areas), and the idea of

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

simply dropping smaller geographic units from consideration ran counter to the project's purpose.

Ultimately, RUPRI settled on counties as the appropriate unit—a choice with certain undeniable benefits. The number of county-level equivalents in the United States (3,141) is large but not overwhelming, and their boundaries are generally well understood and stable over time. More pertinent, counties are also the natural operational geography for many human services providers (e.g., county health and welfare agencies), so counties seemed to mesh well with the project's goals. To be sure, though, counties also raised some disadvantages because they range widely in both population and land area, their size along either dimension can obscure important urban/rural distinctions within county borders, and whole counties might be less likely to stand out with respect to some needs/risk criteria than a purely rural community.

Cinching the choice of county as the level of resolution was what Miller called a "fortunate" bit of timing—active work on the Geography of Need study began around December 2010, coinciding with the release of the first 5-year ACS estimates. The release of the 2005–2009 5-year ACS estimates was the first exposure to ACS data and first update of socioeconomic data since the 2000 census for about 40 percent of counties—and, for purposes of rural data analysis, "this was early Christmas."

The approach taken by Miller and RUPRI was to construct a data set of 12 demographic and economic indicators for all counties. While most of these indicators could be calculated directly from the 5-year ACS data, RUPRI favored data resources from the U.S. Bureau of Economic Analysis and the U.S. Department of Veterans Affairs over the self-report data in the ACS, on income from government transfer payments and veterans status, respectively. RUPRI also turned to the Census Bureau's annual population estimates program to derive the population denominators used in its calculations and to its annual model-based Small Area Income and Poverty Estimates for the county poverty rate. RUPRI's final set of county-level indicators included:

- Percent of the population age 65 and over;
- Working-age dependency ratio;[2]
- Fertility rate;
- Foreign-born population;
- Percent of population in racial or ethnic minority groups;
- Percent of population living in subfamily arrangements;[3]

[2]Segmenting the population into "working-age" (e.g., from ages 15–64) and "dependents" (the complement, e.g., children under age 15 and seniors over age 64), the working age dependency ratio is simply the ratio of dependents to working-age.

[3]Subfamilies are families that live in the household of some other person—for instance, an unmarried mother living with her children in the house of her parents or a married couple (with or

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

- Percent of population with less than a high school education;
- Veterans as a percent of total population;
- Percent of households receiving Supplemental Nutrition Assistance Program (SNAP/Food Stamp) benefits;[4]
- (County) poverty rate;
- Percent of households without a vehicle available; and
- Share of personal income from government transfer payments.

Having constructed a set of indicators, RUPRI derived a very simple index measure: Each indicator was separately sorted and ranked, and a county was assigned a 1-point risk factor if it fell in the bottom 10 percent (and 0 otherwise). Summing the scores across counties produced a Human Services Need Index score between 0 and 12; these scores are mapped in Figure 5-1. Miller said that the red-shaded areas (those with more than 3 or even up to 9) on the map highlight some areas where RUPRI expected to see hardship high levels of human services need: Appalachia, the "Black Belt" of the South, the Mississippi Delta, the Texas–Mexico border region, the California Central Valley, and American Indian lands in the Great Plains and the Four Corners. What also struck RUPRI staff was the vast swath of the country with 0 risk factors and the remote nature of many of the highest-scoring counties, suggesting the possibility of significant transportation barriers and difficulty in service delivery.

RUPRI's full analysis of these data tries to probe what is behind the red-shaded areas on the map—what combinations of the 12 indicators/risk factors seem to show up with greatest frequency, and where. Miller said that it was not surprising to see one frequent combination—low educational attainment and high poverty. But, for some parts of the country, a less-expected combination (high elderly population and veterans' status) proved most prevalent.

Table 5-3 breaks down the index scores by whether the county is or is not part of a metropolitan area. Strikingly, over half of metropolitan counties exhibit 0 risk factors, compared to only about one-third among nonmetropolitan counties. On the high end of the scale—between 3 and 9 risk factors—the difference is even more stark, covering only about 9 percent of metropolitan counties but just over one-quarter of nonmetropolitan counties. Subsequent analysis suggested that noncore counties (those containing no urban cluster with at least 10,000 residents) are "very overrepresented" among the nonmetropolitan counties with high risk factors.

---

without children) living with some other relative. The concept is described in greater detail on the IPUMS project website at http://usa.ipums.org/usa/volii/subfamilies.shtml.

[4]Housing Question 12 on the 2012 version of the ACS questionnaire asks "IN THE PAST 12 MONTHS, did anyone in this household receive Food Stamps or a Food Stamp benefit card? *Include government benefits from the Supplemental Nutrition Assistance (SNAP). Do NOT include WIC or the National School Lunch Program.*" Only a yes/no checkbox response is permitted.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

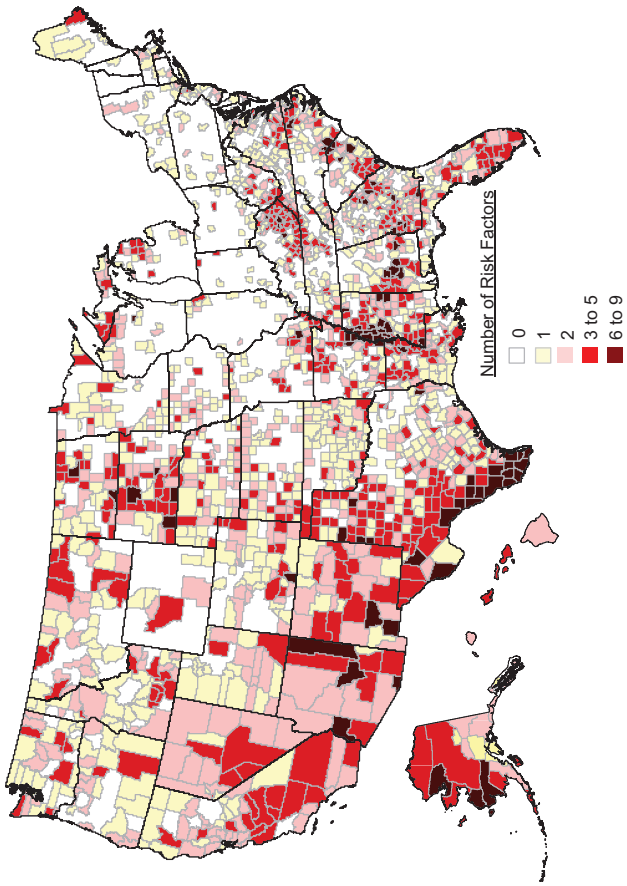


Figure 5-1 Calculated Human Services Need Index for U.S. counties, 2009

SOURCES: Calculated using 2009 Population Estimates, 2005–2009 American Community Survey, and 2009 SAIPE estimates (U.S. Census Bureau) and data from the U.S. Department of Veterans Affairs and the U.S. Bureau of Economic Analysis Regional Economic Information System. Analysis and mapping by Rural Policy Research Institute. Adapted from workshop presentation by Kathleen Miller.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Table 5-3 Rural/Urban Differences in Human Services Need Index, 2009

Number of Risk Factors	Metropolitan Area Counties	Non-Metropolitan Area Counties
0	585 (53.2%)	690 (33.8%)
1	267 (24.3%)	464 (22.7%)
2	149 (13.5%)	346 (16.9%)
3–5	91 (8.3%)	459 (22.5%)
6–9	8 (0.7%)	84 (4.1%)

SOURCE: Calculations by Rural Policy Research Institute, University of Missouri; adapted from workshop presentation by Kathleen Miller.

Wrapping up, Miller commented that the ACS is hardly perfect for studying and understanding rural America. Revisiting the point that a large fraction of U.S. counties fall below the 20,000 population cutoff—and so have to wait for 5-year estimates from the ACS—she conceded that there are grounds to be concerned about the sample size for rural counties, about large margins of error, and less timely estimates (only two nonoverlapping 5-year periods per decade for the purest indication of change over time). But—in short—the ACS "is the best data and often the only data we have" for rural analysis. RUPRI's particular analysis of human services risk factors would not have been possible without the 5-year estimates; the work has helped RUPRI develop hypotheses for further research and helped the sponsoring Kresge Foundation retool its portfolio of grants dealing with human services.

5–C MAPPING ACS DETAIL IN NEW YORK CITY

Shifting from ACS analysis of intensely rural areas to the survey's use in explaining trends in the nation's largest city, Steven Romalewski described recent work in the mapping of ACS estimates and—in particular—first steps in reflecting margins of error in the plots. Romalewski directs the Mapping Service in the City University of New York's (CUNY) Center for Urban Research, which uses geographic information systems (GIS) software to design a variety of online applications using Census Bureau data. Of particular note, and in support of the 2010 census, Romalewski helped develop an easy-to-use mapping application for the Census Bureau's tract-level planning database—a tabulation of 2000 census long-form samples used to calculate an index score (not unlike Miller's additive Human Services Risk Index in the previous section) that indicated the "hard-to-count" nature of each tract. Other applications developed by the CUNY Mapping Service made use of the long-form-sample data from the

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

2000 and earlier censuses (which, he observed, were "maybe not as reliable as everyone thought at the small-area level").

Like Miller's RUPRI, Romalewski's Center for Urban Research and Mapping Service is housed in an academic setting, but works with and for a variety of constituencies with very practical data needs: local organizations, state and city agencies, and other general researchers. Part of the problem the Mapping Service faces in moving more fully into ACS-based products is meeting the expectations that have developed in their partners' (and clients') minds. The Mapping Service has become known for online interactive mapping that allows users to zoom in very closely; their clients really need, and expect to be able to, drill down to fine-grained tract or block group data.

Accordingly, by way of providing initial impressions of the ACS data, Romalewski offered two notes of empathy with Miller's concerns over ACS use for data analysis. First, Romalewski said that he and his colleagues try to adhere to the counsel of the Census Bureau and the National Research Council (2007b) to refrain from comparing estimates for overlapping time intervals; given that the fine levels of geography of greatest pertinence to CUNY's users are only available in the 5-year products, Romalewski also feels the frustration of twice-a-decade comparisons. (Like Miller, he was quick to add that twice-a-decade is far preferable to the alternative.) And, second, Romalewski shared the concern that small counts in individual city tracts and block groups can produce "unacceptably high margins of error."

With those comments as prelude, Romalewski said that—with margins of error now much more prominent in the ACS estimates—he and his colleagues are grappling with two fundamental questions as they update their applications and try to most effectively (yet accurately) present ACS small-area data. Those questions are:

- *Is geographic aggregation necessary?* That is, can we have confidence in visualizing spatial patterns by plotting tract (or even block group) data, or is selective aggregation of some small areas necessary to get acceptable precision?
- *Should (or must) "high margin-of-error" areas be denoted or visually flagged in some way*, as through cross-hatching or other means? Put another way, is it possible to create robust choropleth maps even with uncertain data?

He hastened to add that "I don't have any conclusions" as yet; working through these issues is an ongoing exercise. In that spirit, he spent the remainder of his presentation briefly displaying a mix of applications and first-cut approaches: applications where thorny presentation issues remain to be made in converting from long-form sources to the ACS and initial attempts to flag potential problem areas in static maps.

As examples of his center's ongoing mapping projects, he began by showing screenshots from the Long Island Index project, an interactive map server

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

located at http://longislandindexmaps.com. At that site, interested users can generate overlays from a variety of sources, including data from county planning authorities and the New York and U.S. Departments of Education. The application currently draws population and housing data from the 1990, 2000, and 2010 census summary files, and generates maps down to the block group level. The specific screenshot he chose to display zoomed in on central Long Island (Suffolk County) and displayed the percentage of occupied housing units with no vehicles available. He noted that a number of organizations on Long Island are keenly interested in this specific data item in order to understand access to food sources and for planning transportation routes. Romalewski would return to this specific view—households without cars in central Long Island—to illustrate the Mapping Center's concerns about presenting ACS data. He then briefly displayed a screenshot of a block-level map of the New York metropolitan area map, with each block shaded according to the basic race category making up the plurality of the block's total population. This map server, accessible at www.urbanresearchmaps.org/comparinator/pluralitymap.htm, permits users to look at data for 14 other large metropolitan areas in addition to New York. The specific view Romalewski chose included a vertical slide bar that the user can manipulate; 2000 census data appear to the left of the line and 2010 data to the right.[5] Similar views could be generated from ACS data, albeit not as fine as the block-level counts.

Maps are an especially powerful way of making data interpretable to the broader public, and Romalewski said that this power (and responsibility) motivates the Mapping Center's concern in trying to accurately convey levels of uncertainty in mapped data. He mentioned that he teaches an introductory GIS course for graduate students in the Pratt Institute's city and regional planning program and is struck that even in that setting of fairly sophisticated users, it is common to hear comments along the lines of "The map said *this*." When data are presented in map form, people more readily assume that it is the truth—even though the margins of error underlying the tints of areas on the map might undercut that level of certainty. Romalewski said that he and his colleagues are working with and learning from numerous sources as they explore different cartographic techniques for conveying uncertainty, including David Wong and colleagues at George Mason University (work on software extensions for the ArcGIS package)[6] and the Cornell University Program on Applied Demographics. More recently, an informal working group of interested parties formed in New York City—including the CUNY Mapping Center, the city's Department

[5]Similar formatting is used in an online application portraying legislative districts in New York state, with block-level counts and district boundaries overlaid on each other—pre-2012-redistricting on the left and post-redistricting on the right. This project is located at www.urbanresearchmaps.org/nyredistricting/map.html.

[6]Wong contributed a short methodological overview of his work to overlay coefficients of variation on ACS-based maps to the workshop's case study/agenda book.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

of City Planning, and Ford Fessenden and colleagues at *The New York Times* (see Section 4–C).

Again, Romalewski emphasized, the problem is very much open and he has not yet reached any conclusions about preferred approaches; he moved swiftly through a series of approaches, with their attendant pros and cons. One approach is to visually mask areas with high levels of uncertainty:

- For a map of Asians and Pacific Islanders as a percentage of total population at the census tract level within New York City—2000 census data in one pane and 5-year 2005–2009 ACS estimates in the other—Romalewski sought to flag "tracts with unreliable estimates." Different percentages of the Asian population were portrayed in shades of green (or a light yellow tint for 0–10 percent Asian population) but—for the 5-year ACS pane—tracts with a coefficient of variation of 25 percent or more were shaded grey. Major trends register strongly in the maps—the extreme growth of the Asian population in parts of Queens (see Sections 4–C and 7–A) and historic concentrations in lower Manhattan and parts of Brooklyn. But the "problem" is that there are many such "unreliable" tracts—the ACS-based map is overwhelmingly grey; the map avoids the false precision of the 2000 census pane where many of the ACS-grey tracts show up in the yellow-tinted 0–10 percent Asian category, but the presentation of so many "unreliable estimate" tracts might lead viewers to question the reliability of the data as a whole.
- For a map of tract-level percentages of population living in poverty (income under 150 percent of the poverty line) in the Chicago metropolitan region,[7] Romalewski demonstrated a more overt "mask." The poverty levels were mapped in red-orange shades, and he displayed a version with and without pale blue dots marking tracts with coefficients of variation of 25 percent or more. Again, the attempt at an honest representation of the variability inherent in the estimates winds up calling the reliability of the whole map into question because there are so many blue dots. Perhaps worse, the blue dots very literally mask the picture of poverty where census tracts are smaller and most dense: The dots make it physically difficult to see the shades of poverty in the city of Chicago itself.

Another approach is to confront the margins of error directly—to plot the lower and upper bounds of a confidence interval estimate for a location rather than a single point estimate. To illustrate the point, Romalewski revisited the view of block groups in Suffolk County, shaded based on the percentage of housing units with no access to a vehicle, using 5-year 2006–2010 ACS estimates.

[7]The maps were clipped to make a vertical, portrait-orientation presentation and so focus on the core of the Chicago area in Illinois; the map coverages spill over into northern Wisconsin and northwest Indiana but the view is truncated, as it is for outer-ring Illinois counties like DeKalb and LaSalle.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

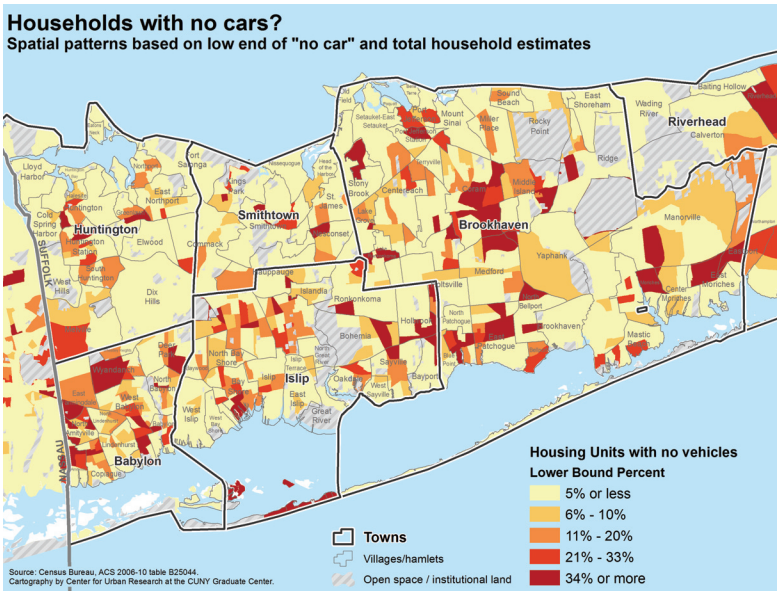


Figure 5-2 Lower bound estimate of percentage of housing units with no vehicle, Long Island, New York, by block group, 2006–2010

SOURCES: Calculated from 2006–2010 American Community Survey, Table B25044; adapted from workshop presentation by Steven Romalewski.

Figure 5-2 shows the map based on the lower bound of an interval estimate for the no-car percentage, subtracting the margin of error from the estimated number of no-car households and dividing by the total number of occupied housing units. He said that this might be considered a "worst case" map for no vehicle access. Most people on Long Island have cars, so low percentages dominate the picture, but Romalewski said that the pattern of block groups with high percentages is fairly similar to the pattern based on 2000 census long-form data. But when the view is switched to the *upper* bound on the estimates (Figure 5-3), the volatility in the data becomes clear; the map is saturated in the red and orange tones connoting the highest levels of no vehicle access because large standard errors drive up the upper bound. It is a "world of difference" between the two maps and—for purposes of envisioning specific transportation routes and policies—it suggests that the ACS data exhibit extreme uncertainty at the fine, block-group level.

Romalewski's next series of maps looked at the same geographic area—Suffolk County, this time looking at census tracts rather than block groups—but a slightly different data source. As discussed in greater detail in Chapter 7,

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

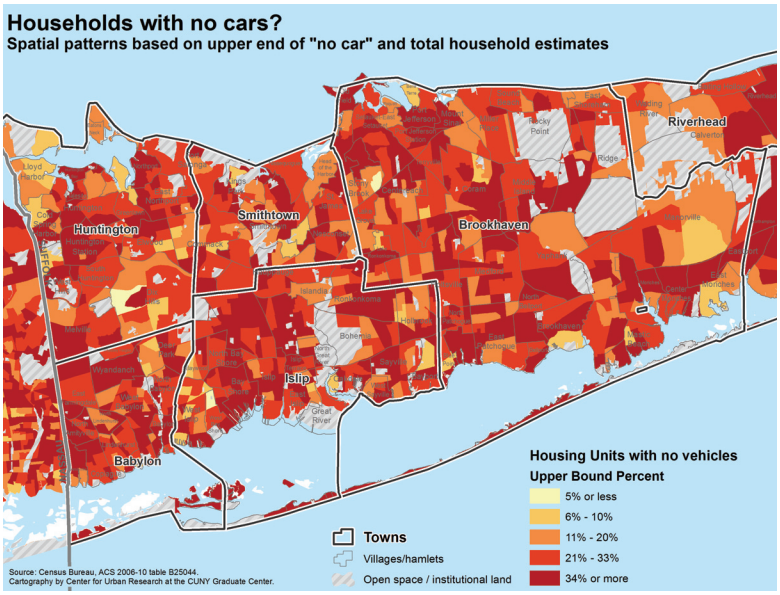


Figure 5-3 Upper bound estimate of percentage of housing units with no vehicle, Long Island, New York, by block group, 2006–2010

SOURCES: Calculated from 2006–2010 American Community Survey, Table B25044; adapted from workshop presentation by Steven Romalewski.

the Census Bureau issued a special tabulation from 5-year 2006–2010 ACS data for use in legislative redistricting and more easily assessing the equity of newly drawn districts. Known as the Voting Age Population by Citizenship and Race tabulation—or CVAP for short—the file contains counts of voting-age U.S. citizens (and their associated margins of error) for geographic units from the nation as a whole down to the census block group level. The maps Romalewski displayed using CVAP data grew out of work that the Mapping Center has been doing with a group involved with boosting civic engagement among the Long Island population; in particular, this set of maps examined the percentage of the total CVAP population in each tract that is of Hispanic origin. Given this application area, one instinctive "answer" to the problem of volatility in the ACS estimates—switching to a more aggregate level of resolution like county subdivisions or defined villages—is not really viable; higher-level geographic aggregates are far from optimal for purposes of planning specific, targeted voter registration drives or activities affiliated with schools.

Using the CVAP data, Romalewski began by repeating the approach of generating separate maps based on lower-bound and upper-bound estimates for per-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

cent Hispanic. The results were similar to, but more muted than, those for the no-car variable by block group; a cluster of high-Hispanic voting-age population in the northwest section of the town of Islip stands out, but numerous tracts flare up to higher levels (and deeper shading on the map) in the upper bound map due to large margins of error. He displayed the results from a few other approaches:

- *"Triangulation"/corroboration with related, auxiliary data:* In this specific setting, he displayed a map that looked at the Hispanic voting-age (18 and older) population from the 2010 census summary data file. This is not the same population because it is not restricted to *citizens*, but it can serve as an intuitive check—or an effective, true upper bound—on the CVAP data. The high-Hispanic-concentration areas in Suffolk County that registered strongly in the CVAP maps show up on the 2010 summary file map—northwest Islip, Brentwood, Huntington. And one could construct a "corrected" map through this triangulation—for instance, an area in East Shoreham shows up in the 11–25 percent Hispanic shading level on the upper-bound CVAP map but that percentage is logically capped at the 6–10 percent category because that is the 2010 census percentage Hispanic-and-voting-age for the tract (without regard to citizen status).
- *Omission of unreliable-estimate areas:* A variant of the masking approach, another take would be to visually omit tracts with highly unreliable estimates (coefficient of variation 25 percent or greater) rather than call attention to them. In practice, this took the form of leaving the unreliable-estimate tracts unshaded (white). Though he repeated that he has made no conclusions, he did suggest that this approach is undesirable because—in general—"you want to try to avoid losing information from the map," and this feels like going out of the way to *not* present the whole picture.
- *Overlay some visual marker on areas where margins of error could shift estimate to a different color category:* In this sample map, Romalewski shaded the Suffolk County tracts in some tint of green based on Hispanics as a percentage of total CVAP population, but overlaid a dot-mesh layer on tracts where the interval estimate could place the estimate in one of the other green-tint ranges. He said that his basic reaction to this picture is that it is "okay to an extent," but it seems to put a very high burden on the map's readers to sort out what the map is really telling them.

He closed by noting that the Cornell Applied Demographics program is doing particular work on techniques to flag potentially unreliable areas in an interactive, online mapping application, focusing the viewer's attention on the possible pitfalls and permitting a more detailed look at the reason for the flags by clicking on them.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

5–D STUDYING DEMOGRAPHIC AND ECONOMIC CONDITIONS IN THE NAVAJO NATION

Using ACS data to craft a demographic and economic portrait of the Navajo Nation is a very different proposition from using it to explore patterns in New York City. The Navajo Nation lands cover a land area in the Four Corners region of the Southwest that is roughly comparable to the state of West Virginia—across which is spread a resident population of roughly the same size as the single neighborhood of Flushing, Queens. However, the ACS is proving just as useful a resource, as Lester Tsosie (from the Navajo Nation's Division of Economic Development) described in his workshop presentation. His workshop slides were actually a direct copy of a briefing that Tsosie's division had prepared for tribal leaders in April 2012 as a briefing on the "economic landscape" of the Navajo Nation, so the core of his remarks at the workshop was actually a pure demonstration of the way in which ACS data are used to address the questions of policy makers.

For the benefit of the workshop presentation, though, he began with some basic overview of structure. The Navajo are one of roughly 550 American Indian tribes in the United States and—in both its membership population and the area of its tribal land holdings—is the largest of those tribes. The tribe maintains its own membership rolls, but respondents to the decennial census and the ACS can self-identify as Navajo in answering the question on race. Per a graph in his slides, the 2010 census reported just over 308,000 self-identified Navajo in the total U.S. population; of these, the majority (about 56 percent) lived "on reservation," within the Navajo Nation's tribal lands. That figure represents a decreased share of Navajo living on tribal land compared to the 2000 census (about 61 percent); Tsosie said that much of this migration from living on tribal land to living elsewhere is principally about Navajo people going to large urban centers for employment (and educational) opportunities.

Functionally, the Navajo Nation is a semiautonomous territory with its own "national" (tribe-level) executive, legislative, and judicial branches. It is divided into 110 chapters that perform some local government functions; these chapters vary greatly in size and population—and 110 chapters partition the on-reservation population of about 173,000 quite finely—but the chapters are a natural and meaningful level of resolution for looking at demographic data.

Tsosie's economic and demographic briefing for tribal leaders centered around four choropleth maps, working with four variables calculated from 5-year 2006–2010 ACS data:

- *Families in poverty:* Per Tsosie's calculations, on the order of 13.8 percent of the total U.S. population has income levels below the poverty line; within the Navajo Nation, that proportion is nearly tripled (37.7 percent). Based on the map, the Division of Economic Development superimposed black circles to draw tribal leaders' attention to 10 particularly prominent

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

clusters of poverty within the Navajo Nation—some single-chapter and geographically compact (e.g., the Alamo and Tóhajiilee chapters in New Mexico) and others more sprawling (e.g., a cluster of roughly 15 chapters centered around the Manuelito and Tsayatoh chapters and spreading across the Arizona–New Mexico border).

- *Percent of civilian labor force unemployed:* Again, viewed as a whole, the Navajo Nation's unemployment rate (estimated for the presentation at 15.6 percent) exceeds—and is nearly double—the national level. For this map, deriving the percent unemployed by chapter from the ACS data, the Division of Economic Development did not superimpose any graphic device to call attention to particular clusters because the areas with the highest unemployment rates (higher than 36 percent) stand out very prominently on the map: a north-south line of New Mexico chapters running from Whiterock to Smith Lake and four chapters in Arizona near the eastern border of the Hopi Indian Reservation, which is geographically an enclave in the main Navajo Nation landmass.
- *Median household income:* Tsosie's department's calculations suggest that median household income in the Navajo Nation ($26,232 per year) is just under half of the median income for the United States as a whole. Mapped by chapter, the higher median-income chapters tend to be on the edges of the Navajo Nation's land area—most strikingly, in a line along the western border of the nation from Lechee and Coppermine in the north to Leupp in the south (or, roughly, from Page, Arizona, to just east of Flagstaff, Arizona). In particular, Tsosie said that his division is using this map (in combination with the unemployment map) to reevaluate the division's strategic development plans, seeing how the maps coincide (or not) with developing commerce centers within the Navajo Nation.
- *Percentage of people whose spoken language is something other than English:* As a crude measure of speakers of the Navajo language, the Division of Economic Development calculated the percentage of each chapter's population that reported speaking a language other than English; presumably, this mainly identifies speakers of the native Navajo language and its dialects. About 70.7 percent of the entire on-reservation Navajo Nation population speak a non-English language, per the presentation's calculations. Though there are pockets of non-English concentrations on the northern and eastern edges of the Navajo Nation's lands, the general impression from the map is a radial pattern—strongest concentration in the chapters in the northeast corner of Arizona and spreading (to lower concentrations) away from that center. This map, and the underlying data, are now being examined by education authorities within the Nation, because it does have implications for the service populations of educational institutions. As an aside, Tsosie also noted that this map also serves as a

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

resource in helping several counties in Arizona, New Mexico, and Utah meet their obligation to provide bilingual voting materials (see Box 7-1 and Chapter 7 generally).

After constructing these maps, Tsosie paid credit to the consistency of the ACS as a data source. The Navajo Nation lands spread across three U.S. states and cover (or otherwise interact with) a host of county and local governments. Touching on so many other jurisdictions, Tsosie said that he knows from experience that it is very difficult for an organization like his Division of Economic Development to coordinate, work with, and obtain information from all those other sources. Information is also not necessarily coded or recorded in the same way in the various jurisdictions; data available in (and obtainable from) the state of New Mexico may be markedly different from that in Arizona. By comparison, he said that the ACS at least provides a stable, consistently-applied-and-coded source for the kinds of data products and tools that he wants to develop for tribal leaders.

Tsosie then walked through a quick demonstration of his divison's website, http://www.navajobusiness.com, from which a variety of additional data analyses and reports are available.[8] The Division of Economic Development fields some data collection activities on its own—for instance, the Navajo Shopping Preference Survey described in one issue of their *Navajo Economic Data Bulletin* in 2012 and a visitor survey—but Tsosie said that they are increasingly turning to the ACS population data and related sources. On one dedicated subpage of the site,[9] Tsosie and colleagues performed a service for the individual Navajo chapters by extracting chapter-specific data from the most recent 5-year 2006–2010 ACS data as well as the 2010 census Summary File 1 tabulation. Admittedly, Tsosie said, some of these extracts are simply PDF files of tables constructed using the American FactFinder interface while others reformat the data into easier-to-work-with Excel files. However, the division's work in putting the documents in one place means that the chapters have a single source and do not need to learn the nuances of American FactFinder on their own. Tsosie said that this repository of reports—and the ACS's coverage in Navajo Nation—has been very well received by the chapters, and has been particularly appreciated by officials from the tribal government and individual chapters involved in preparing grant proposals.

Tsosie also lauded the topic coverage of the ACS as being useful in tribal decision making and resource allocation. One of the tabulations presented on the site—and one that Tsosie asked to be included in the case study/agenda book for the workshop—are chapter-level estimates to the question of whether the housing unit has complete plumbing facilities or not. He said that the questions

[8]See in particular the postings page at http://navajobusiness.com/pdf/Ads/Annoucemts.htm, under the "Doing Business" heading but not, apparently, readily linked from top-level pages.

[9]See http://navajobusiness.com/pdf/Ads/CensusRep.html.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

of whether the housing unit contains a flush toilet or plumbing facilities are often mentioned in discussions of whether the ACS is too intrusive, and they are sometimes lampooned. But he said that the questions are important to ask because—in many parts of the Navajo Nation, in other American Indian reservations, and in other rural communities—the answer to them is "no." Indeed, paging through the chapter-by-chapter listings, quite a few of the 110 chapters have a majority of housing units lacking complete plumbing facilities. If one takes plumbing for granted, the questions might seem frivolous, but in poor, rural areas the information is essential for documenting current conditions and working on changed policies.

Similar to Miller's observation that ACS data are certainly not perfect for rural analysis—but that they are the best and sometimes the only data available (Section 5–B)—Tsosie noted that he has seen some of the same quirks as others when working with small-area, small-population data. Looking at detailed data on household income, he said that he has seen one or two pockets of unusually high values. Like Plyer's anomalous high-wealth section in New Orleans' Lower Ninth Ward (see Section 3–C), the ACS estimates seem to show that "there are supposedly some very rich Navajos out there—we haven't found them yet." But, in his assessment, the benefits of the ACS data outweigh these quirks.

That said, Tsosie said that he needed to mention one special kind of "burden" associated with the ACS and, with it, one direction in which additional information from the Census Bureau could be a credibility boost. That "burden" is the obligation to not only detail the statistical and survey methodology of the ACS, but also explain the precise mechanics of how the data are collected. Though tribal and chapter leaders appreciate the value of the numbers they are beginning to receive from the ACS, there is a great deal of confusion over exactly how the data are collected—on and off reservation lands—on an ongoing basis. The data collection process is much more visible in the decennial census, with its high-profile publicity push and its structured interactions with tribal authorities on how to conduct the enumeration.[10] But exactly how the Census Bureau operates on tribal lands in off-census years, and how it regularly collects ACS information in on-reservation lands—while respecting the reservation's sovereignty—remains a great mystery to outside observers. Based on the comments that Tsosie and colleagues have received, speculation on how the ACS data are actually collected in the Navajo Nation ranges from "put[ting] out a table outside a Bashas' [Grocery Store] to get the information" by interview to "just get[ting] the data off of the different [web] sites out there on the Navajo Nation." Tsosie conceded that he and his colleagues are not really able

[10]In the 2010 census, tribes were given options on the degree to which in-person field visits would be used relative to mail (incoming or outgoing) methods. Most American Indian lands—including the Navajo Nation—were counted via Update/Enumerate methods—field staff knocking on every door, updating their address lists as necessary, and completing the questionnaire through personal interview (rather than simply leaving them to be returned by mail).

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

to dispel or correct these speculations because there does not seem to be good documentation of the true procedures in areas like the Navajo Nation without a great deal of conventional, city-style, mailable addresses. He also remarked that his efforts to "get some clarification" through the Census Bureau's Denver regional office has been met with "very limited success."

Revisiting that comment during the session's discussion period, Tsosie added that part of the concern over how the ACS data are collected on Navajo lands stems from longstanding issues of trust. As he put it, "our historical legacy is not always on good grounds with the federal government," and he conceded that it can be "very difficult to extract data from our people sometimes because of what the federal government has *done* to our people" in the past. But he closed by commending a great benefit of the ACS: its continual presence and its coverage of demographic information within the Navajo Nation serves to "validate our existence"; it "validates our sovereignty" by chronicling that the Navajo are and "will continue to be there on our home lands," and so the ACS is extremely valuable to the tribe.

5-E ACS VIEWS FROM THE COUNTIES, AND DISCUSSION

Asked to moderate the discussion session, National Association of Counties (NACo) Research Director Jacqueline Byers began by noting that the 2010 census was the fourth with which she has worked with the decennial census in a professional capacity, whether in her current role at NACo or when working in state government. Now, working with NACo members and helping them use ACS data, she said that—"finally"—she doesn't need to "apologize that the data is ten years old anymore . . . and it is a wonderful thing." Byers also said that she wanted to follow up on Tsosie's comments on the complete plumbing and flush toilet questions—and remark specifically on the discussion of privacy issues in the workshop session dedicated to ACS burden (summarized in Chapter 8, but a session that immediately preceded this state/local/tribal session during the workshop). She said that she finds it an "interesting contradiction in how people approach privacy" that "people tell everything about themselves on Facebook"—that you can barely slip and fall on the street (or anything else embarrassing) without it "going up on YouTube within half an hour"—yet will "have a problem with telling [the Census Bureau] how many toilets they have or what time they leave for work in the morning."

Turning to counties' perspectives on the ACS, Byers said that counties are an often-underestimated force in public policy; within the states, counties (and county-level equivalents) are typically the administrative units that "are charged with the responsibilities of making things work." Counties' work involves a great deal of planning—developing services and then ensuring that they are delivered to those who need them. Accordingly, "counties do *everything* with

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

data," and so are making extensive use of the ACS. Particularly in small, rural counties, the 5-year ACS data are playing a major role in counties' economic development plans—documenting the education and skill sets in the areas in order to attract new business (and jobs). As mentioned in the previous sessions, counties also make use of the ACS for targeting services for children and for foreign language assistance; for making transportation decisions like road repair programs and development of mass transit routes; and for developing public health and safety plans (and prepare for disasters). She noted seeing a recent news story talking about the millions of dollars being spent by counties to deal with unexpected consequences of tough economic conditions—additional funds to help the homeless and stock food banks to keep up with demand and funds spent to maintain foreclosed properties so that values of surrounding properties do not completely collapse. Increasingly, she said, counties are turning to the ACS to inform these kinds of interventions because they can get the information easily and frequently, especially in large urban counties.

She then asked each of the presenters for their responses to two questions—questions she admitted were "not particularly provocative" but ones important to the current debate on the nature and future of the ACS. First, she asked simply: How would the state governments, tribal authorities, and other client agencies described in the presentations do what they need to do *without* the ACS?

- Miller said that for her applications—trying to look at measures covering the whole country and trends in rural America in particular—"we would be very limited in what we could do and what we could say," to the extent that it is hard to imagine other alternatives. There are some other county-level data resources that might be brought to bear, but none has the same level of detail as the ACS.
- Brower answered that her center would most likely have to turn to state administrative records sources, but that there would be "no comparison" with the range of analysis that can be done with the ACS.
- Romalewski followed up on that comment, saying that he imagined that administrative records from state and local governments would be used more frequently, but he said that these would likely be less consistent and reliable than systematic collection of ACS data. He said that some of this spottiness in records data would arise from inconsistency in collection and coding at the most local levels; even in a large metropolitan county like Suffolk County on Long Island, the county "sometimes doesn't have much say over how the individual towns collect the information," and that's "not a good situation."
- Tsosie said that tribal governments like the Navajo Nation would likely have to turn to the federal Bureau of Indian Affairs (BIA) and try to access the records that BIA maintains. However, he said that he suspected that much of that data "would be outdated and probably not very useful."

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

For more detailed survey-type data, he said that the recourse could be to hire consultants or survey research firms, but that those costs could be prohibitive.

Byers' second question for the panel of speakers—and the rest of the workshop attendees—was what they specifically think would happen to their data projects and their ability to use ACS data if response to the survey were made purely voluntary rather than mandatory.

- Miller said that there are already difficult-to-reach and low-response-rate populations in the mandatory ACS—the most rural areas in the United States counting as one of those. She said that she suspected that a voluntary ACS would impair the effective sample size of the survey, and that the groups that are already most difficult to reach would be more likely to be missed in a voluntary ACS.
- Brower agreed, saying that "we wouldn't have the confidence that we have [now] in the representativeness" of the population. Given her center's analyses, she said that she would particularly worry about a distorted view of low-income or linguistically isolated populations.
- Romalewski said that, not being a statistician, he could not really judge the specific effects of what happens—but he expected that his mapping applications would be worse as a result.
- Tsosie recalled his concern about exactly how ACS data are collected on American Indian lands at present under the mandatory collection; if it is made voluntary, "forget it, [tribal areas] won't get anything."

Byers closed with some personal comments about the legislative debate on the future of the ACS. She said that this workshop—and its assemblage of ACS stakeholders—is an important effort. However, she said, a critical perspective not represented in the room is one that is critical to consider going forward: that perspective is "the average citizen who is filing the complaints" with the Census Bureau and their congressman about the ACS. Byers says that she sees this as a major educational challenge and opportunity—boosting awareness in the mind of the average citizen about how data from the census and ACS affect their lives. For example, the average person putting children on the school bus in the morning might not fully realize that—very often—it is ACS and census data that play major roles in planning the school bus service (e.g., ensuring that the children have a seat in the bus), in making sure that they have a classroom to go to at the end of the ride (school siting to meet demand), that there is a teacher standing in front of the room (school staffing decisions), and so forth.

Copyright © National Academy of Sciences. All rights reserved.

– 6 –

Business, Economic Development, and Data Aggregator Uses

The final workshop session that was focused on a particular sector of the American Community Survey (ACS) user base was designed to cover a range of perspectives on business-related uses. This topic includes the direct use of ACS analysis by private-sector firms for planning and implementation, the use of data by economic development authorities to promote business expansion and recruit new business opportunities, and the recoding and repackaging of ACS data by "data aggregators" to make the data easier to use for downstream users. Despite its origin in and widespread use in the public sector, the ACS has become a heavily relied upon resource in the business community. As hinted at in the discussion period for this session, the relationship between the private sector and the ACS is an increasingly important one because of two arguments that have been stirred up on the debate over the future of the ACS: first, whether the ACS is duplicative of information that could be obtained from other (including business-generated) sources and, second, whether—if businesses benefit strongly from the availability of ACS to make decisions—those businesses should pay more for the ACS's collection and maintenance.

Section 6–A summarizes the use of ACS data in the economic development and workforce community, through the specific example of a development group in Iowa. Section 6–B gives an overview of the experiences of one data aggregator—the Integrated Public Use Microdata Series (IPUMS) project housed at the University of Minnesota—and describes what IPUMS staff know about the users of their versions of ACS files. Three specific business examples

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

of ACS use are then discussed in turn—Acxiom's use of ACS data in marketing and business information management services (Section 6–C), the Conference Board's use of the data to study trends in teleworking and wage inequality (Section 6–D), and AIR Worldwide's incorporation of the data into their models of risk exposure to natural and man-made catastrophes (Section 6–E). David Crowe, chief economist for the National Association of Home Builders, moderated the discussion block for the session, starting with brief remarks on the ACS's role in the housing community (Section 6–F).

6–A ACS DATA FOR ECONOMIC DEVELOPMENT AND WORKFORCE PLANNING PROJECTS

Described by Andrew Conrad as having "a very academic name for a very practical organization," the Institute for Decision Making is an economic development organization affiliated with the University of Northern Iowa (UNI). In his remarks, Conrad briefly introduced five projects (or types of projects) conducted by the institute and its partner organization that make particular use of ACS data.

He said that the first of these—regional workforce analysis—had not really been broached in the workshop's previous sessions, but it relies "tremendously" upon ACS data. He said that he assisted with workforce analyses completed by and for three counties: the Texoma region of southeast Oklahoma and northeast Texas; the Cedar Valley region of northern Iowa; and the Siouxland region spanning parts of western Iowa, Nebraska, and South Dakota. These projects provide businesses and employers in the region in question with a picture of the characteristics of the area's workforce and estimates supplies and demands for specific worker types and talents, and they make extensive use of ACS data as well as employment data maintained by the Bureau of Labor Statistics (BLS). In fact, he said, he knew a few people in the workshop room had gone through training that he conducted at the Census Bureau and that training called for them to use these three regions. As a result, the training attendees "probably know more about these regions than they want to" because the training makes them work with the BLS and American FactFinder websites ("and see what it is really like for us who are data users").

Conrad said that these projects faced a major challenge in that a couple of them started before the first ACS 5-year numbers were released. Consistent with what Miller said about the availability of ACS data for rural areas, 64 of Iowa's 99 counties are under 20,000 population and so have no ACS estimates except for the 5-year numbers; indeed, two-thirds (667 of 950) of Iowa's incorporated places are under *1,000* population. In these projects, Conrad said that they use ACS data for basic demographic (population by age, sex, and race and Hispanic origin) and economic (household income and poverty status) informa-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

tion. However, examinations of the prevailing workforce also depend critically on ACS data on transportation and travel flows (commute time and means of transportation to work), and on the ACS's ability to profile employment earnings by educational attainment. In rural areas, the "laborshed" from which employers may draw workers can be geographically large; hence, these studies typically work with data at the county (or groups of counties) and state level—with some reference to the nation as a whole.

From the broad regional outlook to tract-level studies, Conrad said that the second type of projects for which his organization uses ACS data are neighborhood economic development technical assistance projects. In particular, he referenced work done with the Waterloo Neighborhood Economic Development Corporation (WNEDC) in his own home community (Waterloo being a neighbor city of Cedar Falls, home of UNI). Waterloo contains some of the most diverse neighborhoods in Iowa, defining "diverse" perhaps less stringently than in very large urban centers as census tracts containing about one-fourth (or greater) minority population. The new diversity of many of these neighborhoods is a fairly recent phenomenon, based on an influx in immigrants since 2000. Accordingly, Conrad said that the arrival of ACS data has been welcomed by groups like the WNEDC.

As with the general regional workforce projects, Conrad said that neighborhood-focused economic development efforts rely on the ACS for basic demographic, income/poverty, and industry/employment data at the census tract and, depending on city size, block group. But ACS items of particular interest for this analysis reflect the interests of the client audience. Neighborhood economic development projects are used to inform prospective entrepreneurs, so one very important metric derived from ACS data is an indicator of how many entrepreneurs are already at work in each small area—that is, the number and share of people who are self-employed. Because self-employment can occur either by choice or by need, Conrad said that additional analysis is needed to better understand the context of self-employment. These data are also used to recruit prospective businesses, and those clients tend to be interested in potential employment pools (daytime population in neighborhood areas) and household characteristics by neighborhood (e.g., what type of income might be tapped from a retail standpoint?).

The third project type Conrad described in which ACS data are used in economic development studies are asset mapping studies, at the state or regional level. He raised as an example a project done principally with two Iowa state agencies—Iowa Workforce Development (the state employment agency) and the Iowa Economic Development Authority—along with affiliated regional and local economic development authorities. This project attempted to take a high-level view of the regional economic development terrain in the state—constructing, for each of 16 state-defined marketing regions, asset maps to help assess the comparative strengths and weaknesses of the regions in terms of

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

strategic workforce assets. He said that these asset maps made use of a host of ACS data in addition to the standard demographic and income/poverty breakdowns—data on travel/commute time, educational attainment, health insurance and costs, and prevailing housing values and rents. The asset map documents made use of these and other variables at the state and county level, along with a U.S. total benchmark. Singling out the educational attainment variable, Conrad said that ACS data have been useful in documenting and understanding a phenomenon common to many rural areas: high school graduation rates that can actually be very strong, but vastly smaller numbers and shares of individuals with any post-secondary education (including 2-year as well as 4-year programs). Conrad also noted that ACS data involving health insurance information can be useful as a proxy for employment benefits offered in a particular area.

A fourth type of economic development study using ACS data is the derivation and comparison of regional metrics—with a particular eye toward understanding change over time (or the lack thereof) rather than a pure snapshot at a time period. He used as an example an analysis that his institute conducted with the Iowa City Area Development Group. A public/private partnership, this group wanted to better understand the drivers of economic growth in the Iowa City area; Conrad said that the area has grown due to the presence of the University of Iowa and its affiliated hospitals as well as expansion of private-sector firms like ACT, Inc., Proctor and Gamble, and Rockwell Collins (the latter headquartered in nearby Cedar Rapids). In particular, they were interested in using data both to establish a benchmark (current economic development) and to project and estimate future economic activity and the possible impact of new development strategies. ACS data used in developing this set of metrics included earnings by educational attainment, travel/commute time and means of transportation to work, and income/poverty status, as well as basic demographic splits. For use as a benchmark, this work involved use of ACS estimates at the place (city/town), county, state, and national levels; as Conrad said, "Iowa City wants to know what other metro markets are their competitors," nearby and nationwide, and regional economic development bodies generally have a "hit list" of other localities that they compete against (and against which they need to be compared).

As a final example, Conrad turned to a pure state-level measure—the development of an Iowa Competitiveness Index. Unlike some other states, Iowa does not have a statewide chamber of commerce; instead, the Iowa Business Council brings together representatives from about 20 of the largest employers in Iowa (including firms like John Deere and Company, Pella Corporation, Rockwell Collins, and the major state universities). As Conrad put it, the Business Council's role is not to lobby but still to "push issues"—as it has for a number of years, for instance, on the status of pre-kindergarten education and services. Like the Iowa City example of regional metrics, the Business Council sought a way to benchmark the state economy and track progress over the years. Accordingly,

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

in 2011, the Council published the first iteration of the Iowa Competitiveness Index, a set of key indicators that are scored and combined to produce general assessments of the state's economic strength along five broad dimensions: economic growth, education and workforce readiness, governance and fiscal matters, health and well-being, and workforce demographics and diversity. The index has been updated for 2012, and the Council intends to maintain it as a regular series. Several of the indicators—the levels of which are compared and ranked across all the states and then scored by their "overall competitiveness trend" (improving, no significant progress, or worsening)—are derived from a variety of sources, including data from the Iowa Department of Education (by way of the National Center for Education Statistics) and the Iowa-based ACT, Inc., college testing service. However, several of the indicators are drawn from the ACS, such as per capita income and measures of racial and ethnic diversity.

In closing, Conrad said that he wanted to emphasize that economic development is a process, not an event. Some elected officials mistakenly think of it as a discrete event—a particular ribbon-cutting or ground-breaking—when it is a complicated (and sometimes lengthy) process. He said that he recalled working on a specific project for many months—working with Target Corporation about 10 years ago on the placement of a 1.3 million square foot distribution center. Conrad's Institute for Decision Making was involved in early work on the project for a long while before they even knew the specific company interested in making the move, and it continued as Target "kicked the tires in the community" before making its final decisions. But Conrad said that the main point is that economic development is not only a process but one that involves data-driven decisions at each step. Like other companies, Target had to approach the decision from a market analysis perspective—where might they be opening or expanding stores, how are they going to serve those markets, and where would the workforce for both the stores and distribution center come from. At even more detail, wanting to enter the community as an "employer of choice," Target wanted to understand the wages and benefits that it would have to provide to be competitive. Conrad said that data analysis from the ACS (including the use of health insurance as a proxy for benefits, as described above) has proven useful in these kinds of projects.

Conrad also argued that ACS-based analysis is critical for entrepreneurs. He recalled a specific example from Waterloo, of the WNEDC working with a local Hispanic couple interested in opening a *tienda* in the area. That case involved detailed analysis—possible only with the ACS—of 13 census tracts, using the data to try to pick an ideal location based on the demographics of the area and the access to transportation networks (anticipating that much of their clientele would be foot or public transit traffic). As another example, Conrad said that he had worked with Pella Corporation on numerous site location projects, profiling small-area local workforces and trying to determine which areas have workers with the educational attainment and skills needed for the facility and

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

whether—if one built a 400-employee facility in a town of 5,000 population—housing would be there for the new workforce. Unfortunately, he said, the ACS data were not available at the time of this project.

During the discussion period for the session, Conrad was asked by his colleague in economic development work—Patrick Jankowski (Greater Houston Partnership and member of the workshop steering committee)—to comment on particular uses of ACS data to profile and promote areas to foreign corporations and foreign investors. Conrad replied that he has worked in such cases before and agreed that the ACS data are very helpful in building cases for foreign investors—with two curious caveats. First, he said that he is struck that foreign firms are "shocked at the level of detail that we are able to provide" through public data resources. It is not only the ACS that yields this response—when economic developers are attempting to weave a data-based story for investors, data resources like the Bureau of Labor Statistics' Quarterly Census of Employment of Wages and its Occupational Employment Statistics program also play important roles. But the level of demographic detail from the ACS surprises foreign firms—which occasions the second "caveat," which is that foreign investors will commonly engage other consultant firms to review the data work by economic developers as a check on whether the numbers are correct.

Put briefly, Conrad said that professionals in the economic development community "use the ACS data on a daily basis"—and communicate impressions and findings from those data to their end clients. To be sure, he said, ACS data are not the only important data resource; they rely on other government-produced data series such as those produced by the Bureau of Labor Statistics. But, he said, the ACS is becoming increasingly critical to making the kind of "informed decisions" that help communities and local governments cultivate entrepreneurs and businesses, and continued availability of the data is critical. He noted that other speakers had "wish lists" to make the ACS more useful and relevant; from the economic development perspective, Conrad said that the addition of a question on multiple jobs—how many jobs each person actually holds—would be invaluable in sorting out dynamics that are presently murky. Above all, though, Conrad stressed the importance of the 5-year ACS estimates to a state like Iowa and its roughly 3.1 million population—a state in which two-thirds of counties have less than 20,000 population and only 22 communities have populations sufficient to obtain 3-year average estimates.

6–B IPUMS: COMPILING AND DISSEMINATING ACS DATA PRODUCTS

Funded by the National Science Foundation and the National Institutes of Health and administered by the Minnesota Population Center (MPC) at the University of Minnesota, the IPUMS project produces (and makes available for

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

free) microdata series for research use that are recoded from the originals to use consistent labels and concepts across time (past iterations) and space (other countries).[1] For the purposes of this workshop—and among U.S. data users—the term "IPUMS" is generally used to refer to products from the IPUMS-USA subproject, including harmonized, consistently coded microdata samples from all U.S. censuses since 1850[2] and—since 2001—the ACS.[3] The other two branches under the IPUMS label are the IPUMS International subproject that enables linkages across multiple iterations of several foreign censuses and IPUMS-CPS, which applies consistent coding to multiple microdata series from the Current Population Survey (CPS) samples. Across all its data sets, MPC currently maintains on the order of 800–850 million individual person records—and it distributes them to a steadily growing user base. Katie Genadek of the IPUMS user support team said that MPC averaged slightly less than 100 gigabytes of data downloaded by users from its servers in 2001, around the time that its web distribution strategy came into place; as of 2010, the data volume moved by MPC had grown 10-fold to about 1 terabyte per week, on average. The IPUMS-USA samples, including the 1-percent microdata samples from the ACS (yearly and multiyear samples), are the most widely used of MPC's data resources, and currently have about 30,000 registered users on MPC's website.

Speaking for the IPUMS user support team, Genadek remarked that her workshop presentation would be different from the rest of the workshop because she has *used* ACS data extensively but is not actively in the business of *producing* anything from the ACS—"no map, no table, nothing." But the perspective she said she would bring to the session was that of a data aggregator—repackaging and adding value to the base ACS products and brokering their use by a broader class of data users. Significantly, she noted that she can also speak to the ways in which users manipulate and analyze the IPUMS version of ACS data files.

Genadek said that the "value added" to the IPUMS version of the public ACS files runs along four basic lines—three related to data content. The first is the database's titular integration. Even in the relatively short lifespan of the ACS, there have been changes to the exact coding of responses to particular questions, to the labels associated with specific responses on the questionnaires, and to un-

[1] Additional information on IPUMS is available from the project's website, http://www.ipums.org.

[2] There is one exception to IPUMS' census coverage; most of the original schedules (the ledger-type "questionnaires" on which personal information were recorded) from the 1890 census, from which microdata would be drawn, were destroyed by fire in 1921.

[3] Though IPUMS is well known (and a widely used term) among data users, the preferred way that the term should be *pronounced* is less clear. As an aside to her presentation, Genadek remarked that the preferred in-house pronunciation is *ih*-pums (soft I)—consistent with the pronunciation of the word ("integrated") providing the initial—although *eye*-pums (hard I) is commonly used. As she said, "we are not related to the Apple products"—and try to avoid association with them—making the pronunciation distinction more important in recent years.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

derlying geographic and other codes. To the greatest extent possible, IPUMS recodes the public files to be able to facilitate more direct comparisons across years—or to make pooling of different years' samples possible. The second component of value added is the detail of the online documentation of the data sets. The samples themselves are described with care, as are individual variables; the history of individual variables, including the comparability of data across samples and changes in the underlying questionnaire texts, is recounted in detail. The third content-related value-added component is "new" variables derived by MPC and included on the IPUMS files. These include "crosswalk" occupation and industry variables (for which consistency across year and beween samples is particularly hard to maintain because of changing definitions—and the changing nature of industries). The IPUMS-coded ACS data also include derived variables to reconstruct family and subfamily interrelationships, to permit analyses using families as the basis for analysis rather than household/housing unit. The family relationship variables are derived from various patterns in responses as well as the general census/ACS question on relationship to household head—an item that technically asks about the relationship of each person in a housing unit to "Person 1" (the one who completes the questionnaire or who is interviewed), who is intended to be but is not necessarily the effective head of household—or, for that matter, a family member of some people in the housing unit[4]).

Genadek said that the fourth value-added component is the online interface and data extraction system, which gives users great flexibility in selecting as many or few variables as they might like from as many samples (census or ACS) as desired. The IPUMS site has also taken care over the years to provide results in many output forms, in whichever way is most convenient for the user to access and analyze. User-specified extracts can be output in formats for common statistical analysis packages (SAS, Stata, SPSS); recently, facility was added to generate results in basic comma-separated value (CSV) format, for easy input into Microsoft Excel or other spreadsheet software. For users without access to statistical software, the MPC platform has built-in access to the Survey Documentation and Analysis (SDA) package developed by the University of California, Berkeley, Computer-Assisted Survey Methods Program. Hence, it is possible for users to directly analyze all of the IPUMS-USA (census and ACS microdata samples) through online queries. Genadek said that she is part of the IPUMS user support team that has proven to be a popular aspect of using the ACS data; she and her small team of colleagues field email and telephone inquiries about the data and using the data, with the goal of responding to all questions within three days (and, more often, within a few hours). At the workshop, she said that

[4]Similar to earlier versions, the main body of the 2012 ACS questionnaire begins with the parenthetical instruction: "(Person 1 is the person living or staying here in whose name this house or apartment is owned, being bought, or rented. If there is no such person, start with the name of any adult living or staying here.)"

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

she had specifically checked her email for queries from IPUMS ACS users and found about 2,500 such email messages.

Of all the user-specified extracts generated from IPUMS-USA—samples from the 1850–2000 decennial censuses and the 2001–2010 iterations of the ACS—some version of the ACS was included in 55 percent of the queries. Of those IPUMS requests asking for ACS samples, 137,029 have asked for one or more 1-year PUMS files; smaller but sizable numbers of extracts have asked for the more recently available 3-year (5,726 extracts) and 5-year (2,141) samples. Also a fairly recent development, the SDA-based online analysis component of the IPUMS interface has generated 250,000 user-requested tables using ACS data—13,000 of those in the month preceding the workshop alone.

Turning to the question of who uses the IPUMS-produced files, Genadek noted that academic scholars are the largest segment of IPUMS' user base. As evidence, she displayed a graph of the number of citations to three MPC databases—including, and with trends driven principally by, IPUMS—each year from 1993 to 2011, as found in Google Scholar searches. Similar to the graph of MPC data downloads, references to IPUMS and MPC resources in the academic literature began to grow precipitously around 2002 and have increased by roughly 50 percent over the last 3 years, hitting a level of just over 1,100 citations in 2011. As Genadek indicated, the graph includes publications from non-IPUMS data maintained by MPC; however, Genadek said that she believed that the IPUMS versions of ACS products account for a large part of this growth and that the ACS "is driving a lot of this research right now."

Using basic information from the user profiles defined on the MPC site, Genadek briefly displayed a pie chart showing the rough breakdown of IPUMS users' backgrounds. Indeed, users with academic affiliations make up about 77 percent of the known users; the largest share of these are users affiliated with an economics department (31 percent), followed by demography/sociology (16 percent). Genadek says that MPC and IPUMS staff are aware of IPUMS data being the source for "tons of dissertations"—graduate students may be "our number one user"—and an increasing number of class assignments and projects. But other blocs of IPUMS users are significant—state and local government agencies, private industry, journalists. Genadek said that, at the federal level, IPUMS data had been accessed and used by—among others—the U.S. Government Accountability Office, the Bureau of Labor Statistics, and the Federal Reserve Board (and its member banks).

IPUMS' online extract system is such that IPUMS staff can determine how frequently users check specific ACS variables for inclusion in their extract data sets. Genadek reported that the data item on race is "by far number one" in terms of included variables. But some ACS-specific variables (or, at least, variables that are consistently reported across small areas by the ACS) are interspersed with the standard demographic variables, suggesting that users are not just using the ACS for general population estimates. To wit, the second-most-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

included variable in IPUMS extract is employment status—a major driver in the workforce and employment studies described by Conrad (Section 6–A). Person age ranks third, but place of birth and educational attainment both slip into the rankings ahead of sex and Hispanic origin. Rounding out the list of the most commonly accessed variables in IPUMS ACS, in descending order, are relationship to household head, metropolitan area, state, marital status, income, and language spoken at home.

Genadek said that she also combed through the emails to the IPUMS support team and characterized the most frequent questions in terms of both topic and technical concern. In terms of topics, she said that the IPUMS staff are most often contacted for clarity on poverty estimates and the relationship of poverty levels to the ACS microdata. Next most frequent are questions of a geographic coverage nature—why, in the microdata sample, do people appear from one specific county and not another? Questions about migration and mobility of people within the ACS sampling framework are common, as are questions about the derived variables on family and subfamily composition. Questions about occupation coding, measures of physical or mental disability, and income round out the topic-based questions. On more technical matters, IPUMS staff commonly field questions about how income variables are adjusted based on the rolling, multiyear sample design; they also receive and address questions about weighting, allocation flags, and variance estimation, as well as more general questions on the interpretation and communication of estimates from multiyear samples.

In communication with IPUMS users, Genadek commented that she also has a sense of common wishes—things that IPUMS users would really like to have included in the ACS. In her assessment, the biggest thing that she hears from users is desire for the month of the survey's administration to be coded in the microdata. In addition to possibly permitting some glimpses at seasonal populations and very precise construction of age variables, the month of interview could also allow advanced users to apply their own income adjustments to the finance and income questions. She conceded that she was not sure exactly how much of a concern the identification of survey month would be from a confidentiality standpoint, but it remains a major part of users' wish lists for the survey. Another common request is inclusion of a clearer and more detailed relationship question. The IPUMS coding tries to reconstruct family structures as much as possible, but the form of the relationship question and the way in which the question is phrased can obscure some family relationships depending on who answers the question.[5] Also high on IPUMS users' wish lists is the re-

[5]A previous National Research Council (2006:133) report, following Schwede (2003), cited a simple example of this masking phenomenon: "Consider a case where a man and woman live together, unmarried, along with the woman's child from a previous relationship. If the man is the census respondent, the woman may be reported as an unmarried partner or an 'other nonrelative,' while the child would likely be 'other nonrelative'; the biological link between woman and child is obscured. If the woman is the respondent, the biological link between her and her child would be

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

instatement of two variables that were available in previous census long-form samples but were dropped over time. The lifetime fertility question—number of children ever born to a mother—was dropped from the 2000 census long-form sample and has not resurfaced in the ACS. Questions on the place of birth of mother and father have been absent even longer, having been dropped from the 1980 census long form; the Census Bureau announced plans to add parental place of birth questions to the ACS questionnaire in 2013, but withdrew those plans in May 2012.[6] But, Genadek said, much though IPUMS ACS users might appreciate such small revisions, her sense is that the users' main concern and interest is for *minimal* change in the future. IPUMS users tend to be interested in long, historical time sequences—for that, one wants continuity in content, form of question, and sample size. Hints or rumors of content being changed in major ways or dropped together—periodically, for example, the ACS question on ancestry is thought to be vulnerable—are disturbing prospects for policy-relevant analyses of specific ancestry groups such as Brower's profile of specific Asian subpopulations in Minnesota (Section 5–A).

Genadek closed by referencing two additional MPC products, including the National Historical Geographic Information System (NHGIS)—essentially, a mapping analogue to the extended time series across censuses and data collections that can be calculated from IPUMS data. Small-area data (down to the finest level of resolution possible in earlier years) can be extracted in mappable form from the 1790–2010 censuses and, now, from ACS 1-year and multiyear products. Like the IPUMS extraction tool, users can pull multiple tables and variables from multiple years at once, and shapefiles for use in geographic information systems software for the different time periods can be downloaded as well. The NHGIS interface also gives users access to a summary file not available from interactive tables on the Census Bureau's American FactFinder site: a 5-year summary file from the 2006–2010 ACS that can be queried down to the block group level. She said that MPC plans to disseminate all of the ACS files as they become available and are processed—not as a replacement or competitor to FactFinder, but as an alternative for people who find FactFinder hard to navigate.

---

preserved, but it would be ambiguous to family and household researchers whether the male is the child's biological father or not."

[6]The Bureau posted notice of new ACS questions in the *Federal Register* (77 FR 18203–18205) in advance of its filing a request for three additional years of clearance from the U.S. Office of Management and Budget (OMB) under the Paperwork Reduction Act. The decision to withdraw the parental place of birth questions was communicated in a memo updating that clearance package, viewable at http://www.reginfo.gov under Information Collection Review number 201202-0607-003.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

6–C ACS IN BUSINESS: MARKETING SERVICES AND INFORMATION MANAGEMENT SERVICES

Formerly a statistician with the Census Bureau, Matthew Christenson (Acxiom Corporation) began his remarks by noting a basic difference between his former and current employers in summarizing the process of getting his slides approved for presentation. At the Census Bureau, the overriding worry was always content; with Acxiom, a marketing services company, the focus is *branding*. Putting the point succinctly, Christenson began by saying, "I work for a company called Acxiom"—and "after this presentation you may not be able to forget that name," as it is "probably the most commonly used word in the presentation." Christenson gave an overview of how ACS products factor into Acxiom's work, but the broader point he alluded to in his opening—the comparison of a focus on branding with the identity crisis of sorts noted in the media perspectives session (Chapter 4) and other points in the workshop, in which the ACS has struggled for recognition in some circles as something other than general "census data"—would recur during the discussion period.

Headquartered in Little Rock, Arkansas, Acxiom was founded in 1969 and went public (on the NASDAQ exchange) in 1983. Christenson said that Acxiom's roughly 6,100 employees are distributed across offices on five continents (including several U.S. locations other than the Little Rock headquarters); in fiscal year 2012, Acxiom reported just over $1.1 billion in revenue. He observed that Acxiom's target clients are Fortune 1000 companies and counts among its clients many of the largest companies in the United States (and the world) in industries ranging from credit card issuers to airlines to telecommunications/media. Acxiom also engages in work for several federal government agencies.

Christenson described Acxiom's work as being three distinct but related businesses operating under the same company. The first of their constituent business is what might be called "recognition services"—helping client businesses recognize their individual customers regardless of how they make contact with the company (direct phone, email, or letter contact, or electronically through web browsers). This permits linkages to other data to be drawn when a person contacts the client, to make customer service more efficient; Christenson said that recognition services are also used as part of companies' fraud prevention plans. The second constituent business—and the branch in which Christenson works—is marketing services, identifying potential customers and establishing contact with them. More colloquially, Christenson said, people "typically accuse" people in marketing services "of being responsible for your junk mail, for the junk emails"—but, done more correctly and precisely, he thinks of marketing services as being responsible "for everything you *don't* get." That is, the goal is a marketing outreach program allowing businesses to target and engage potential customers but without ads for retirement services being

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

sent to young people, for ads for Depends undergarments to be sent to households with young children, and so on. Finally, Acxiom operates an information management services businesses—which is to say that its computer servers are commonly used to host and manage information in online commerce and banking transactions, a line of work that means that the company hosts some of the largest data sets in the world.

Under Acxiom's marketing services umbrella, Christenson outlined three key data products that make use of ACS estimates—or to which ACS estimates serve as a complement—in various ways:

- Based on an input name and address, Acxiom's *InfoBase Enhancement* product returns a host of individual- and household-level data (demographic, financial, housing, and marketing propensity) that can be appended to the records for further analysis.
- *InfoBase List* is "our equivalent to the census," as Christenson put it—it is as complete a roster of "economically active individuals and households" as can be derived from commercial and other records sources, and this is the data resource primarily used for "customer prospecting."
- The *Market Indices* product is a distillation of demographic and economic information that can be queried by Acxiom clients by address—providing a detailed small-area look at the immediate area surrounding an address—or directly appended to a client by matching to address.

Of these, Market Indices is the most heavily dependent on ACS estimates—consisting, in present form, of over 500 preprocessed data elements from the ACS, derived for levels of geography ranging from the nation as a whole down to block group. The product is so intertwined with the ACS that (as Christenson mentioned during the closing discussion) Acxiom is now beginning to brand new releases of the product as "Market Indices ACS." Specifically, Market Indices aggregates and assembles data from the ACS 5-year summary files; Christenson said, in candor, that this work is done with estimates from the 5-year summary files, and that Acxiom has little to no use for the 1-year or 3-year estimates. With equal candor, he commented that "we don't throw out very much of the data" from the summary file; the data items pulled for use in the product are fairly exhaustive of the questions on the ACS—a limited number of population count variables, but many more variables expressed as percentages, rates, or measures central tendency (mean/median). In terms of the value added to the ACS files, Christenson said that what Acxiom "sells" with the Market Indices product is not so much the data themselves as "the speed [with] which we can append it to huge datasets," on the order of hundreds of millions of records per month—approaching real-time matching and delivery.

Like the data aggregator role described by Genadek concerning IPUMS, Acxiom's work with Market Indices depends vitally on meeting end users' needs—in this case, the needs of Acxiom clients. However, in a pure business

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

environment, Christenson said that he could not really speak to *how* Acxiom's clients use the data—Acxiom sells the data to their clients' competitors, too, so clients don't really "clue us into" their uses. But, in broad strokes, companies use Market Indices for online (and offline) targeting of marketing materials, for the selection of prospect lists, and to improve customer care. Through a better understanding of local conditions, companies use the products to devise strategies to reduce customer churn and turnover; they also use it to develop cross-sell or up-sell opportunities in marketing approaches (e.g., a customer buys one product and is told—based on local information—about other products that might be of interest, or bigger/better versions of the product).

The ties between ACS estimates and Acxiom's InfoBase products are less direct, yet still important. ACS data are used to construct several of the "modeled elements" that go into the individual- and household-level data "appends" that comprise InfoBase Enhancement and that are available as an add-on to InfoBase List. Acxiom itself uses data from ACS 1-year PUMS files to build at least 75 of these modeled elements; still others are constructed (possibly including ACS data as well) by Acxiom's consulting organization; distributions of the modeled elements matched to InfoBase List are also compared with distributions of related variables in the PUMS files, as a check on whether the models seem to be working properly. These modeled elements range from demographic characteristics (e.g., educational attainment, length of residence) to market relationships and behaviors (e.g., behavior as an investor, media preferences) to online behaviors (e.g., behavioral models to predict Twitter or Facebook usage).

Being matched to person-level attributes rather than just address or geographic location, the modeled elements in the InfoBase products are applied to vastly more records each month or year than the block-group-limited Market Indices file—billions of records each month, Christenson said—but the general uses clients likely make of the data are roughly the same. Though the ultimate fine-grained data (individual and household level) that can be estimated by these products are Acxiom's stock in trade—ultimately, the items that Acxiom's clients generally find most valuable—Christenson said that they and their clients both use ACS estimates alongside the more specific analyses; the breadth and completeness of the ACS data are important to their users. Christenson said that he also understands that Acxiom clients are increasingly importing their information into geographic information systems (GIS) packages for mapping and trade-area analysis. More generally, speaking of the product construction process, Christenson said that Acxiom finds the ACS extremely valuable in its model-building exercises; their sense is that the quality of the ACS data increases the precision of their models (and ultimately the modeled elements), and access to the data makes creating and tweaking the individual models easier. The role briefly mentioned above of distributions of PUMS variables being used as a quality check on modeled elements is also important to the company; Christenson said that some results from models for individual- and household-level

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

data could charitably be described as "funky," and "the ACS helps smooth out the rough spots." Hitting a theme raised in Chapter 4 and that would come up again in the discussion period, Christenson also said that Acxiom's making use of the ACS imparts credibility to the products—it "carries quite a bit of weight with our customers if we say we are using Census [Bureau] data."

Christenson closed by noting that, from his and Acxiom's perspective, the ACS is a uniquely useful data resource because of its breadth, depth, and completeness, and it is useful both as a product that Acxiom can develop and make available to users as well as a source for its own modeling and estimation work. From a business perspective, the ready availability of annual releases of small-area data is greatly beneficial; constantly vying with competitors, it is good to feel current and up to date in estimates and projections (and not to feel like competitors have a leg up somehow).

Like other workshop speakers, he said that he would certainly like "more" from the ACS—but in a slightly different way. He said that he regularly hears from the company's salespeople and clients who are interested in ever more granular data—frustrated that household-level (or, for that matter, individual-level) data are not readily available from the ACS. From the marketing standpoint, their clients are greatly interested in data down to ZIP+4 Codes—a cut that could be as fine as 2–3 households. The reason why these incredibly fine-grained data are not available is fairly straightforward—respondents' privacy must be respected and personally identifiable information not disclosed—and so Christenson said that the ACS block group level is very valuable to Acxiom. But one concern that he wanted to express concerned the granularity of the categories used to calculate estimates. Until the arrival of the ACS, he said, Acxiom had been using the 2000 census-based Summary File 3—developed its products around the categories used there—and encountered significant problems when the categories used in variables like income or age ranges were made more coarse in the ACS files. Put bluntly, he said, "we don't care about confidence intervals"—in the sense of being paralyzed by the reported standard errors. Dealing with uncertainty is an accepted part of the bargain; Acxiom has statisticians, and its clients have statisticians, to sort through those issues, and they would rather be in the position of working with finer categories and making their own conclusions about what numbers are fine to use and which are not.

6-D ACS IN BUSINESS: UNDERSTANDING TELEWORKING, WAGE INEQUALITY, AND HOUSING

Though most people may be familiar with the Conference Board through its highly visible Consumer Confidence Index, Gad Levanon (director of macroeconomic research) emphasized that the Conference Board's analysis and research

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

work spans a wide range. A membership research organization that counts hundreds of private and public corporations as members, the Conference Board conducts lines of research in economics, labor markets, human capital, and other topic areas important to businesses. Levanon noted that the ACS is a fairly recent discovery for the Conference Board but that it has already factored significantly into several projects; like many other users, awareness of the ACS data—and the availability of those data in sufficiently fine-grained form—really came about with the release of the first 5-year ACS estimates and the PUMS files. Like previous speakers, Levanon said that—prior to the ACS—his analyses depended heavily on the Current Population Survey (CPS); also like previous speakers, he has found the ACS's sample size relative to the CPS makes a lot of things possible analytically that could not be done before. In his workshop comments, Levanon said that he would review three Conference Board projects relying on ACS data—one still in progress—as case studies.

First, he described analysis that the Conference Board has done on teleworking—working from home. The ACS question on mode of transportation to get to work includes "worked at home" as a response option,[7] which permits study of people who usually telework (even though the analysis might omit people who occasionally telework). Other ACS questions and variables allow the analysis to get as close to the teleworking community as possible—including full-time workers who work for some employers and excluding those who are self-employed.[8] ACS data (including the not-yet-full-scale collection between 2000 and 2005) suggest that the overall percentage of people who primarily work from home roughly doubled over the past decade, even though the percentage is very small—from roughly 1 percent in 2000 to just over 2 percent in 2010. Levanon explained that this overall percentage presents a distorted view because it includes a lot of people with occupations where the telework percentage is essentially (or necessarily) zero; to wit, elementary school teachers cannot primarily work from home by nature of the job, linked to the workplace (the school). Still, the overall percentage does suggest an escalating trend; growth seems to have been particularly rapid in the last 5 years relative to earlier in the decade.

[7]Person Question 31 on the 2012 ACS questionnaire asks "How did this person usually get to work LAST WEEK?" and permits the respondent to check multiple responses from the following: car, truck, or van; bus or trolley bus; streetcar or trolley car; subway or elevated; railroad; ferryboat; taxicab; motorcycle; bicycle; walked; worked at home; and other method. Answering "worked at home" routes the respondent to Question 39a, skipping over questions on commute time and unemployment/layoffs. Person Question 30 also asks "At what location did this person work LAST WEEK?," and that location could presumably be compared with the housing unit address.

[8]Person Question 41 on the 2012 ACS questionnaire asks about the nature of each person's current or most recent job activity, including two self-employed categories (depending on whether the person's business is incorporated or not incorporated) as options; responses for different classes of private- and public-sector employment (e.g., work for a private nonprofit organization or for a state government) are also permitted.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Levanon said that the great advantage of the ACS in this analysis is the level of detail in the coding of occupations. The ACS microdata permit researchers to drill down to very detailed levels of occupations—roughly the equivalent of the 6-digit Standard Occupational Classification (SOC) codes used in other federal statistical data products. From the experience of working with the occupation variables, Levanon raised an issue for future work and clarification by the Census Bureau—working to make the occupational codes more comparable over time, because changes in the occupation coding can make it difficult to construct coherent time series. Indeed, Levanon noted, he and his Conference Board colleagues worked with IPUMS data—and its more aggressive coding to promote comparability over time (see Section 6–B)—on this project, and they ultimately wound up primarily using IPUMS' OCC1990 variable—a recode to 389 occupational categories based on the occupation codes used in the 1990 census.

Levanon said that it is evident from looking at the work-from-home data by very detailed occupation that there are some work types for which work-from-home is very common, with potential teleworkers making up 10 percent or more of the workforce in some categories. Levanon's tabulations show one detailed category has a work-from-home percentage that dwarfs all others: medical transcriptionists, with 44.6 percent of workers in that class reporting that they work primarily at home. Other job categories with sizable work-at-home contingents (over 10 percent) include sales engineers/sales representatives and travel agents—jobs in which employers might have incentive to save costs by reducing or eliminating office space for people who may not need to be in the office very often. Other occupational categories with fairly strong telework levels are information technology-related careers, such as web developers, computer network architects, and computer hardware engineers.

Comparing work-from-home estimates from early in the decade (averages from ACS files from 2001–2003) with those late in the decade (2008–2010 average), Levanon also observed that some occupations have experienced particularly strong growth in telework. For some categories, this suggests that networks and remote access tools have developed sufficiently over the decade to make employers more comfortable with the telework option: For instance, telework among records clerks grew from 0.9 percent (2001–2003) to 5.5 percent (2008–2010) and among travel agents from 1.8 percent to 5.9 percent. Insurance underwriters and bill/account collectors were job categories where telework more than doubled over the decade. Still other job categories with fairly strong work-from-home levels did not experience the same amount of growth, which Levanon said was likely due to the fact that improvements in technology could not or would not greatly affect the ability to work from home—these include door-to-door sales vendors (a more modest growth from 5.1 percent to 7.1 percent) and clergy and religious workers (among whom work-from-home decreased slightly, from 6.7 percent to 6.2 percent).

Levanon said that he and his colleagues have also observed interesting dif-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

ferences in teleworking propensity by region. In their models of teleworking behavior, they included dummy variables to capture state-level effects (controlling for individual characteristics); classifying and mapping these state-level effects, he said the regional differences are striking. In general, teleworking rates are considerably higher in the West—likely because "hubs of technology" and places with high concentrations of the industries most likely to enable telework are on the West Coast; high levels of telework are also found in the extremes of the Eastern Seaboard (New England, Georgia, and Florida), while the lowest rates are in the Deep South and the Rust Belt.

Of this examination of telework behaviors, Levanon made clear that this is not especially complex work—indeed, he said, this is a quite "simple usage of the microdata." But he said that he thinks that it speaks to an important base of potential data users—namely, people in the human resources field interested in trends in the workforce. Levanon said that, as best he knows, the Conference Board's specific analysis of teleworking using the ACS data had not been done before, and he said that the range of interesting kinds of things that can be learned from ACS microdata might spur additional work—and interest in using the microdata—among parties new to the data.

The second ACS-related Conference Board project that Levanon described is one that is still in progress, but one that again makes use of the ACS's strength in providing consistent measures across a broad range of geography. Specifically, the Conference Board is studying wage inequality, by any number of factors. As an example of the work and the ACS's utility in it, he displayed the graph shown in Figure 6-1. The graph provides a general sense of geographic differences in wage inequality through a relatively straightforward metric; making use of ACS microdata, he calculated the ratio of the 90th percentile of total wages for full-time workers in each state to the 10th percentile. Printed at a small size, he noted that it is hard to read, but some of the states at the top—with the largest ratios and hence the greatest spread in wages—include California, Texas, New Jersey, the District of Columbia, Georgia, and Virginia. At the bottom—smaller ratios and less spread across wages—are South Dakota, Maine, Vermont, North Dakota, Wisconsin, and Iowa. The dark vertical bar shows the ratio for the nation as a whole. He said that this fairly simple univariate slice from ACS microdata spurs questions and areas for future probing; based on the states on the high and low ends of the spectrum, "one can speculate that ethnic diversity" might be an important determinant of the inequality, and that is something that can be examined at finer levels of aggregation. The national-level vertical line—and the fact that so many states are below the line—suggests interesting trends as well; in early looks at the change in the ratios over time, Levanon said that shifts in the national-level ratio are more pronounced than are shifts in the distribution of inequality across the states. Again, he observed, these are interesting phenomena that "I don't think can be done [using] any other data source" with such a level of confidence.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

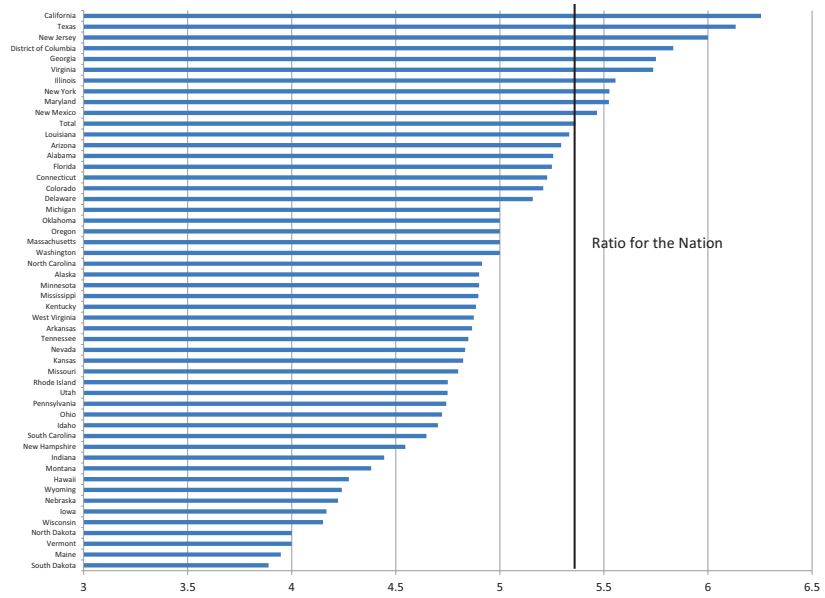


Figure 6-1 Ratio of 90th percentile household income to 10th percentile, by state

SOURCES: Calculated from American Community Survey data; adapted from workshop presentation by Gad Levanon.

Finally, Levanon described work that he had done on housing characteristics and demand in support of the Demand Institute, a new joint venture sponsored by the Conference Board and Nielsen. Given the current economic climate, one factor that the Demand Institute was interested in learning about is the phenomenon of "doubling up"—multiple families or individuals sharing the same housing unit. Within that "doubling up" population, an important subgroup is young adults continuing to live in the parental home. Levanon said that the ACS data were extremely useful to illuminate some points about this population. He briefly displayed a chart using data from 2006 and 2010 ACS files examining the percentage change in home ownership rates, which varies inversely with age group; home ownership rates dropped by almost 18 percent among 20–24-year-olds and almost 14 percent for 25–29-year-olds, compared to only a 0.5 percent change among 65–69-year-olds. Part of this drop among young age categories seems attributable to "doubling up"; as shown in Figure 6-2, almost half of 20–24-year-olds and a quarter of 25–29-year-olds postponed independent household formation and continued to live with their parents, as measured in 2010 ACS data. Even among 35–39-year-olds, 8 percent continued to live with parents—a

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

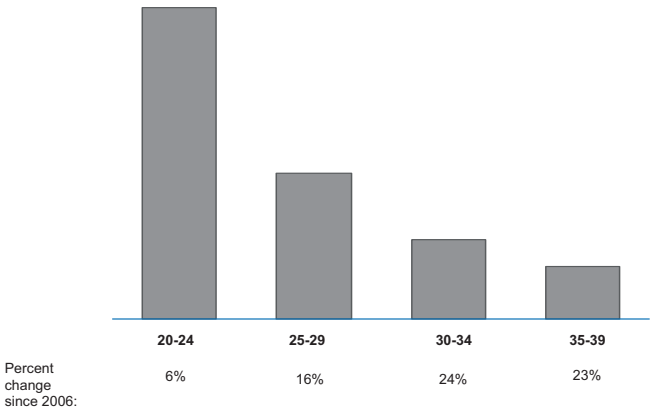


Figure 6-2 Percent of people living with parent by age, United States, 2010

SOURCES: Calculated from American Community Survey data; adapted from workshop presentation by Gad Levanon.

23 percent increase over 2006 estimates. The contrast with pre-recession numbers (2006) is particularly interesting and suggests the recession's role in driving housing decisions. Like the other analyses, these general observations raise interesting follow-up questions that can be addressed at the state level (or other aggregations) using the ACS, and this work permits one to see whether "doubling up" is particularly concentrated in states hardest hit by the recession.

Work with the housing data also occasioned Levanon to comment favorably on the variety of questions in the ACS. Even seemingly obscure ones like the number of rooms in the home[9] can yield interesting insights from the ACS data. In this case, Levanon observed that home sizes decreased during the recent housing crisis. Detached single family homes[10] with 5 or less rooms ticked up slightly from about 23 percent to about 29 percent of the stock of detached single family homes; apartments with 3 rooms or less rose from about 38 percent to just short of 50 percent between 2006–2009, though the percentage dropped (and so apartment sizes grew) between 2009 and 2010. He concluded that these kinds

[9]Housing Question 7 on the 2012 questionnaire; part a asks about the total number of rooms and part b asks for the number of bedrooms, "count[ing] as bedrooms those rooms you would list if this house, apartment, or mobile home were for sale or rent."

[10]Housing Question 1 on the 2012 ACS questionnaire asks about the general nature of the home, asking "Which best describes this building?"; responses include "a one-family house detached from any other house," "a building with 5 to 9 apartments," and "boat, RV, van, etc."

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

of shifts in housing are important to businesses and other parties interested in understanding the demand for housing in the nation or in specific regions.

6–E ACS IN BUSINESS: RISK ASSESSMENT AND INSURANCE

Headquartered in Boston, AIR Worldwide has been modeling the risks of natural catastrophes since 1987, and today does so in more than 90 countries. It is a member of Verisk Insurance Solutions at Verisk Analytics; as that label implies, many of AIR's clients are insurance and reinsurance companies seeking to understand and manage their risks, but AIR's client base also includes financial institutions and government entities. Cheryl Hayes, senior research manager at AIR, commented at the workshop about the ACS variables that AIR uses in its research projects and about the general importance of the ACS to the field of catastrophe modeling.

Hayes said that AIR began its work by developing models of hurricanes, tornadoes, and earthquakes and their attendant impacts and risks in 1987, but that catastrophe modeling did not come to the forefront until two extremely costly natural disasters, 1992's Hurricane Andrew and 1994's Northridge earthquake. The magnitude of damage and costs in those two disasters sparked awareness of the major impacts that could come from catastrophes, and companies began to realize that catastrophe modeling could help them better manage their financial and personnel risks. The industry, its models, and its methods evolved over the next decade, to be updated and expanded yet again after a major disaster, this time the terrorist attacks of September 11, 2001; models for terrorist behavior and workers' compensation were first introduced by AIR in 2002. By at least one metric, Hayes said, AIR and its clients have begun to see the benefits of catastrophe modeling; while 11 companies became insolvent as a direct result of Hurricane Andrew in 1992, only three such insolvencies was recorded following the record-breaking 2004 and 2005 hurricane seasons (punctuated by Hurricane Katrina in 2005; see Section 3–C).

The way in which ACS data enter this process is that they are crucial for developing the Industry Exposure Database that is a core component of a catastrophe model. An Industry Exposure Database is essentially a representation of the built environment that may be impacted by a catastrophe; it includes counts of buildings in a particular geographic area and information on characteristics of those buildings, such as their occupancy type, date of construction, and floor area. These variables contribute to another key feature of the database, which is the corresponding replacement values of those buildings. Hayes said that AIR built its first Industry Exposure Database for the United States when the company started in 1987 and has been updating it annually ever since—and, all along, the primary sources of information used to build and update the database are the decennial population censuses, the economic censuses, and, now, the ACS.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Elaborating on that point, Hayes sketched the general processes involved in constructing the database, beginning with the generation of risk counts, which are the numbers of dwellings and establishments. Information that they try to assemble for properties include the manner and type of construction—e.g., whether it is a high- or low-rise structure, and made of wood or concrete—because that information affects the potential vulnerability of the structure. In addition to occupancy type, the floor area of properties is a particularly important variable because it is essential to valuing the property. In addition to the censuses, this information (along with data on construction costs) is derived from housing surveys and from property costing and construction reports. The replacement/rebuild costs are calculated from the data on square footage, using different multipliers based on construction type and height; costs are also adjusted based on local and regional variation in costs of materials and labor. When these elements are combined and then benchmarked against other sources—data from clients and reports from the insurance industry—the result is a robust Industry Exposure Database.

Hayes said that the decennial census provides requisite information on the number of housing units—total housing counts. Displaying a map showing the percentage change in housing counts between the 2000 and 2010 censuses, Hayes noted that many of the states showing the highest rate of change in housing counts are the southeastern states—Florida through North Carolina—that are also prone to the specific disaster of hurricanes, which means that they have an increasing number of properties at risk of damage.[11] While the census sheds light on the contours and gross change in housing, the ACS is essential to generate detail on the characteristics of the housing. The specific housing items included in the ACS questionnaire are fairly blunt, but Hayes noted that many of them provide key clues for modeling:

- The type of structure, number of units in the building, and the year the structure was built all hint at the vulnerability of the structure; mobile homes are more likely to be damaged or destroyed by events like tornadoes than large complexes.
- Combined with other information from building permits or industry analyses, the year the structure was built can be used to infer things about the likely type of materials (e.g., whether masonry or wood was more likely to be used in particular states at a certain time).
- The ACS variable on the number of rooms in the house can be used as a rough proxy for the floor area of the housing unit (and structure), which in turn is a key part of the replacement value calculation.

[11] Texas, also on the Gulf Coast, also ranked in the high growth in housing rate category on Hayes' map; much of the Mountain West also experienced major growth in the number of housing units, and Nevada stands alone as the most extreme housing growth rate—and in the impact from the housing crisis and recession.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

- Other variables like per capita income and employment status have bearing on the valuation (replacement cost) of a property.

Of course, Hayes said, the ACS is not AIR's sole source of information, but the critical variable of floor area is a good example of the ways in which AIR finds the ACS essential to fill in gaps. She displayed a graph depicting the average residential structure for small areas across the country, based on data that AIR is able to acquire directly through various means. More to the point, the map also showed—shaded in dull grey—the areas in which the direct data are not available. The coverage of the direct data sources is spotty—weak in rural areas in the West but also surprisingly low for some midwestern and eastern states. But, Hayes said, the ACS data provide great assistance—using the ACS variable on the number of rooms in housing units as proxy for floor area, the same map drawn with ACS data has vastly fewer "no data" holes, and the direct measures can be compared with the ACS-based proxies for areas where they overlap in order to judge the quality of the proxy measurement.

In general, Hayes concluded, the ACS is vital to AIR's annual update and maintenance of a robust Industry Exposure Database and, once updated, that database drives a number of important analyses and modeling efforts. From an ongoing research standpoint, Hayes said that the exposure database is continually used along with historic loss estimates to develop, validate, and recalibrate its core catastrophe models. The models and the exposure database are used for the main thrust of AIR's work, part of which is to generate real-time estimates of losses—before events occur and as they are unfolding—to enable better planning and management of reserves. The data and methods are also used to validate the losses reported by individual companies—and so assess the quality of their own data—so that they can better manage their financial risk.

Hayes wrapped up her discussion by displaying a table showing AIR's estimates of the total insured value of properties in each state that borders the Atlantic Ocean or the Gulf of Mexico, along with the percentage of that value corresponding to coastal counties—and so at greater risk of hurricane damage.[12] Hayes said that table speaks to the importance of catastrophe modeling to companies writing insurance in areas like Florida—where AIR estimates that almost 80 percent of the insured property value is in coastal areas. Moreover, the development of these models over time has suggested rapid growth in this value—an annual increase in coastal counties' property valuations of roughly 7 percent. She ended by commending the ACS as an integral part of Industry Exposure Databases, and so argued that the continuance of the ACS is vital to the clients who rely on AIR's models to manage their risks. She conceded that many of AIR's clients "might not know it"—might not fully appreciate how big a role that Census Bureau and ACS data play in the modeling and their resulting abil-

[12] The same table is available at http://www.air-worldwide.com/_public/images/pdf/AIR2008_Coastline_at_Risk.pdf, which also spells out which counties are defined as "coastal."

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

ity to manage catastrophe-exposure risk—but that the survey's value is not lost on AIR and the insurance industry generally.

6–F ACS FROM THE CONSTRUCTION AND HOME BUILDING PERSPECTIVE, AND DISCUSSION

Moderating the discussion session, David Crowe (chief economist, National Association of Home Builders [NAHB]) said that he wanted to take the opportunity to make clear that the clientele and users in the housing industry value the ACS in the same way as other business sectors. He said that housing is one of the most distinctly different and variable commodities from place to place—the old saying stressing the importance of "location, location, location" has never been truer. As suggested in Hayes' presentation, Nevada experienced a staggering home construction boom—and has been hit accordingly hard by the subsequent bust—while other states (e.g., Texas) were not as hard hit. The way we know such things, Crowe said, is through data sources like the ACS. He observed that NAHB has made extensive use of ACS data—in analyses for its client builder firms to understand the differentiation in housing trends across space.

In addition to documenting the effects of the recent economic trends, Crowe said that the ACS has made possible analyses of important aspects of the broader real estate industry, and he offered two specific examples. The first is home remodeling, which Crowe noted has become a much more productive component of the real estate industry than new construction. ACS estimates were instrumental in NAHB's construction of county-level estimates of remodeling expenditures, which in turn permit NAHB clients to understand the strengths and weaknesses in this increasingly important segment of the marketplace. As another example, Crowe said that NAHB has used ACS data to derive "affordability indexes"—factors to understand how housing affordability varies across locations and by demographic groups.

Listening to the presentations in this session, Crowe said that the common thread is "the use of data to make appropriate and intelligent decisions"—nothing more elegant or sophisticated than that—but that each presentation is also "a good demonstration that [there] isn't an immediate transfer from the Census Bureau to the client." Intermediaries are needed to compile the information, to develop it further, to construct products of the necessary level of detail, and to provide assistance in handling and interpreting the numbers—all of which helps the clients make informed decisions.

Opening the question period, Crowe asked the speakers a form of the same overview question that had been raised in previous sessions: What would each speaker do, and what (if any) data could serve as a backstop, if the ACS were no more? Conrad answered that he and other economic development organizers

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

would have to turn to private data vendors—and hope that "they are somewhere in the ballpark with their numbers" relative to what is available in the ACS. However, he emphasized that he works with small communities, which are particularly limited in their budgets and so would sharply constrain their ability to acquire the data; the availability of ACS data as a public good is greatly important to users. Genadek offered that IPUMS users would likely turn to the CPS but—in the interest of continuing research across time—would have to hold out hope that the ACS content would return in the form of a decennial census long-form sample. Hayes—from the perspective of a company reliant on statistical modeling—said that AIR would have to do a lot more modeling on its own, pulling together historical data and doing whatever it could to model and project from that base. But, she emphasized, the loss of the ACS "would definitely put us at a big disadvantage." Levanon concurred with Genadek that, for the Conference Board's labor market research, the alternative would almost surely be the CPS—but, as noted in other presentations, the sample size limitations on the CPS are such that the CPS is a "*far* second-best" alternative in many applications. For detailed analysis of occupations like those that are now possible with the ACS, researchers like the Conference Board would face the stark options of either using larger aggregation groups or not doing projects at all, and that would be significantly damaging. The unique nature of Acxiom's data resources is such that Christenson said that one reaction would be to position the company's private records-data-based holdings as an alternative for some users' needs. But, as Christenson emphasized in his presentation, Acxiom depends on the ACS data as well; were the ACS to cease to exist, he said that Acxiom would have to treat the last release of ACS data like a decennial census data file: freeze it and continue to use it in its models going forward, but recognize that it will be less reliable over time. Crowe summarized, adding his and NAHB's viewpoint that what would result is less definitive decisions—or at least poorer information feeding decisions.

With the floor opened to questions, Lester Tsosie (Division of Economic Development, Navajo Nation) asked about using ACS data to study labor migration—whether it is possible to extract information on between-county migration or, ideally, between tribal reservation areas and surrounding areas, from the PUMS files or other tabulations. Conrad answered that he does not typically use PUMS data much due to the nature of his projects. Reiterating that the basic nature of economic development projects is telling a story with a variety of data sources, he said that he commonly uses the county-to-county migration data developed by the Internal Revenue Service based on annual tax returns and—specific to mobility within labor market areas—data from the Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) program. Scott Boggess (U.S. Census Bureau) commented that county-to-county migration tabulations from the ACS (based on the question on place of residence 1 year ago)

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

are available on the Census Bureau website—albeit not directly from the American FactFinder interface, where those tabulations do not naturally fit.[13]

Crowe asked another question of the set of presenters, asking them whether they think that their clients ultimately have a good sense of where the information comes from—do they know that the products being delivered to them are based on ACS data, or just general "census data" from the Census Bureau? Christenson answered first: "the term 'American Community Survey' confuses every single person I talk to, internally and externally." He sees very little awareness of exactly what the ACS is and that, historically, Acxiom and others wind up using "Census Bureau data" as a descriptor because the ACS label simply does not ring a bell. Hayes agreed, recalling that references to "ACS" were unclear to many, if not most, of the participants of an AIR client conference in April 2012; they, too, turned to general descriptions of "census data," though they also tried to introduce the concepts and terminology of the ACS. Conrad admitted to using "census data" as a descriptor, but more as a convenience; in the economic development arena, so many data sources and so many acronyms can crop up in products that "census data" clears some of the acronym clutter. However, he argued that—if users like himself and his fellow presenters rely on the data so much—it is incumbent on them to try to inform and educate clients. Another workshop participant added that "ACS is a brand name that nobody knows" on Capitol Hill; in conversations with constituents, congressional staff find that "census data" is less confusing. In short, Crowe summarized, what he concluded from this exchange is that—"to use an Acxiom phrase"—the ACS "needs branding." Christenson acknowledged this, and added that Acxiom itself is starting to label new releases of its Market Indices product as "Market Indices ACS," to start to build awareness.

Ken Hodges (Nielsen, and co-chair of the workshop steering committee) said that he wanted to interject a point that perhaps runs to the contrary. He recalled a session about 8 years ago, when his company was still known as Claritas, where he ran a session on the ACS at a well-attended client conference. At that time, awareness of the ACS and its basic features seemed to be quite high. Genadek added that academic researchers in general—such as make up a large share of the IPUMS user base—seem to have high awareness of the ACS nomenclature; in academic circles, at least, the "branding" of the ACS has gone fairly well. Crowe agreed that intensive data users know the ACS very well but that he is not surprised that end-stream clients of business users or data aggregators are not fully aware of the ACS. He said that the challenge, and the point that he wanted to raise in starting the line of questioning, is that if people do not know that "the ACS" is a valuable piece of information, "they will never come to the defense of it" if the survey is vulnerable. Christenson commented that the calls, emails, and questions he has received concerning the ACS have

[13] These migration data are available at http://www.census.gov/hhes/migration/data/acs.html.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

probably increased 10-fold with the legislative moves challenging the ACS (see Section 1–B); a possible "silver lining" of the recent increased scrutiny of the ACS is that awareness of the survey may grow as well.

Constance Citro (Commitee on National Statistics) asked Crowe and the presenters about the housing content of the ACS—what use is made of the questions on the ACS about mortgages and utility bills, because those questions are among the most difficult for respondents to answer. Conrad replied that he regularly makes use of the ACS data on average monthly rents and mortgage payments in characterizing the housing markets in smaller, rural areas. In those rural areas, realtor data—if they exist—-can be particularly sensitive to the role of realtors as sellers; the ACS measures might be less direct but may better reflect prevailing conditions than volatile under- or overstatement of values that might creep into realtor data. Crowe added that these ACS housing variables are important to NAHB's calculation of affordability indexes because they get at payment burdens and loan-to-value ratios. He added that NAHB has also made use of the ACS's data item on property taxes paid. The property tax rate in a particular locality is often stated as percent of the value of the home or property, but is often based on something other than the true market value of the property. He said that the ACS variables may provide a better measure of the actual value of the home and, thus, a better basis for comparison across areas. NAHB has published county and even subcounty maps to illustrate particularly high and low property tax rates, based on ACS data. Christenson added that, historically, his clients would most commonly request information on median income and median home value; now, his users are growing much more interested in the more-recently-added ACS questions on first and second mortgages and on health insurance questions.

Roderick Little (U.S. Census Bureau) asked the presenters for their views on the issue of voluntary versus mandatory response to the ACS, and whether they would see a voluntary ACS as having a major impact on their work. Genadek said that she agreed with the comments from presenters in other sessions—that mandatory is better than voluntary if only to stave off declines in the response rate (and effective sample size).[14] Conrad echoed Miller's comment earlier in the workshop (see Section 5–B) about the impacts being potentially worse for rural areas; repeating that Iowa has only 22 counties with populations 20,000 and larger, his bottom-line concern is that the loss of mandatory response could start the slide to exacerbating the divide between "data haves" and "data have-nots." In sum, his big concern is approaching a situation where, for Iowa, there will be a lot of information on Des Moines but nothing about the rest of the state. Citro concluded the session by summarizing the brief, hypothesized re-

[14]There followed a brief colloquy on exactly what the response rate to the ACS is—the response to the initial mail sample, the weighted response rates including follow-up collection by phone or personal interview, and so forth.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

sults: that voluntary response would cause mail response to decrease considerably, meaning that either additional funding would be needed to get back to the current effective sample size or data might have serious issues concerning bias and representativeness.

Copyright © National Academy of Sciences. All rights reserved.

– 7 –

Legal and Social Equity Uses of ACS Data

Important though information from the decennial census has become for applications ranging from public health surveillance to genealogy, the census is inherently political. Its constitutional mandate is to inform the redistribution of power through the apportionment of seats in the U.S. House of Representatives and, arguably, its most important secondary purpose is providing the raw material for the redrawing of political districts at the federal, state, and local levels. Inheriting as it does from the census and its previous long-form samples, the American Community Survey (ACS) likewise plays an important political role. ACS demographic data—and their detailed information on variables like primary language spoken at home, educational attainment, and physical or mental disability—play important roles in assessing the fairness of new districting plans and securing appropriate access to voting materials and public resources.

Section 7–A summarizes work in using the ACS to implement the Voting Rights Act's provisions for language services for the diverse population of New York City, while Section 7–B describes the ways in which ACS data have been used in legal cases challenging redistricting plans. Expanding the scope to include the ACS's role in monitoring social equity generally, Section 7–C summarizes work on studying disparate impacts in housing. To distill the themes from these presentations, the workshop steering committee asked Terri Ann Lowenthal—former staff director of the U.S. House of Representatives subcom-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

mittee with primary oversight of the census and current co-director of the Census Project—to moderate the discussion for this section (Section 7–D).[1]

7–A IMPLEMENTING VOTING RIGHTS ACTS LANGUAGE REQUIREMENTS IN QUEENS, NEW YORK CITY

In 2006—anticipating that the 2010 census would not include a long-form sample—the federal Voting Rights Act was amended to explicitly direct that ACS data be used to identify localities that must provide bilingual voting materials and assistance in areas with significant numbers of linguistically isolated households. Details of these requirements and examples from the first-time application of ACS data to these "Section 203" language-assistance requirements are described in Box 7-1. The mandate is etched in federal law but its implementation is—like the conduct of elections generally—inherently a function of state and local governments. As mentioned in the box, the New York City borough of Queens is obliged by the Census Bureau's ACS-based calculations to provide foreign language election services in four language categories. In his presentation, Joseph Salvo (director of the Population Division of the New York City Department of City Planning [DCP], and presenting joint work with colleague Peter Lobo) described the challenges of implementing these services for Queens' Asian Indian population.

Starting to operationalize the problem, Salvo said that DCP focused on what he called the CVLEP population—U.S. citizens of voting age (18 and over) who are limited English proficient (LEP). With that, the basic task may be restated as deploying language services to election poll sites with high concentrations of Asian Indian members of the CVLEP population. However simple that task might seem, stated in a few words, it is extremely complex on at least three conceptual lines:

- First, the "Asian Indian" category is far from homogeneous linguistically—India has nearly two dozen languages listed in its constitution. Queens' population of 2.3 million would rank it as the sixth largest U.S. city if treated on its own; Salvo added that almost half of the borough's population was foreign-born and that Queens was home to roughly half of the Asian population of New York City. The combination of the multiple languages under the "Asian Indian" category and Queens' size means that there are likely to be sizable pockets of several Asian Indian languages. Indeed, Salvo's analysis of 2010 ACS Public Use Microdata Sample (PUMS) data found many Asian Indian languages among Queens' CVLEP population; the top five such languages are Bengali, Panjabi, Hindi, Gujarati, and Urdu.

[1]One speaker on the agenda, Kimball Brace of Election Data Services, was to speak on the use of ACS data in the construction of districting plans, but was unable to attend.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Box 7-1 Legal Mandates: Determination of Areas Requiring Bilingual Voting Materials

In 1975, P.L. 94-73 amended the Voting Rights Act to prohibit discrimination of access to voting by providing English-only voting materials in areas with concentrations of non-English language minorities; these provisions were included in a new Section 203 to the act, so the bilingual assistance requirements have come to be known as "Section 203 requirements." The new law instructed the director of the Census Bureau to determine political subdivisions with such concentrations of language minorities or illiterate persons. Further amendment in 1992 (P.L. 109-246) provided additional detail on jurisdictions to be covered by these bilingual requirements and publish the same in the *Federal Register*, and 2006's P.L. 109-246 changed the generic reference to "census data" as the source of the determination to explicitly direct that the American Community Survey (ACS) data be used. At the same time, the legal change provided that these determinations be updated every 5 years, rather than the implicit once-a-decade update based on the census long-form sample.

The current law mandates that states or subdivisions thereof are "covered" (and hence must provide bilingual voting materials) if (excerpting from 42 USC § 1973aa-1a(b)(2)(A)):

> the Director of the Census determines, based on the 2010 American Community Survey census data and subsequent American Community Survey data in 5-year increments, or comparable census data, that—
>
> (i) (I) more than 5 percent of the citizens of voting age of such State or political subdivision are members of a single language minority and are limited-English proficient;
>
> (II) more than 10,000 of the citizens of voting age of such political subdivision are members of a single language minority and are limited-English proficient; or
>
> (III) in the case of a political subdivision that contains all or any part of an Indian reservation, more than 5 percent of the American Indian or Alaska Native citizens of voting age within the Indian reservation are members of a single language minority and are limited-English proficient; and
>
> (ii) the illiteracy rate of the citizens in the language minority as a group is higher than the national illiteracy rate.

(An entire state may be subject to Section 203 requirements under the 5 percent linguistic isolation threshold, but subdivisions under the state like a county are exempt from those requirements if their local share of that particular language minority group falls under 5 percent.) Under the same section of law, the director's determinations are held to be "effective upon publication in the *Federal Register* and shall not be subject to review in any court."

In October 2011, the Census Bureau published the first set of Section 203 determinations using 2005–2009 ACS data, listing covered states, county-level equivalents, and cities (76 FR 63602–63607); eligibility for Section 203 coverage was considered for 7,892 possible jurisdictions and 64 possible language minority groups (including Hispanic [Spanish-language assistance] and numerous American Indian and Alaska Native dialects). In addition to the new data source, the new determinations made use of a new methodological approach. As detailed by Joyce et al. (2012), the Bureau constructed hierarchical models after partitioning the data into a set of mutually exclusive "minority estimation groups," generating estimates using empirical Bayes methods.

(continued)

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Box 7-1 (continued)

As examples of some 2011 determinations:

- Three states were deemed to be covered under Section 203—California, Florida, and Texas, all for Hispanic/Spanish-language assistance.
- As further examined in Section 7–A, Queens County, New York, was deemed covered with respect to four groups (Hispanic, Asian–Asian Indian, Asian–Chinese, and Asian–Korean).
- As discussed in Section 5–D, the Navajo Nation covers a land area roughly the size of West Virginia, spreading across parts of Arizona, New Mexico, and Utah. In all, 3 Arizona counties, 7 New Mexico counties, and 1 Utah county are obliged by the Section 203 determinations to provide Navajo language assistance.
- The nation's most populous county—Los Angeles, California—qualified under 8 separate categories (Hispanic and 7 Asian languages, including "Other Asian—Not specified").
- Relative to the prior set of determinations (from 2002; 67 FR 48871–48877), Vietnamese became a covered group in three additional large counties. It remained covered in Los Angeles, Orange, and Santa Clara Counties, California, and Harris County, Texas, and became covered as of 2011 in Alameda and San Diego Counties, California, and King County, Washington. Similarly, coverage of Filipino language groups grew from 5 counties (Los Angeles, San Diego, and Santa Clara, California, and Honolulu and Maui, Hawaii) to 9 (adding two Alaskan boroughs [Aleutians East and West]; Alameda, California; and Clark, Nevada); one previously covered county, Kodiak Island Borough, Alaska, dropped out of coverage.
- One jurisdiction qualified for Section 203 bilingual coverage for Bangladeshi speakers—the city of Hamtramck, Michigan.

- One way to simplify the problem is to focus on the segment of the population that identifies "Asian Indian" as their race, because many who are not Asian Indian also speak Indian languages. Indeed, Salvo noted that his ACS PUMS tabulation found a roughly even split within the single largest Asian Indian language group in the CVLEP population: Roughly half of CVLEP Bengali speakers self-identified as Asian Indian race, but roughly half did not. Three of the top five language groups—Panjabi, Hindi, and Gujarati—had much higher levels of self-identification as Asian Indian race (80 percent or higher), but were not complete. Urdu, the national language of Pakistan, is also classified as an Asian Indian language, but only about 13 percent of Queens' Urdu-speaking CVLEP population checked the Asian Indian race category.[2]
- Because the basic objective is to provide *election* language services, it is natural to focus on the finest-grained operational unit of electoral geography: the poll site (the actual location at which the language services must

[2]Pakistani is one of the explicitly mentioned write-in examples under the "Other Asian" race category, two items below the "Asian Indian" category, in the ACS race question: Person Question 6 on the 2012 questionnaire.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

be deployed) and its covered jurisdiction. However, this electoral geography needs not—and does not—neatly correspond with statistical geography. Salvo noted that there are 669 census tracts and 316 poll sites in the borough of Queens; clearly, the arithmetic does not work to evenly match tracts to poll sites. In fact, Queens' poll sites and their coverages are defined so that—on average—a poll site is built from parts of about five census tracts.

Salvo described the entire Voting Rights Act compliance project as a function of three basic components: ACS data, administrative data, and input from the Asian community. Having outlined some of the conceptual complexity of the problem, Salvo said that he wanted to reveal the bottom line at the outset and then describe how DCP arrived there, based on findings from (and limitations of) each of the components. They made the simplifying assumption to base calculations and analysis on the population that self-reported Asian Indian as their race, and the ultimate decision was to provide—in designated poll sites—written assistance in Bengali and oral language assistance to all other Asian Indian groups in the Hindi language.

The first major task in this project was to pull tract-level estimates of the CVLEP population for Asian Indian language groups, using 5-year ACS data (2006–2010)—and this work experienced an immediate, curious snag. Census Bureau estimates for the tracts in Queens yielded tract-level CVLEP counts for Gujarati, Hindi, and Urdu speakers but not the two major language categories of Bengali and Panjabi; those languages were lumped into an "Other Indic" language category. Salvo said that DCP had to bump up to a higher level of geography—Public Use Microdata Areas (PUMAs)—and look at the percentages within the "Other Indic" category who speak Bengali or Panjabi. They then applied those percentages to the "Other Indic" counts in each of the tracts in a particular PUMA to approximate the share of Bengali and Panjabi speakers in each tract.

As a second major task—to corroborate these estimates—DCP requested a special tabulation from the Census Bureau: tract-level tabulation of CVLEP estimates for the Bengali, Panjabi, Hindi, Gujarati, and Urdu language group. They initiated the request cognizant of the "massive disclosure issues" associated with it, but felt the need to seek the tabulation to determine if their work was "grounded in reality." The agreement was far from perfect but it did suggest that DCP's approximations were in the ballpark. For instance, the special tabulation identified 18 Queens tracts with 50 or more CVLEP Bengali speakers while the department's approximate calculations from ACS data indicated 28 such tracts; the two sources agreed on 14 tracts and the 14 tracts flagged only by DCP's calculations were all geographically proximate to the 14 jointly agreed-upon tracts. This provided some reassurance that the ACS-based calculations could be useful in identifying small areas of interest.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

The third major task was to start the process of identifying needs at the poll site level, and was the point at which administrative data and input from the local community entered the mix. DCP sought lists of Asian Indian surnames compiled by various immigrant advocate groups, for each of the major Asian Indian language groups, and matched them to Queens voter rolls to generate counts of each language group by poll site. This provided auxiliary information on the concentration of CVLEP persons for each language group that could be contrasted with the ACS-based estimates—effectively, borrowing strength in finalizing estimates.

The fourth task was to confront the geographic mismatch mentioned at the outset: converting the ACS-based tract-level estimates to poll site levels. To do this, the Asian Indian CVLEP population by census block had to be obtained, which could then be aggregated perfectly into poll site areas. To get the CVLEP population by census block, a key assumption had to be made that the census block-to-census tract distribution of Asian Indian voters would also hold for the ACS CVLEP tract-level population.

The result of this task was, for each of the major Asian Indian language groups, poll-site-level estimates of the number of self-identified Asian Indians speaking that language who are also CVLEP. A map of those estimates could then be compared directly with the surname-and-voter-roll-based map for the final step: selecting the poll sites at which to supply language assistance. DCP settled on some uniform rules, to recommend that a poll site be identified for language assistance provision if it ranked high on both the ACS- and administrative data-based maps (say, for Bengali, a poll site with 50 CVLEP speakers in the ACS and 35 Bengali surnamed voters in the administrative data). Again, similar geographic clustering in both sets of data provided some reassurance—some "modicum of reliability"—that appropriate sites were being flagged.

The final details of the decision came about from input from the community. Bengali was the most common of the Asian Indian languages found in the data for Queens so it was a fairly natural selection. DCP learned from the Asian Indian community that Panjabi speakers commonly understand Hindi as well, Hindi being the national language of India. Because that permitted two of the major language groups to be combined, DCP reasoned that Bengali and Hindi would provide the most effective combination of languages to offer.

Salvo commented that "the great thing about this project is this—we are going to know whether this worked or not, because there's going to be hell to pay" if it does not. Time—and flak from reaction as voters go to the polls in November—will ultimately determine whether this analysis steered language assistance resources appropriately. Of the experience, Salvo said that this work is indicative of the challenges federal mandates place on localities. But, more positively, he views it as a good illustration of how political (anecdotal and qualitative), statistical, and administrative data can all be brought together to craft solutions that work. Moreover, he argued, it serves as an argument for more

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

model-based estimation using ACS data, up front by the Census Bureau, because it is what the local end users are having to do to fulfill their needs.

7–B ACS DATA IN REDISTRICTING STUDIES AND CHALLENGES

Criteria for the construction of political districts rank among the central protections in the federal Voting Rights Act. To preserve the right of protected, demographic minority groups to elect representatives of their choice, the act requires that so-called majority minority or supermajority districts be drawn in areas where there are sufficient concentrations of protected groups (rather than diluting their voice over multiple majority-driven districts). The act further specifies that general political redistricting lines cannot be drawn in ways that weaken the political power of the members of these protected groups. With that in mind, Jeanne Gobalet (Lapkoff and Gobalet Demographic Research, Inc.) began her remarks at the workshop by preemptively answering a question raised in several of the discussion sessions. She argued that, if the ACS were suddenly no longer available, "we will be back in the Dark Ages of the 1980s," when important legal and voting equity determinations had to be made more through guesswork than solid quantitative evidence.

In her applied demography work, Gobalet said that she has used ACS data for several years in a variety of political and legal settings: in housing discrimination cases, studying different kinds of housing tenure and occupancy rates; in review of jury selection systems, ensuring that juries chosen in trials truly represent their communities; in assessing fair access to education by examining private school enrollment rates within public school districts; and in challenges of racially polarized voting. However, with the time constraints of the workshop, Gobalet said that she would focus on only one area where ACS data have usefully been brought to bear: comparing and assessing different scenarios for political redistricting.

Specifically, Gobalet said that the three case studies described in her presentation share three basic characteristics:

- Each case involved a California county governed by a Board of Supervisors, with those supervisors elected by district; the districting scenarios all concerned the supervisor districts.
- Each county in question has a large Hispanic population, many of whom are immigrants.
- Each case depended critically on the special Citizen Voting Age Population (CVAP) tabulation prepared by the Census Bureau from 5-year ACS

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

data; two different vintages of CVAP tables have been produced (for 2005–2009 and 2006–2010).[3]

Procedurally, the work is complicated by the fact that the target districts are sufficiently small that census blocks are the necessary building blocks for supervisor districts—not higher-level geographic aggregates. So, Gobalet said, she goes through a process similar to what Salvo described for inferring finer detail from the available ACS data; specifically, she develops citizenship rates (by different racial and ethnic categories) at the census tract level from the CVAP tabulation, then applies those rates to voting-age-population counts—by census block—from the 2010 census. She added that she uses the CVAP tract-level data even though the CVAP releases estimates down to the block group level in order to reduce the problem of large margins of error.

The motivating question behind Gobalet's first case study is one of basic feasibility: *Can* districts with Hispanic majorities be drawn to properly comply with the Voting Rights Act? The county she used to illustrate the fundamental challenges of drawing such lines is Monterey County, including the city of Salinas. To answer the basic question, she laid out a two-step approach—map the Hispanic CVAP shares for various subareas of the county and then compare those rates with the corresponding share of the total voting age population that is Hispanic (without regard to citizenship). The first map (making use of the ACS's information on citizenship) might be thought of as the true eligible-to-vote map for Hispanics in the county, while the other—limited only to decennial census data provided in redistricting data files—reflects the information actually used to draw the lines. What Gobalet observed through this process is a negative correlation between Hispanic citizenship rates and Hispanic population shares—the higher the concentration of Hispanics in a census tract, the *lower* the citizenship rate (and thus the effective eligible voter population). She said that this demonstrated that—to preserve Hispanic voting rights—simply clearing the 50 percent threshold of Hispanic population in a district in this county would *not* ensure that the district would include a majority of Hispanics eligible to vote. An effective "majority minority" district in this case would have to have an extremely large Hispanic population share to meet the Voting Rights Act requirements—raising the possibility that such a district would look (on the surface) like an extreme case of "packing" the district.

For her second case study, Gobalet described a case in which her firm was engaged to examine four different redistricting scenarios that would partition a county (unnamed) into five supervisor districts; the task was to determine which of the scenarios comply with Voting Rights Act requirements, with the slight complication that Hispanic advocate groups had gravitated toward one particular plan where they sensed the plan would give them the ability to elect Hispanic

[3]The CVAP files and support documentation are available at https://www.census.gov/rdo/data/voting_age_population_by_citizenship_and_race_cvap.html.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

supervisors in two of the five districts. In this project, Gobalet said that she examined numerous data sources, but arguably the most compelling work came down to the type of analysis outlined before: work with tract-level information from the ACS CVAP tabulation to estimate the Hispanic CVAP population in each district, for each of the four plans. Gobalet said that the real difficulty in this work comes in calculating the margins of error for these districts, formed by piecing together Hispanic CVAP shares from smaller-level geographies. Ultimately, though, she said that she was able to derive the margins of error and thus compute 90 percent confidence intervals for the Hispanic CVAP share by district.

Each plan—arbitrarily labeled A, B, C, and D—shared some basic traits. Hispanics made up the majority of the total population in four of the five supervisor districts in all of the plans; indeed, all of the plans included at least one district that is at least 70 percent Hispanic. The same holds true restricting attention to the *voting-age* population—Hispanic majorities in four of five districts. But working with the ACS to add in consideration of citizenship (and thus baseline eligibility to vote), Gobalet found that only two of the plans—B and C—included an effective Hispanic majority district, in which the 90 percent confidence interval for the Hispanic CVAP share of the population lies completely above 50 percent. Both Plan B and C included a second district where the confidence interval overlapped 50 percent, suggesting that there might be a second majority Hispanic district (but not necessarily); only in Plan C did the point estimate for the Hispanic CVAP share in this second possible district exceed 50 percent. Neither Plan A nor Plan D could be said with confidence to have an effective majority Hispanic district; A had two districts with the Hispanic CVAP confidence interval overlapping 50 percent and D only had one.

Gobalet said that the "punchline" of the second case study underscores another reality in the application of the ACS to legal and political matters: namely, that data do not always dictate the decisions ultimately made. In this particular case, the Board of Supervisors ultimately chose Plan D, the one with arguably the weakest claim to any effective majority Hispanic district. Gobalet briefly mentioned follow-up research in this case that paralleled parts of Salvo's work—examination of surname lists from the county's voter rolls. Even with the large Hispanic shares of the total population and the voting age population for most of the districts in *all* of the plans, none of the districts—in any plan—has a majority of voters (registered and actual) with Hispanic surnames. Gobalet said that this work further reinforces the notion that proposed districts with very high Hispanic concentrations may not really have effective Hispanic majorities; it determined that none of the districting scenarios really have the effective Hispanic majorities that advocates had hoped would obtain for two of the five districts. Gobalet added that the county in question faces a strong possibility of Voting Rights Act litigation.

The third case study follows up from the methodology briefly mentioned

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

at the end of the second case study, in the interest of answering the question of whether local administrative data on voters can be used instead of or in addition to the ACS data in Voting Rights Act analyses. Returning the focus to Monterey County, California—a jurisdiction for which she has done a lot of redistricting work—Gobalet said that she has had the opportunity to do detailed geocoding and surname analysis of the voter rolls. She emphasized that surname analysis might be feasible for Spanish-origin names but is not believed to be reliable for identifying blocs of other minority groups; African American surnames are not distinctive enough to be singled out, and she said that the methodology has been found to have limited utility for identifying Asian American voters. But—if workable to identify Hispanic voters—the surname-based voter roll data has the attractive feature of being linked to precise addresses (and thus exact geographic locations) that can be linked to arbitrarily small geographic areas.

Gobalet said that her work with the Monterey County registered voter data has suggested fairly strong concordance between the 2006–2010 ACS CVAP tabulation and the surname-based Hispanic voter shares, with the latter computed separately from registered voter rolls as of November 2008 and the rolls as of November 2010. Monterey County is subject to U.S. Justice Department preclearance of districting plans under the Voting Rights Act, so she conducted this work in the context of evaluating two alternative districting plans partitioning the county into five districts—one developed by county authorities (and ultimately adopted) and a last-minute alternative proposal backed by the county's largest city, Salinas. Across both proposals and both vintages of the surname-based Hispanic shares, the Hispanic CVAP percentage and Spanish-surnamed registered voter share differed by no more than 5 percent for any district. She concluded that it is reassuring that both data sources seem to be measuring the same basic thing so, in some cases, it may not be necessary to rely solely on the Hispanic CVAP data from the ACS. That said, she reiterated that the CVAP data have more utility for examining concentration of minority groups other than Hispanics.

She closed by repeating her bottom-line conclusion from all three case studies—that the ACS-based CVAP special tabulation has become essential to ensure compliance with the letter and intent of the Voting Rights Act. Absent the ACS, things would necessarily return to the days of basing important determinations on guesswork and hunches.

7-C STUDYING DISPARATE IMPACTS IN HOUSING

Andrew Beveridge (Queens College and Graduate Center, City University of New York, and Social Explorer, Inc.) began his remarks by describing himself as "sort of an accidental demographer; I stumbled into this field without knowing too much about it, after graduate school." Getting involved in civil

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

rights issues and then serving as president of the Yonkers, New York, school board changed the focus of his work and research, and he has since been actively involved in using demographic data in a variety of areas under the broad heading of civil rights: redistricting challenges, jury selection systems, review of New York Police Department policies, and the topic that he elected to focus his workshop presentation upon, the study of housing disparity. He started his comments by observing that the standard tabulations from the ACS provide much information about the differences (and potential disparities) between various racial and ethnic groups, and analysis is even more powerful using the ACS PUMS data to tailor areas to the question of interest and choose how the tabulation is done; the frequency of ACS releases also make it preferable to the once-a-decade long form. Hence, as Gobalet did before him, Beveridge stated up front his view that—were the ACS to go away—the types and the richness of disparate impact analysis that are now possible for understanding housing conditions would be impossible.

Reviewing the established standards for demonstrating disparate impacts—not just in housing but also in other areas like employment discrimination—Beveridge emphasized that the key feature of such cases is that one need not prove *intent* to discriminate. The policy or action in question may be facially neutral yet still be found to have a disparate impact on minority groups if all three of the following are proven:

- The policy or action affects some members of some minority group more than the rest of the population.
- The housing referenced in the policy or action is more likely occupied—or more likely to be occupied—by minority groups.
- The building, failure to build, or destruction of housing like that covered in the policy or action is likely to have a disparate impact on such minority groups *in general*.

Beveridge said that the ACS is emerging as a particularly valuable tool in analyses on the second and third of these points.

Beveridge used as an example a recent filing in a case to block the construction of subsidized housing in Kenosha, Wisconsin. A developer proposed building low-income housing through the use of Low Income Housing Tax Credits (LIHTC) and the project was readily approved by the municipal government, but soon experienced serious public opposition. The city responded by blocking the project; the developer sued to force development, and so the question occurred: Are there disparate impacts associated with the city's blocking the new housing development? As additional context, defining the parameters of the study and the affected population, Beveridge noted that the LIHTC program is targeted toward construction of units serving people who make less than 60 percent of the median income in a particular area. The proposed new housing was also to include some units open to people with Section 8 public housing choice

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

vouchers, including the disabled, so the composition of the Section 8 waiting list would also become an issue.

Beveridge's analysis in this case made use of ACS data for the PUMAs in the city of Kenosha, for the surrounding Kenosha County, and the tri-county region comprised of Kenosha County and its immediate neighbors to the north and south (Racine County, Wisconsin, and Lake County, Illinois, respectively).[4] For exploratory purposes, he began by plotting a choropleth map shaded based on the percent of the population coded as non-Hispanic black; he overlaid on this map a layer of circles, centered on the location of existing LIHTC-funded projects in the area and drawn with the radius proportional to the number of units in the project. Zooming into the city of Kenosha, Beveridge observed that the proposed development site was in an area with moderate to strong minority populations but that is not already being served by similar LIHTC projects of any substantial size.

A next step in the analysis was to compare the composition of the Section 8 housing choice voucher waiting list for Kenosha County with the underlying demographics of the county as reflected in the 2005–2009 ACS PUMS data. Beveridge concluded that the disparate impact of getting into new Section 8 housing is fairly evident: the number of non-Hispanic blacks on the voucher wait list represents almost 75 percent of the non-Hispanic black population of Kenosha County as a whole, while wait-listed non-Hispanic whites make up just under 3 percent of the county's total non-Hispanic white population.

Those initial steps suggested the possibility of a disparate impact case, but fuller analysis requires examination of income data. The Census Bureau now regularly prepares a special tabulation from ACS data for the U.S. Department of Housing and Urban Development (HUD) known as the Comprehensive Housing Affordability Strategy (CHAS) data; earlier vintages of CHAS were prepared using 1990 and 2000 census long-form-sample data.[5] The CHAS data provide household counts (as fine grained as the place/city level) of households in need of housing assistance and with incomes that qualify for HUD programs like LIHTC—including shares of households that fall below various percentages of the prevailing area median income defined by HUD. For the Kenosha example, Beveridge looked at data for households meeting the two basic criteria to be eligible to move into the proposed housing development: household income not exceeding 60 percent of area median income and median rent not exceeding 30 percent of the area median income. Again, Beveridge concluded, evidence

[4]Kenosha County also shares shorter borders with Walworth County, Wisconsin, to the west and McHenry County, Illinois, to the southwest, and both those counties show up in some of Beveridge's analytical maps. Kenosha County and city is bordered on the east by Lake Michigan.

[5]To date, the Census Bureau has produced two iterations of CHAS data using 3-year ACS data (for 2005–2007 and 2006–2008); per HUD's website, HUD requested and the Census Bureau delivered in early 2012 a 5-year CHAS data file for 2005–2009. See http://www.huduser.org/portal/datasets/cp/CHAS/data_doc_chas.html.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

of disparate impacts are fairly clear; only a small percentage of Kenosha's non-Hispanic white population meets the eligibility criteria for the proposed development (roughly 8 percent), but just over 20 percent of the city's non-Hispanic black and Hispanic populations qualify. Similar disparity is evident when looking at the entire tri-county area.

A final piece of evidence in the case concerned the eligibility of disabled persons for entry into the proposed development; in addition to the income requirement (not exceeding 60 percent of median area income), households must have one member classified as disabled to be eligible for Section 8 vouchers. Running the eligibility criteria again, this time for disabled (technically, households qualified as disabled) versus nondisabled, the divide was again stark; Beveridge found that just over 25 percent of the nondisabled population in Kenosha met the eligibility requirements compared to nearly 50 percent of the disabled population.

Ultimately, Beveridge said, the disparate impact among African American and Hispanic households of *not* building the LIHTC housing became evident to all parties. The city of Kenosha and the developer agreed to settle the breach-of-contract case out of court, with the developer winning a multimillion-dollar settlement for damages and expenses from the city.

Beveridge concluded his remarks by briefly discussing his other work with census and ACS data, through Social Explorer—directing workshop users to the company's website at http://socialexplorer.com—and offering a few general comments on census and ACS data. First, having worked with *The New York Times* on several projects, he commented that he agreed with Ford Fessenden's approach to mapping (see Section 4–C)—to consider the meaning of small-count areas or gaps but not to let them interfere with telling a story with data, and to drop them if necessary. On a related note, and concurring with Fessenden's and Salvo's approaches (see Section 7–A), he recalled a favorite example in making the point that it is possible to become overly paralyzed by margins of error—even the "perfect," "real" data of the decennial census has curious anomalies. He said that he distinctly recalls getting a phone call shortly after the release of detailed data from the 2000 census, from a *Times* reporter with a curious question: "What's going on with the 59 people that live in Central Park?" The reporter had noticed, amidst the census block counts, that population total corresponding to the park; the reporter called the Census Bureau for clarification, and said that the Bureau staff did not have a good answer to the question. As it turns out, Beveridge said, there has been a small population credited to Central Park since the 1980 census, but there is not a very clear answer to the question of *why* they are counted there. A large part of it might be homeless people encountered during operations like 1990's S-Night (Shelter and Street Night) and 2000 and 2010's Service Based Enumeration program, but some might be from data swapping to preserve confidentiality.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

7-D THE LEGAL AND POLITICAL CLIMATE OF THE ACS, AND DISCUSSION

Moderating the discussion session, Terri Ann Lowenthal (Funders Census Initiative and the Census Project, and former staff director of the U.S. House subcommittee with oversight of the census) began by commending the presenters for what she called a strong set of presentations on issues that all, in some form, tie back to the original purpose of the census itself. The constitutional mandate for the decennial census is to fairly allocate representation through apportionment and, later, redistricting; through the long series of Supreme Court cases on redistricting and the enactment of related law like the Voting Rights Act, issues of ensuring social equity—fairness in representation writ large—have emerged as major uses of census data, so it is good to see the ACS data satisfying the same mandates.

That said, by way of summarizing both the session and what she had heard in the workshop as a whole, Lowenthal said that she had noted three "big disconnects" from the political and policy perspective that she said need to be considered going forward. First, Lowenthal observed a disconnect between ACS data users and the current political climate. She noted that several presentations and discussion sessions had ended in some "wish lists," some for more questions or topics, some for more frequent products, and many for a larger sample size. Though understandable and sensible *within* the community of data users, Lowenthal argued that users need to recognize that such wishes are "complete anathema to what is going on in the political environment right now." She argued that it is a major, looming communication challenge for all sides; ACS data users "have to hear" and listen to what ACS critics are saying (even though, Lowenthal senses, the critics are, "frankly, vastly in the minority") and understand their arguments. At the same time, data users have to recognize that the language of some of their wishes—"we really need a larger sample" mapping roughly to "we want more people to get that burdensome form that you don't want *anyone* to get"—are not arguments that will carry weight with critics.

The second "disconnect" Lowenthal observed is between the Census Bureau and the respondents to the ACS ("otherwise known as the American public"). She said that the Census Bureau has made efforts to explain why specific questions are asked in the ACS—and, to a lesser degree, how those questions and the answers to them benefit the public. But she said she is concerned that the Census Bureau is not answering the concern in quite the right way. She recalled that, during her time as staff director for the House census subcommittee in the mid- to late 1980s, the Census Bureau went about addressing the problem of justifying the questions on the census long form by compiling "The Notebook"—a binder that listed each long-form question and attempted to outline the legal and regulatory requirements behind the question. "The Notebook" has endured, and continued into the ACS era, albeit now in a digital form: roughly, "we collect

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

data, or publish data, on journey work because Title *X* or Section *X* of *this* act says that we need these data to implement *this* program." This information is all well and good, but Lowenthal argued that it is simply "too esoteric for the average citizen, and dare I say, for the average member of Congress." She said that specific examples of ACS use like the work presented at the workshop are likely to be "much more powerful" and persuasive to decision makers. In short, she said, she would encourage the Census Bureau and ACS stakeholders "to really talk in more basic terms" and concrete examples in discussing the need for the ACS.

Continuing, Lowenthal argued that ACS users—and the specific examples of ACS uses for important purposes—need to take care in how they justify some of the ACS questions that draw the most confusion and ire:

- Lowenthal observed that the questions "that are being ridiculed the most right now are those on disability." Read cold, absent any context, the language of the core disability question—about having "serious difficulty concentrating, remembering, or making decisions" or having "difficulty dressing or bathing"—comes across as "a really bizarre personal question," and it is unfortunately easy to lampoon the question as such.[6] What has been missing from the discussion, she argued, is a clear explanation that the question is not asked to pry into any individual person's life situation but rather to get information about various dimensions of disability—pieces that can be constructed to derive information on the disabled population generally—to make important policy decisions. Beveridge's example of the disability question being essential to judge fairness in housing policy is one of many others that could be then used to make a solid case for the question.
- Another frequently lampooned or criticized question is the one at the heart of journey-to-work data: "What time did this person usually leave home to go to work LAST WEEK?"[7] Absent context and good examples, the question is easy to criticize as trying to dupe respondents into revealing the best times to ransack their homes. Again, Lowenthal suggested, a missing step in the logic is that "they [the Census Bureau] really don't care when you leave the house and when you get home"; they're asking the question to get a sense of when roads in the vicinity might experience peak traffic and when transportation routes are being most (and least) utilized.

[6]See Person Question 18 on the 2012 questionnaire; the text excerpts from parts a and c of the question, and part b asks "Does this person have serious difficulty walking or climbing stairs?" Question 19 adds: "Because of a physical, mental, or emotional condition, does this person have difficulty doing errands alone such as visiting a doctor's office or shopping?"

[7]Person Question 33 on the 2012 ACS questionnaire.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

- Finally, some of the housing stock questions—the number of rooms and the number of bathrooms—are commonly criticized; Lowenthal said that what is missing is a rationale (with examples) for asking a series of questions rather than asking the overly blunt and likely misleading question "Is this house overcrowded?"—"because if you're mad at your wife that day, you just might say it is."

The third disconnect Lowenthal noted—and, she admitted, a hard one to resolve—is between Congress and the work that it does. She said that the Congress that asks "Why are we gathering these data?" and hammering the ACS as intrusive is the same Congress that periodically has to reauthorize programs in major surface transportation bills—seemingly missing the connection that the allocation of funds under those acts, and all of transportation plans that local and state governments have to submit, relies heavily on the journey-to-work data that are currently being collected systematically only by the ACS. The same phenomenon holds true in other policy domains as well—all the more reason, Lowenthal concluded, that clear examples of how ACS data (at fine levels of geography) that are essential to responding to congressional mandates are going to be "more powerful arguments for lawmakers than many other things that we can say."

Turning to specific questions, Lowenthal asked the panel of presenters the same question that was raised in other discussion sessions—what would you do, and how would your analyses change, if the ACS were no longer available. Gobalet's presentation had discussed the possibility of Hispanic surnames from voter registration rolls as one possible alternative data source, but—though those data were workable in the specific Monterey County example she examined—she concluded that the ACS data were more flexible and ultimately more useful. Accordingly, Gobalet answered Lowenthal that she would see no alternative but to go back to the 2000 census long-form data and to the "last," most recent iteration of ACS data. Lowenthal probed on that point and asked how long that would last; Gobalet answered that, at some point, it would be clear that those antique data are effectively useless, but that the basic argument that she would have to make is "these data are useless but they are all we have." Beveridge said that his sense of what would happen is that some of the commercial data vendors—"some of whom may even be in this room"—would derive products that are more model-based—necessarily involving more conjecture, being based in part on old data—and that there would be considerable uncertainty in exactly how the estimates are derived. Salvo said that his work with Lobo pushed ACS data to their limits, but that having those data gives him a distinct power: the ACS data are far from perfect, and no data are completely neutral, but the ACS data impart power by forming "a basis for discussion" and policy debate. He said that his concern is that, were the ACS to go away, that basis for discussion

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

would be lost, and the entire decision making process would become a function of politics.

On specific alternatives—the use of surname lists having been invoked—Salvo added that it was challenging to come up with a comprehensive list for Asian Indian languages, but the advocacy groups were up to the task. It is, he said, a task that "borders on insanity" and was tremendously difficult. But—just as Gobalet found for the Hispanic surname list in Monterey County—Salvo conceded that it was interesting that the process yielded good results, when "combined with other items and combined with reality checks on the ground." (Or at least, he hastened to add, the process yielded good results—with the "goodness" yet to be determined with the performance of poll sites in November.)

With the floor opened up to general questions and comment, Steve Murdock added a comment concurring with Gobalet about what would happen without the ACS. He recalled a time in the mid-1980s, working in Texas and being delivered a bunch of data from the Texas Department of Human Resources—"just incredible breakdowns," for all manner of demographic subgroups, for all 254 counties in the state. When he called the department and asked how they had arrived at what was presented as current and detailed data, they replied—essentially—that they had "made up population estimates for the counties and we used the same rates, or whatever, that were in the 1980 census." His concern is that the same form of crude extrapolation, from increasingly old and unreliable data, would be the alternative if the ACS were to be discontinued.

Lester Tsosie (Navajo Nation) asked the presenters to comment on the social equity uses of the ACS in light of the changing demographics of the nation, with longstanding racial and ethnic minorities becoming more significant segments of the total population. The state of New Mexico is already at or near "majority minority" status, with about half the population in the 2010 census being of Hispanic origin; do the presenters have a sense of implications—more injustice or less injustice—from changing demographics, first in states like Arizona and New Mexico and later in other parts of the nation, particularly if the ACS were weakened? Gobalet answered bluntly that her sense is that neither the letter nor the intent of the federal Voting Rights Act—as currently written and interpreted by the courts—can be enforced without ACS data. Beveridge replied to Gobalet's point that, unfortunately and ironically, "there are people who would be fine with that." That said, he agreed that—without the ACS data—it is simply impossible to accommodate the demographic changes Tsosie referred to with respect to voting.

Alan Zaslavsky (Harvard University) said that he was struck in this session, and in some earlier sessions, that the point is made about small counties in rural areas suffering from lack of ACS samples—even though, operationally, their rates are higher in some cases because of concern about producing estimates for local civil divisions. Given that much of the work described here involves very small areas (and very small areas within urban areas as well), are there impli-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

cations about sampling these different kinds of areas in being able to meet the needs for ACS data? Salvo agreed, recalling that he listened to Andrew Conrad's presentation about economic development uses in Iowa (see Section 6–A)—and its recitation about how few cities and counties in Iowa have populations over 20,000—with jaw agape; "we have neighborhoods in New York with 30,000 people, 40,000 people." So he conceded being a bit embarrassed to make things sound like an argument for an increased sample in a borough like Queens—but "the law is the law, and Section 203 implementation requires detail," and that is going to push the limits of the ACS sample even in the densest of areas. He agreed that more effective ways to sample in rural areas need to be considered—indeed, he said it would be a real death knell for the ACS if the sample in New York City were increased at the expense of rural areas, because the ACS would lose value as purely "an urban area/New York City thing." Lowenthal said that she wanted to raise and put on the table an idea that Ken Hodges (Nielsen) has raised in the past—instead of trying to increase the ACS sample "in fits and starts," try to find a way to convince Congress to link the sample size to some automatic measure (like the estimated number of housing units in the nation). She said that such incremental, automatic changes to sample size might be more palatable than getting approval for larger chunks of sample.

Struck by Lowenthal's description of "the most ridiculed question" on the ACS, Stephen Tordella (Decision Demographics, Inc.) said that he had to wonder if putting legislators on the spot might be an effective argument—countering ridicule of the disability question by asking, in reply, "Are you ridiculing the disabled?" Beveridge answered that his own reaction to Lowenthal's discussion of the disability question was that it would be a good idea to get disability advocates involved and engaged in discussions of the ACS. Likewise, with some other questions, he wondered whether it would be a good idea to get civil rights advocates more engaged in ACS-specific efforts. Lowenthal agreed with Beveridge, adding that the Leadership Conference on Civil and Human Rights is currently one of the most active organizations in support of the ACS and the census; she said that getting disability advocates more involved could bolster the rationale for the question, and is certainly consistent with resolving the disconnects she described in her opening remarks.

Copyright © National Academy of Sciences. All rights reserved.

– 8 –

The Burdens of the ACS, and Closing Discussion

As described in Section 1–A, the Workshop on the Benefits (and Burdens) of the American Community Survey (ACS) sought to emphasize the benefits of the ACS to a wide array of data users while also giving its burdens—its challenges and drawbacks—an honest and appropriate airing. The workshop presentations and discussions summarized in previous chapters reflect different aspects of burden, including the relative disadvantage rural areas confront in access to and accuracy of ACS estimates compared to more populous areas (Sections 5–B and 6–A), questions and concepts that do not completely mesh with pressing policy interests (Section 2–A), and restrictive data embargo protocols that can hinder the work of groups serving as "interpreters" of the data (Section 3–B). In addition to burden-related material in these and other individual presentations, the workshop steering committee devoted a separate session to issues of burden, assembling a small group to speak about a selection of important aspects or components of burden:

- The workshop was intended to focus on nonfederal users of ACS data, and so no speakers or applications from the federal executive agencies were included in the workshop program. Yet a major "burden" associated with the ACS is that it needs to *fill the role of the previous census long-form-sample data in informing general policy*. The committee felt that the U.S. Government Accountability Office (GAO) is uniquely positioned to speak to the application of ACS data to the full sweep of policy decisions and could also speak to the potential costs and benefits of a voluntary ACS. Ron

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Fecso summarized GAO perspectives on behalf of himself and GAO senior analyst Kathleen Padulchick (Section 8–A).

- Relatedly, and extending themes raised during the state, local, and tribal perspectives session (Chapter 5), the committee sought someone to speak about the challenges of *communicating estimates, and uncertainty, to state and local decision makers*. Accordingly, Warren Brown (Cornell University) discussed examples from his time as state demographer in Georgia (Section 8–B).
- As touched on in the newly swirling legislative discussion of a voluntary ACS (Section 1–B), the ACS is continually open to criticism along *privacy and confidentiality* lines; some questions on the survey—such as the presence or absence of a flush toilet in the housing unit or what time each person leaves home for work—are routinely challenged as being intrusive or tantamount to identity theft. Hence, the committee asked Barry Steinhardt (former associate director of the American Civil Liberties Union) for perspectives on how one privacy rights advocate reads the ACS (Section 8–C).
- Michigan state demographer Kenneth Darga was asked to comment on the *respondent burden* issues associated with the ACS—how long it takes to complete the questionnaire, whether the resulting information and data justify the "imposition" of the survey on a large annual sample, and—consistent with the new debate—how response to the ACS might change if the survey were made voluntary (Section 8–D).
- Finally, the committee asked longtime demographic consultant Stephen Tordella to play devil's advocate—to get a sense of and comment on the general *complaints and concerns about the ACS and its content* raised by respondents in the public, and how those complaints are registered with decision makers (Section 8–E).

At the workshop, each of the speakers gave a short opening statement before the floor was opened up to discussion; that round of questions and answers is summarized in Section 8–F.

This chapter is also an appropriate point to summarize the brief period of closing discussion for the workshop. The workshop steering committee invited Steve Murdock (Rice University)—former director of the U.S. Census Bureau and former Texas state demographer—to wrap up the workshop with brief remarks on what he heard at the workshop and on the prospects of the ACS. This discussion is summarized in Section 8–G.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

8–A MAINTAINING RELIABLE INFORMATION FOR POLICY ASSESSMENTS

Ron Fecso (U.S. Government Accountability Office [GAO], on behalf of himself and Kathleen Padulchick) began with a short introduction of GAO's role as an investigative arm of the legislative branch. Formerly known as the General Accounting Office, GAO's name was changed about 10 years ago to reflect that the core of its work had changed from primarily auditing programs to evaluating them, across many dimensions. Most of GAO's studies are done at the request of Congress—often 800 or more jobs a year, but Fecso said that the office is also uniquely positioned to occasionally scan the environment and inform Congress of developments and trends that it should know about—"and the ACS, believe me, is one where we try to get their ear as often as we can."

Fecso reiterated that the basic goal of GAO is to provide Congress with information that is "objective, fact-based, non-partisan"—"totally nonpolitical"—and that, as a result, GAO is extremely serious about the quality of the data that it uses in its reports. Though it most often relies on already available data, GAO will occasionally conduct its own data collections as studies warrant, including surveys of local governments and school systems, but it is not equipped for doing big population surveys. Fecso said that GAO makes a point of conducting a data reliability assessment on data sources available to it, be they public- or private-sector-generated sources, and that this assessment includes such factors as the competence of the source, the reasonableness of the resulting estimates, and the soundness of the methodology. If the data do not meet GAO's standards, "we cannot use them."

One concern as the ACS was launched was whether the estimates from the survey could satisfy all the functions and demands then placed on estimates from decennial census long-form samples. Fecso said that the breadth of applications to which GAO has used the ACS (after satisfying itself of the ACS's fitness to the task) testifies to the fact that the ACS has proven itself a desirable replacement. Some of these wide-ranging applications of the ACS in GAO studies (with their GAO report number, to facilitate easy reference to the studies at http://www.gao.gov) are:

- Study of veterans' housing characteristics and the affordability of rental housing among low-income veterans (GAO-07-1012); ACS data revealed a surprisingly large group of veteran renter households with low incomes (about 2.3 million) but ultimately suggested that—at the time—veterans were not significantly different from nonveteran renter households in having problems affording housing.
- Study of differences in educational attainment and income among detailed Asian and Pacific Islander demographic subgroups (GAO-07-925); ACS data suggested, for instance, higher propensity for Asian Indians and Chi-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

nese persons in the United States to hold college degrees than other subgroups like Vietnamese and Native Hawaiians.

- Evaluation of the National Flood Insurance Program (NFIP) and the possible financial impact of changes in policy and premium rates (GAO-09-20); in this case, ACS data on such variables as median household income and median value of owner-occupied homes were used to target a sample of counties for intensive case-study analysis—somewhat similar to AIR Worldwide's use of ACS data to characterize areas and properties at risk of catastrophes (see Section 6–E).
- Comparison of alternative methodologies for allocating grant monies for vocational rehabilitation programs across the states (GAO-09-798); the ACS data proved ideal for the analysis not only for its detailed information on a wide variety of disabilities—disability rates being a key factor in the need for vocational rehabilitation programs—but also its coverage of the nonhousehold "group quarters" population, including people living in group homes or specific rehabilitation facilities.
- Assessment of the possible relationship between limited English proficiency and *financial literacy*, or awareness of consumer finance issues (GAO-10-518); while the study required external data (besides the ACS) for the financial literacy component, GAO appreciated that the multiple questions on the ACS gave it great flexibility in defining and constructing its own standard definitions of "limited English proficiency" for small areas.
- Description of the characteristics of women in managerial positions in the workplace (GAO-10-892R); in a 2010 update of a report originally done in 2001 using Current Population Survey (CPS) data, GAO was able to use ACS data to examine pay and demographic differences between women and men in management positions—for a wider range of industries than was possible with the CPS data.
- Profiling the demographic and economic characteristics of one extremely specific industry, people working in early child care and education (GAO-12-248); among the interesting findings, about 93 percent of the workers in this field who have a bachelor's degree do *not* have a degree specifically in early childhood education.

Fecso noted that GAO has occasionally been called upon to examine and advise upon technical aspects of the ACS. GAO's legal staff was asked by Members of Congress to render an opinion on the legal justification for the ACS, and their resulting memo (GAO report B-289852) is still an important part of the ongoing voluntary-versus-mandatory-response debate because GAO concurred that the Census Bureau has the legal authority to make the ACS mandatory. In 2002, GAO's *The American Community Survey: Accuracy and Timeliness Issues*

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

(GAO-02-956R) forecast the decline in response rate—and with it the decline in accuracy—that could accompany a switch from mandatory to voluntary ACS collection. Most recently, earlier in 2012, GAO issued a report looking at the whole portfolio of federal household surveys and the role of the ACS, making the case that the ACS was successfully able to add a question on field of degree that permits the ACS to serve as the sampling basis for the National Science Foundation's National Survey of College Graduates but that—large though it is—the ACS's current sample size is too small for it to be piggybacked upon for other follow-up survey designs.

Fecso closed by offering brief comments on general ACS issues. He reiterated his, and GAO's, concern that a switch to voluntary collection could degrade the quality of the ACS estimates. Additional funding—and additional sample units—could get around the basic problem of a lower initial sample size, but he said that he would still be very worried about the biases that might be introduced by nonresponse. He acknowledged that the ACS does genuinely incur burdens as well as benefits—it can be a lengthy questionnaire to complete, and it can be a challenge to communicate. That said, he expressed hope that communication and education could provide reassurance to ACS respondents, first that only a very small percentage of the population receives the survey and, second, that completing the survey is positive civic participation (rather than rote civic duty). He said that he appreciates the privacy and confidentiality concerns associated with the survey, and the arguments about the questions being overly intrusive, but—in his assessment—the wealth of information that those seemingly invasive questions can provide outweighs the burden. To that extent, he said (similar to Terri Ann Lowenthal's discussion in Section 7–D) that concrete, accessible examples of ACS like those at the workshop should really be developed and spread around to bolster the case for the survey.

8–B TRADEOFFS: USING A FEDERAL SURVEY TO DRIVE STATE AND LOCAL GOVERNMENT DECISIONS

Warren Brown (Cornell University) began his remarks by commenting that a frequently raised argument in favor of the ACS—not one dwelt upon at all in this workshop, but invoked in other settings—is that it is essential to the federal government and specifically to the distribution of more than $450 billion of federal funds. Important—and true—though this might be, Brown said that he recognized that this argument is not very persuasive in some circles: There are those "who would like to kill the beast" altogether and feel strongly that such large sums should not be collected and redistributed in the first place. That said, he argued, there is a legitimate question underlying this counterargument—what are the proper roles of and relationships between federal and state governments?—in light of which it is useful to consider the role

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

of federally collected data like the ACS. "Even if we went back to the Articles of Confederation"—weak federal government and strong state and local governments—Brown argued that it would still be incumbent on the state and local governments to provide some services to their populations; would it or would it not be short-sighted to shun a federally collected data collection that provides detailed information on *all* states and localities?

Brown continued that he wanted to use his time to talk about the uses of ACS data by state and local governments in addressing the needs of their residents—and, in that light, the burdens that the ACS in its present form creates for state and local users. He said that the basic point he wanted to express is that the ACS creates tradeoffs for users—fix or relieve one burden, create or exacerbate another—and he said that he would briefly illustrate his point through an example of work done with the Georgia Division of Aging Services, during his 3-year period at the University of Georgia as state demographer.

Over his 35-year career, Brown said that his role has been to serve as a "data intermediary"—making the data from statistical agencies more useful to analysts, policy makers, and general staff members of state and local governments. In that capacity, he said that he has heard—extensively—-complaints about the inadequacies and burdens of government statistics. In the days of the decennial census long-form sample, the recurring complaint was that the data are so out of date as to be useless. When the ACS came online, the complaints changed dramatically. Some challenged the sheer volume of the data and found it to be almost *too* timely—the thrust of the complaint being that "this flow of information is overwhelming." Others found more subtle, technical grounds for complaints; officials from small-population areas with significant group quarters populations (e.g., prisons, college dormitories, nursing homes) critiqued the "erratic and inaccurate" nature of the ACS group quarters data.[1] Significantly, state and local government users also complained about the burdens of interpreting multiyear period estimates relative to the point-in-time estimates of the old long form and of grappling with more prominent margins of error; so many state and local programs have specific cutoffs or thresholds that estimates that look like they are bouncing above and below those thresholds are highly disconcerting.

With that as context, he moved to the example. The Georgia Division of Aging Services administers a variety of programs to deliver services to the elderly, to persons with disabilities, and specifically to military veterans with disabilities—the central objective of which is to enable people who wish to remain in their own homes (without having to be institutionalized) to do so. Consequently, the division has to coordinate extensively with a variety of service providers, to streamline delivery approaches so that the maximum of services

[1]To that end, Brown held up a copy of—and commended—the recent National Research Council (2012) report on measuring the group quarters population in the ACS for offering recommendations on resolving these problems.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

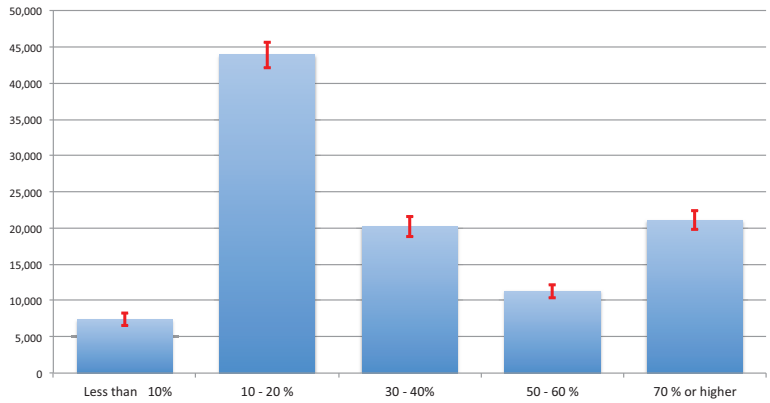


Figure 8-1 Veterans by service-connected disability rating, Georgia, 2008–2010

SOURCE: American Community Survey, 2008–2010, Table B21100; adapted from workshop presentation by Warren Brown.

can be provided with a minimum of cost—"a laudable goal," Brown said, and undoubtedly one that is replicated by similar agencies in other states as well.

When the Division of Aging Services came to him with a problem, Brown emphasized that they did so with tabulations in hand: "They used data from the American Community Survey; I didn't bring this data to them," or have to bring it to them. The heart of the problem is that one specific question on the ACS makes it, seemingly, ideal for the division's purposes. The U.S. Department of Veterans Affairs assigns service-connected disability ratings—expressed as a percentage in the sequence 0, 10, 20, through 100 percent—to affected veterans, and the ACS asks about those disability ratings.[2] To their thinking, the division hoped that the ACS would be able to provide them "statistically reliable information to provide more accurate assessments," Brown said; it seemed as though the ACS would be a very valuable tool in generating regional estimates of current and anticipated demand for services specifically aimed at veterans with disabilities. Indeed, Brown said, the division had used the ACS to explore the age of veterans, their living conditions, and their military service—all information collected in the ACS.

Figure 8-1 displays the estimates—and the upper and lower confidence limits

[2]Person Question 28a on the 2012 questionnaire asks "Does this person have a VA service-connected disability rating?" If the answer is "Yes (such as 0%, 10%, 20%, . . . , 100%)," then the respondent is asked to report the rating using the five categories implied in Figures 8-1 and 8-2: "0 percent," "10 or 20 percent," "30 or 40 percent," "50 or 60 percent," or "70 percent or higher."

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

on those estimates, shown as red bars—for the service-connected disability rating question, based on 3-year ACS data (2008–2010) for the state of Georgia as a whole. The confidence bands are quite tight—"but we don't deliver the program at the state level." Instead, the program for outreach to veterans with disabilities is administered through regional offices for the aging, so there is a need to drill down to finer levels of geography.

The top graph in Figure 8-2 shows the same service-connected disability estimates (again from 2008–2010 ACS data) for the largest geographic subpopulation in the state: the Atlanta metropolitan area, in which over half of Georgia's total population resides. Looking at the graph, Brown said, "you can do some program planning at that level"—"the confidence limits are still pretty acceptable," and there is still some clear separation in group sizes between the extremes of disability rating groups. But the bottom graph shows the results for the Macon metropolitan area—"not a small population," at around 250,000 total population, roughly 17,5000 veterans in total, and about 2,300 of those having a service-connected disability rating. And, there, the division came to Brown, wondering what to do. The confidence limits widen a great deal and, from this picture, Brown said that "it is difficult to justify how many persons in need" of the veteran-specific services are in the Macon area. Yet "this is the only information they have, to set those kinds of planning objectives."

Brown closed by arguing that what is needed for effective programming in state and local government programs is "reliable, consistent, accurate, precise estimates, as best we can get them." Brown said that the example of the disabled veterans brings home to him that the ACS creates, and will continue to create, tradeoffs. Making responses voluntary might benefit the privacy of individual responses, but the resulting diminution of the sample size could hurt the accuracy of estimates. Brown said that his worry is that these shifts in quality might disproportionately "create tangible negative consequences for some of our most vulnerable residents: youth, the elderly, and disabled veterans."

8–C INTRUSIVENESS AND PRIVACY CONCERNS

Barry Steinhardt (Friends of Privacy USA) opened his comments with an important disclaimer; some people in the workshop know him from his lengthy career with the American Civil Liberties Union (ACLU) and, in particular, from his service as ACLU's representative to the Census Advisory Committee. However, he emphasized that he does not represent or speak for the ACLU today. What he said he does speak from is the perspective of a long-time observer of the intersection between the census and privacy and confidentiality issues—in which capacity he asked to start by stating his "incredible respect" for the Census Bureau's privacy safeguards. In his assessment, the Census Bureau is the federal government agency with the most effective program for "protect-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

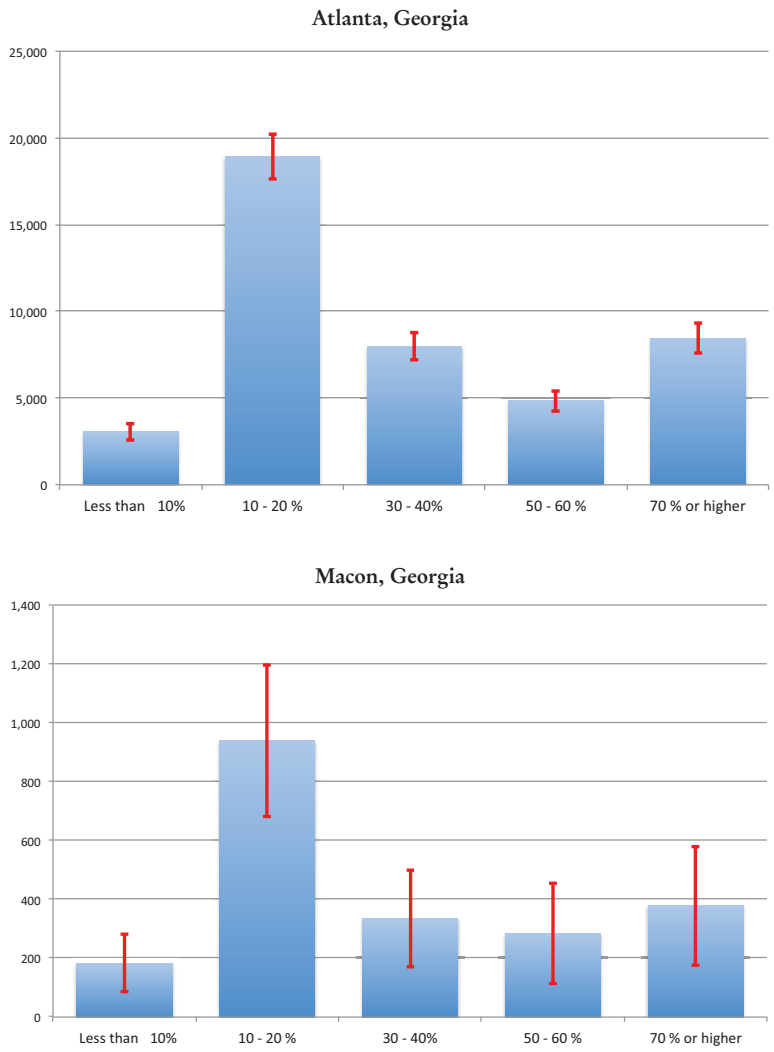


Figure 8-2 Veterans by service-connected disability rating, Atlanta and Macon, Georgia, metropolitan areas, 2008–2010

SOURCE: American Community Survey, 2008–2010, Table B21100; adapted from workshop presentation by Warren Brown.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

ing the privacy of the respondents and the integrity and the privacy of the data that it receives." Steinhardt added that he currently sits on the Privacy Advisory Committee of the U.S. Department of Homeland Security (DHS) and that "they largely roll their eyes when I set the Census Bureau as a model" for DHS to follow; still, he finds the contrast between the Census Bureau and DHS in terms of privacy protection to be remarkable.

Steinhardt also stated his conclusion that "it is very difficult and nearly impossible to glean personally identifiable information from aggregated census data," in large part because of the Census Bureau's well-developed techniques (including data swapping) for minimizing identifiable information. He knows of no example since World War II—the "infamous example" of disclosure of information involving Japanese Americans—of the Census Bureau willfully or inadvertently releasing personally identifiable information about an individual, which is an extremely commendable effort. Steinhardt did note that one might not be able to identify particular individuals in aggregate census data but one can tell when there are some outliers—far different in some characteristic than the norm, within some fairly discrete answer. How much information one can glean from outliers and whether marketers and others in the for-profit sector make use of them are open and interesting questions, Steinhardt observed.

With that said, and recognizing the many "extraordinarily valuable role[s]" played by the ACS, as illustrated by the presentations at the workshop, Steinhardt said that it has to be recognized that some of the questions on the ACS involve personally sensitive information. Moreover, many of the ACS questions are inherently highly personal in nature—Steinhardt said that there is "no other way to get the value out of the ACS *without* asking those questions," using the personal answers to produce estimates at aggregate levels.

So, getting to the way in which a privacy rights advocate reads the ACS, Steinhardt said that the first natural question is how well the Census Bureau protects the data. As he said before, "the answer is 'very well' "—the Census Bureau has a "sterling record of protecting the data." He recalled—from his tenure at the ACLU—actively urging people to fill out the decennial census form *and* the ACS questionnaire, saying that they could do so "with great confidence that their data would be protected." So, on that level, the ACS is not troubling.

However, the question of whether responses to the ACS should be voluntary or mandatory is another matter. Steinhardt said that, in his opinion, "it is a close constitutional question" as to whether the Census Bureau can compel responses to all of the questions on the ACS. To be clear, he continued, "I am hardly a constitutional originalist on this"—census practices have had to change, and have changed, since ratification of the Constitution and the conduct of the 1790 census, and he does not want to try to parse exactly what the founders could have possibly envisioned regarding the census. Moreover, there is "certainly a whole range of questions which I believe people have a constitutional obligation to answer." But, he suggested, "I think it is a stretch" to argue that questions of

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

"whether or not you heat your home with natural gas" and whether that natural gas is in a bottle or comes through a pipe[3] flow directly from the constitutional mandate to apportion the Congress.

In the voluntary-versus-mandatory debate, Steinhardt observed, it is frequently mentioned that the Census Bureau has never prosecuted anyone for failure to answer the questions. In the cases that he is aware of, where litigants have challenged the penalties for nonresponse, Steinhardt said that the courts have uniformly declined to hear the question; because the Census Bureau has not brought charges against anyone, the litigants have "no credible fear of prosecution" and so the courts rule that they lack standing to bring the case. In his opinion, Steinhardt said that he thinks "the courts are wrong in that," and that actual judicial examination of the issues would be very interesting—and "an extraordinarily close question about whether or not you can be required to answer *all* the questions."

Concluding his remarks, Steinhardt said that "if you ask me as a privacy advocate what I worry about," the mandatory-versus-voluntary issue is not at the top of the list. Again commending the Census Bureau's "sterling record" in resisting such entreaties, Steinhardt argued that a breach in the Census Bureau's protections and the misuse of census or ACS data by other elements in the federal government would be vastly more damaging to the enterprise of the ACS. He said that he appreciated hearing about the manifold uses of the ACS, that he thinks that the ACS plays an extremely valuable role, and that Congress should continue to support the ACS. He is, however, firm in his conclusion that the question of whether all of the ACS questions should be mandatory responses is a close constitutional question.

8-D IDENTIFYING (AND REDUCING) RESPONDENT BURDEN

The Paperwork Reduction Act of 1995 requires federal agencies like the Census Bureau to submit proposed "information collections"—essentially, any gathering of information from 10 or more respondents—to the U.S. Office of Management and Budget (OMB) for clearance, at least once every 3 years for an ongoing survey like the ACS. OMB, in turn, is asked to weigh—among other things—whether the demands that the collection puts on respondents' time are appropriate (and reduced to the extent possible). In its most recent request for clearance of the ACS (viewable by searching for Information Collection Review 201202-0607-003 on http://www.reginfo.gov), the Census Bureau estimated that the standard ACS housing unit questionnaire takes about 40 minutes to complete on the mailed paper form and about 27 minutes administered

[3]Housing Question 10 on the 2012 ACS questionnaires asks "Which FUEL is used MOST for heating this house, apartment, or mobile home?" The first two response categories, out of nine, are "Gas: from underground pipes serving the neighborhood" and "Gas: bottled, tank, or LP."

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

by phone or personal interview. (The actual time to complete the survey depends critically on the number of people in the household—and, with it, the number of person-level questions that must be answered.) Based on these assumptions, and including quality control interview and data collection from group quarters as well as households, the Census Bureau said that it anticipates the average annual respondent burden for 2012–2015 to be roughly 2,435,568 hours across 3,805,200 respondents—a large commitment of time and resources that invites continued and active discussion to keep that burden in check.

Asked to comment specifically on issues of respondent burden, Kenneth Darga (Michigan state demographer) framed the problem by noting that true respondent burden includes at least two major components: demands that the survey puts on respondents' time and effort (as described above) and the cost/burden of revealing personal information. He said that his comments are intended to place these components into context through comparison with other demands for time and effort and with disclosure of personal information.

He first asked the workshop audience to consider that various levels of government—federal, state, local—place myriad demands on individuals' time and effort (not to mention, in some cases, their money). Among these, the federal government demands compulsory military registration (and, in some wartime periods, compulsory service); it also directs that individuals wait in line for Transportation Security Administration screening in order to travel by air. Depending on the letter of the law in individual states, state government makes a strong, mandatory claim on young peoples' time: mandatory school attendance from anywhere from 10 to 14 years. And state and local governments both demand that people wait at traffic lights, and that they keep their lawn mowed and their sidewalks clear of snow in winter.

Against this backdrop, Darga argued, is the particular demand by the federal government to complete the ACS questionnaire—a demand, roughly speaking, that is made of a specified household "about once every 45 years." As has been said of the ACS, Darga commented that it is conceptually possible to make any of the other demands that he listed voluntary rather than mandatory—but, as Brown had suggested in his remarks, there are tradeoffs. If waiting at traffic lights were made voluntary, it could appear on the surface like giving people a freedom that they did not have before—yet, at the same time, Darga said that the move would work to take away or impair the special freedom of being able to drive safely. Likewise, as others have argued, it might appear that making the ACS voluntary would free up individuals' time, but it could introduce major costs in impeding good and relatively unbiased data. Darga reasoned that the ACS is the least of the demands that he listed, especially amortized over a 45-year period, but that the fundamental problem is that the ACS is a little-understood (and hence much-resented) demand; people instinctively understand why they have to stop at traffic lights or mow their lawns, but the reasons for completing the ACS questionnaire are less clear.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Darga added that, to be sure, the actual time and effort needed to complete the ACS varies considerably across households. Most fundamentally, most of the questions on the ACS are asked about each individual in the household; larger households necessarily require longer times to complete the questionnaire. Other factors that can inflate the time needed to complete the ACS questionnaire—in essence, Darga completed his analogy, mowing the lawn several times rather than just once—include language barriers and literacy issues. He said that he would shortly offer some suggestions for reducing this burden but that, in his mind, the relevant question is, essentially: Is completing the ACS as important as spending the same number (or more) minutes cutting grass, waiting at traffic lights, or—generally—doing any of the large number of other things that various governments require us to do for the public good?

Turning to the second broad component of burden, Darga echoed Steinhardt's conclusion that—for some people—some of the information called for by the ACS questionnaire can be sensitive. Someone's actual level of educational attainment might be less than most people think it is; racial or ethnic background might be a deep family secret; in some contexts, and in a climate of concern about identity theft or government intrusiveness, even the basic data item of a person's *name* might be deemed sensitive. On this point, Darga said that it is impossible for the Census Bureau to fully anticipate which data items might be deemed sensitive by which particular individuals, so its only natural response—a commendable one—is to zealously guard all personal data as confidential. Continuing, Darga suggested an important contrast with other demands for personal information: "when the [Internal Revenue Service] or an insurance company or a police officer wants information about you it is generally because they want to make some sort of decision about you." But the Census Bureau is fundamentally different in that what it really wants is aggregate information, not personal information—the catch being that it is not possible to arrive at those aggregate data without asking respondents inherently personal questions and then adding the responses together.

Under the broad heading of disclosing personal information, Darga said that he wanted to make two other points (and draw two additional contrasts between the ACS and other situations). One is that it should be acknowledged that some (but definitely not all) of the information requested by the ACS is already known and disclosed to many individuals and organizations. Tax records, employer files, driver records, local assessor records, conversation with neighbors, and so forth—Darga said that if the FBI or some other entity wants specific information about a particular person, they can get it from these kinds of sources. But the contrast with the ACS is important to bear in mind. These other data sources are well suited to the revelation of personal information but *not* producing the aggregate characteristics of an entity like a census tract, which is the ACS's strength. Another point is that information about specific individuals is not available from the ACS—and, arguably, would not be very useful if it were.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Using himself as an example, Darga said that "anyone who wants information about me can forget about getting it from the ACS," for two basic reasons:

- Procedurally, the confidentiality safeguards lauded by Steinhardt prevent browsing or look-up of individual information for purposes other than compiling statistics or evaluating the survey, and
- Logically, one could not *possibly* find information about Darga from the ACS for the simple reason that "my household has not gotten an ACS form yet"—the same statement that could be made by the vast majority of persons and households in the country.

Put more succinctly, Darga said that "if Big Brother wants information about me, he does not really need the ACS"—"his best bet is the Internet."

Darga concluded by offering four specific suggestions that could be considered for reducing respondent burden and alleviating privacy concerns:

- *Improve the layout of the ACS questionnaire to make it easier and quicker to report information for large families:* Darga observed that the current ACS questionnaire can seem intimidatingly long but that a major reason for that is cosmetic in nature—the same core of questions is repeated five times, for each member of the family. So, he said, the person responding for a large family "needs to read through four pages of questions five separate times"—frustrating for many people, and actually difficult to do in households where language or literacy concerns exist. Darga suggested that strategies for developing and deploying a large-family form—a columnar format where each question would only have to be read once and then answers for each person made across a row—could shorten the form and make it less imposing.
- *Make better provisions for complex households:* For living situations like co-op student housing or group homes, the ACS data collection task is one of collecting all the information of households of many unrelated residents—and, Darga said, it is tough to envision or rely on a "house secretary who is going to take responsibility for answering the ACS" on behalf of 20 unrelated residents. In these kinds of households, the Census Bureau's effective "all-or-nothing" approach to collection—assuming that the response will be automatically coordinated for everyone in the household—may not be effective. It might be worth considering asking the lead respondent to make the roster of people in the household but to permit checkboxes to be filled for people (or subhouseholds) where the Census Bureau would be better off making a separate contact (by mail or phone). This kind of approach could prove helpful in settings where people do not want to disclose personal information to a housemate or where some household members are unavailable when the (single) form is being filled out.
- *Suggest an alternative for people who do not want to provide their legal names:* Darga said that simply reporting one's name can take respondents

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

aback; they might fear identity theft, they might fear exposure of immigration status, or—for other reasons—they might not want to give their name to a government agency for any other purpose. So, he reasoned, one possible fix would be to take the person's name out of the equation. Structurally, Darga said that the primary roles played by person name in the current ACS are to simply "keep people straight while answering the questions" (e.g., to ensure that "Person 3" is actually the same person through the whole questionnaire), and to help follow-up workers clarify missing or contradictory information. It may be feasible to achieve both of those objectives by allowing respondents to use or claim a nickname or an alias; Darga added that deemphasizing the need for reporting names could help the Census Bureau make clear its interest in aggregate information rather than "building a master database of personal information on individuals."

- *Consider the use of an incentive to help reverse attitudes toward participating in the ACS:* Acknowledging that "it is probably not feasible to include a cash incentive payment in the Census Bureau's budget" for the ACS, Darga suggested "many politicians—and taxpayers—*do* like tax cuts." Darga suggested that some small tax credit for people who have submitted a complete ACS form could be a reasonable way to offset a sample household's time and effort in completing the survey. Indeed, he remarked, "we might even start to see people complaining that they *haven't* received an ACS form instead of complaining that they *have*."

8-E RESPONDENT COMPLAINTS AND CONGRESSIONAL REACTION

Acknowledging that he had been asked by the workshop planners to serve as a sort of "designated complainer," Stephen Tordella (Decision Demographics, Inc.) opened his remarks by stating that "the ACS really *is* a burden; there is no way of getting around that." But, he continued, his remarks are meant to underscore two basic points, the first being that the "burden" of the ACS is—and should be—-partially borne by Congress, one of the survey's most important stakeholders. As to the second, he recalled that he used to work for a sales-driven organization, in which the constant mantra was "nothing happens until somebody makes a sale." Tordella said that it is equally true that, for the census and the ACS, "nothing happens until people send in their responses." Hence, his second major point—that respondents should be viewed as "the most valuable commodity we have"—"they should be revered and treasured, not threatened," and they deserve more prominence in census and ACS operations.

First, specific to the burden argument, Tordella said that he had talked to a lot of people in the weeks before the workshop, to get their reactions to some of

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

the ACS questions. From those conversations, Tordella concluded that the data and the underlying estimates are certainly important but that even the advocates of the estimates have to concede that there is something ridiculous, something odd about asking people "when they left the house that morning or whether they have toilets." He said that his own wife's reactions—the morning of the workshop, previewing his comments—were telling. Asking her whether she knew about the question about toilets, she shrugged, and admitted that she did not understand why the Census Bureau would be asking about toilets. He followed up: "How about if somebody asked you what time you left the house this morning?" His wife's immediate response was that the question "seemed kind of creepy," and Tordella suggested that "there is no way that it will ever not seem creepy to some group of people."

No one knows the real number of complaints lodged against the ACS, Tordella said—certainly not the congressional staff members he talked to about the ACS. The Census Bureau "has a little better idea," keeping a record of correspondence that they receive, but even that misses the silent complaints—questionnaires not filed out of anger or aggravation. Still, he suggested that the correspondence suggests some of the flavor and the magnitude of complaints about the survey—even if 1 in 1,000 people complain, 1 out of 1,000 of the roughly 3 million households reached by the ACS each year, "that is still 3,000 complaints," from which insight can be gleaned.

From his interactions with Census Bureau staff, Tordella said that he received information about some broad categories of reactions to the ACS. The most voluminous of the complaints are those objecting to the perception of intrusive and invasive questions. He quoted from some of these complaints:

- "This is a lengthy questionnaire that asks very personal questions that are frankly no one's business."
- "Why would you ask, or need to know, what time I leave for work each day, and how long it takes me to get to work? Nor do I understand what my monthly bills have to do with it?"
- "I find this survey very invasive and really none of the Bureau's business."

A second broad category of complaints comes from people who did not return the mailed questionnaire, and so complain about the telephone and field visit follow-up steps. Again, Tordella recited some quotes from the correspondence to the Census Bureau:

- "I spoke with a Census rep last month and since then I have been getting five to ten calls a day from random people claiming to be employees of Census."
- "I called the number for assistance [and] was horrified at the rude and insulting and harassing and threatening language used by the person supposedly there to provide assistance. Naturally, this made me more suspicious of the survey of its intended purpose."

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

- "A worker came to my home and she was very hostile. She said that it was mandatory that we fill out the Census. We told her we already did the Census. She said we had to fill it out. We told her that we were under the impression that the Census was voluntary. She went on to say that there was personal information that was needed. We did report her to our local police department."

Following the theme of the last quote, the next set of complaints concerns the mandatory nature of response:

- "In which country am I living [if this is mandatory]?"
- "I am only responding so I do not have to pay a fine—but I am EXTREMELY uncomfortable about providing this information" (Tordella emphasized that "extremely" was written in all-capitals).
- "Being forced to complete a survey like this gives me the feeling of being a disenfranchised federal tax-paying citizen with no option but to shut up, fill out the survey used to distribute money that we don't have, and pay higher taxes."

Other complaints are fewer in number but can be particularly vocal and passionate. Tordella said that worries about whether the ACS questionnaire is really a scam seem to be particularly acute among the elderly and infirm; their surrogates or caretakers will make appeals and inquiries raising that concern. Other complaints question the constitutionality of the ACS and the degree of overlap between the ACS and other agencies' data sources. He closed the recitation of common complaints by reading a longer quote from one letter:

> I find this American Community Survey to be appalling, invasive, and intrusive, and none of the government's business and I intend to let my senators and congressmen know how I feel. Take note of that.

And, Tordella concluded, "a good number of them do." Even if it is 1 in 1,000, "these are legitimate and heartfelt beliefs," and that is the backdrop against which the ACS must operate. He said that these kinds of complaints are never going to go away completely—but neither are they a completely new phenomenon. He recalled being interviewed on Wisconsin Public Radio during the preparations for the 1980 census; reviewing some of the reported complaints about the ACS brought back memories of that day in Wisconsin because many of the same issues and claims of harassment by the census were raised then.

The dots are not hard to connect concerning the way these kinds of ACS complaints register on Capitol Hill, Tordella said. Congress is necessarily "part of the complaint bureau for the ACS"—their staff members field these kind of complaints and the ACS, being in continuous operation, spurs such complaints every week of the year—"and Congress controls your budget." To be sure, he added, Congress is a diverse body just as the American public is diverse, and so contains a range of views on the role of the government in the census and

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

ACS. Tordella noted his sincere personal belief that "there *should* be some people in Congress who hate the census and everything it is about—just to keep the Census Bureau on its toes." But there are many other members of Congress for whom the ACS only exists—to the extent that they recognize it—as a source of "continual pain," the source of those complaints from constituents.

From his conversations with congressional staff members, Tordella said that he concluded that there is a feeling out there that the Census Bureau is doing little to alleviate the perception problems with the ACS. "Press these congressional staffers about the idea of a voluntary ACS and the cost and quality implications, and their response is, '*You* fix it.'" Use cost savings from Internet data collection, or work out some new methodology, but "just go and fix it."

With that in mind, Tordella suggested that one possible solution—or at least "a place to start"—would be for the Census Bureau to recognize that "respondents really should be king." Newspapers have ombudsmen to take readers' perspectives in mind and challenge editorial approaches—Tordella asked "why shouldn't the *respondent* have an ombudsman" appointed at the Census Bureau? There are mechanisms within the Census Bureau that serve to protect respondents' rights—a chief privacy officer and a Disclosure Review Board, for instance—but those are little understood (or appreciated) by the public. A highly visible ombudsman and a citizen's advisory panel, "where people feel like they can be heard," could go a long way to improving perceptions of the ACS.

Tordella then briefly displayed some recent screenshots of the Census Bureau's homepage to suggest that "there should really be a lot made on the website about respondents." On the particular day he visited the site, one of the four "top stories" on a few-seconds rotation at the top of the page actually did speak to respondents: "Your Response Makes a Difference for Small Businesses" read the headline, above a link to more information about the 2012 Economic Census. That is good, but seemingly a one-shot deal—if you have received a questionnaire from the ACS or some other survey you have to scroll all the way to the bottom of the page and find, in small type, the link "Are You in a Survey?" in order to start having your questions answered. He argued that the placement of the link has probably been tested in some way—"Census tests most things"—but he concluded that it ought to be much easier for respondents to find supporting information, to find justification for the questions they are being asked, and to feel as though their concerns are heard.

He closed by nothing that these kinds of "respondent relations" steps might not alleviate all of the complaints, but that the Census Bureau could still benefit from them—and learn from the examples of other agencies. Arguably, he said, the federal agency "that has the most teeth and the most penalties and the most influence over the American people" is the Internal Revenue Service (IRS), which went through a substantial modernization in the 1990s. That modernization succeeded, he said, not solely through upgrades in information technology,

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

but also through concerted attention to taxpayer needs and concerns. These days, "even the IRS is kinder and gentler"—surely, he said, there must be ways for the Census Bureau to be kinder and gentler as well.

8–F DISCUSSION

In the general discussion session following the speakers' opening statements, Darga's suggestion of finding some way to incentivize ACS response drew a variety of reactions. Patrick Jankowski (Greater Houston Partnership) asked other session speakers to comment on that specific proposal, and Fecso said that he worried about the idea "snowballing to all the other federal surveys"—ultimately, the financial costs of the incentive might outweigh "what you are getting out" of the incentives in terms of good response. Recalling his immersion in the complaints expressed about the ACS, Tordella replied that people might look at an incentive program and ask "why are we giving away more tax dollars for nothing?"—an incentive might boost response, but it could also create more hard-set opposition to the survey. Brian Harris-Kojetin (U.S. Office of Management and Budget) reminded the workshop audience that—just the week before the workshop—the House of Representatives approved an amendment to an appropriations bill (albeit not the Commerce, Justice, and Science bill that funds the Census Bureau) prohibiting the use of money as a respondent incentive in a survey, so the legality of federal survey incentives is a matter of considerable current debate.[4] Dan Weinberg (U.S. Census Bureau) argued that even the idea of a tax credit, rather than a cash incentive, is infeasible because the Census Bureau would have to violate its own confidentiality provisions in order to tell the IRS who had completed the survey; Darga countered that some kind of stub or "receipt" from the ACS response could be attached with a tax return if the respondent wanted to claim the credit, but Weinberg argued that simply confirming a person or household's inclusion in the ACS could constitute a Title 13 violation.

Alan Zaskavsky (Harvard University) addressed a question to Brown and his specific example of estimating service-connected disability ratings among veterans (though, he noted, other speakers in the workshop sessions could address it in their own fields). Would clients or end users like the Division of Aging Services accept a move to much more model-based estimates? Such estimates might be smoother and have smaller intervals—at the risk of possibly looking

[4]On June 6, 2012, the House voted 355–51 in favor of an amendment to H.R. 5325, the Energy and Water appropriations bill for fiscal year 2013, offered by Rep. Scott Tipton (R-Colorado). Motivated specifically by $2 and $20 incentives offered in a 1,000-household Bureau of Reclamation survey on attitudes toward removing dams on the Klamath River, the amendment's language swept wider: "None of the funds made available by this Act may be used to conduct a survey in which money is included or provided for the benefit of the responder." The whole bill, as amended, was passed that same day, and awaits consideration in the Senate.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

more like a constant rate across the different disability rate categories. Put another way, is the problem for the division that the margins of error are so large and the intervals so wide, or is it that the statistical evidence does not differentiate much between the categories? Brown indicated that he thought such model-baed estimates could certainly win acceptance, and could be viewed as credible compared to estimates with much larger margins of error. Brown said that his sense is that people generally accept that the highly publicized monthly unemployment statistics are comprised of model-based estimates.

Terri Ann Lowenthal (Funders Census Initiative and the Census Project) and Steinhardt engaged in a colloquy over Steinhardt's assertion that the mandatory nature of the ACS is a close constitutional question. Lowenthal said that she had to "respectfully but firmly disagree" about mandatory response being a close call; she mentioned discussions in the first Congress about the content of the 1790 census (asking for more than a simple headcount) and court cases upholding the constitutionality of the content on the previous census long form. In her mind, the constitutionality of mandatory response "seems to be on pretty solid ground." Steinhardt replied that there are legal scholars who see the question as similarly clear-cut—in the opposite direction. He said that what he finds significant is that the courts have never really addressed the issue head-on—that is, dismissing cases for lack of standing is not quite the same as strongly supporting the constitutionality of the survey content. He reiterated his comment from his opening statement that "the level of detail in the ACS"—down to the method of delivery of natural gas—"is such that I think it is a close question." Lowenthal answered that she saw Steinhardt's point but sees things in a different way. She said that the issue is not really whether the framers could have possibly envisioned specific questions, but the more general issue of "whether the government has a right" and the authority "to gather information for the public good."

8-G WORKSHOP CLOSING REMARKS, AND DISCUSSION

Asked for his comments on the main topics of the workshop and the prospects of the ACS—from his perspective as a long-time researcher, as former Texas demographer, and as former director of the U.S. Census Bureau—Steve Murdock began with a slight apology for not being able to sit in on more of the workshop. However, he said, one of the reasons that he could not be present for longer was a previous speaking engagement for officers of the American Association for Affirmative Action. When those officers learned where Murdock was heading next, he said that several asked him to "please let everybody know how important the ACS is to us—we hear it's under attack in Congress."

Murdock said that the ACS is "critical to what the country does for decision making," and he commended the workshop speakers for providing very

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

clear indications of the value of the data in a wide range of applications. He said that he wanted to make two basic points in his short remarks, the first of which is that the ACS sample size is unlikely to ever be so large as to completely mitigate the challenges that confront users of small-area ACS estimates. He described preparing a speech in a community of about 25,000 people, and he said that he commonly prepares standard demographic profiles—from the ACS—for such speeches, together with his Rice University colleague George Hough. One of the standard figures that he likes to display is change in poverty rates over time, by race. In this particular instance, the 1990 and 2000 census data both showed the percent of Hispanic households in poverty in this community as being roughly 36 percent; generating the same variable using 2005–2009 ACS data, the figure was 11 percent. Clearly, something seemed off, and Murdock said that he decided not to use the city-level figures in his talk—and he did not. But, "sure enough," he was challenged during discussion by someone in the audience who noted that he had shown these data for the nation as a whole, for the state, for large metropolitan areas, and that he had presented all manner of *other* data for the particular community, but not that specific poverty figure. As it turned out, this particular questioner worked for the local mayor and approached Murdock after the talk, saying that the local government was currently trying to figure out how to allocate funds to various community groups—she asked whether they could cut the distribution to Hispanics in the community because the ACS shows so many few Hispanics in poverty, and Murdock explained that this probably was not the case. The questioner replied, "Oh, so you're saying these data aren't really all that good?," and Murdock tried to explain that the numbers have to be understood in a broader context.

Murdock explained that he brought up this story because perceptions matter a great deal; consistent with Tordella's earlier remarks, Murdock said that it does not take a majority to create major problems for the ACS—single, vocal complaints from key stakeholders can do just as much damage.

Acknowledging that it might sound to some "like I'm a traitor to the cause"—"I am not"—Murdock said that his second main point is that it is prudent for the survey's stakeholders to give serious thought "about what happens if our defense of the ACS Alamo doesn't work." The history of the battle at the Alamo suggests that many of the casualties at the fort "died so bravely because there was no way out." Murdock emphasized that he "has absolutely no doubt in the utility of the ACS data"; having led the Bureau, he also has "nothing but complete and total admiration for the Census Bureau staff that does this work." Nonetheless, he suggested that stakeholders need to "start talking quietly, constructively," about viable alternatives and about managing the tradeoffs that might come through alternative sampling sizes or aggregations of periods over time. He closed by expressing his hope that this workshop could be followed up with one that asks: "What would you do? What *can* we do? What can we suggest?" Work should certainly continue to support the current ACS de-

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

sign, and hopefully it can succeed, but he suggested that, unfortunately, "noble causes and noble rationales don't always work politically."

At the end of Murdock's remarks, steering committee co-chair Linda Gage (California Department of Finance, retired) asked whether anyone from the Census Bureau had any summary comments that they would like to offer. Jim Treat (chief, American Community Survey Office, U.S. Census Bureau) said that he came to the workshop with "high expectations of what I was going to hear," and that those expectations had been exceeded by all the presentations and discussions. He said that he and his colleagues have a great deal of information to take back—examples of specific data uses to flesh out and emphasize, challenges in communication to work on—and he thanked the presenters and participants for their efforts. Constance Citro (Committee on National Statistics) added that the day and a half had been "fantastic," outlining the breadth and depth of uses of ACS data—alone and in combination with other data sources—for a wide range of policy areas. She noted that the workshop would contribute, in part, to the work of a new National Research Council panel on ACS technical issues, and credited the Census Bureau with taking a hard, open look at the whole ACS program and finding ways "where it can be focused, improved, and made more useful." Like Murdock, she said that she did "not want to be defeatist in any way," but that she agreed that ongoing examination of the ACS demands serious alternatives—ways to make the ACS "as cost-effective as possible in a very tough political environment."

Copyright © National Academy of Sciences. All rights reserved.

References

Council on Asian Pacific Minnesotans (2012). *State of the Asian Pacific Minnesotans*. St. Paul, MN: Council on Asian Pacific Minnesotans. Accessed: http://www.capm.state.mn.us/pub_demo.html [August 2012].

Herriot, R., D. Bateman, and W. F. McCarthy (1989). The Decade Census Program—A new approach for meeting the nation's needs for sub-national data. In *Proceedings of the Social Statistics Section, American Statistical Association*, pp. 351–355. Alexandria, VA: American Statistical Association.

Joyce, P. M., D. Malec, R. J. A. Little, and A. Gilary (2012). *Statistical Modeling Methodology for the Voting Rights Act Section 203 Language Assistance Determinations*. Revised from January 24, 2012, original. Research Report Series, Statistics #2012-02. Washington, DC: U.S. Census Bureau. Accessed: www.census.gov/srd/papers/pdf/rrs2012-02.pdf [August 2012].

National Cooperative Highway Research Program (2007). *A Guidebook for Using American Community Survey Data for Transportation Planning*. NCHRP Report 588. Washington, DC: Transportation Research Board.

National Cooperative Highway Research Program (2011, August). *Producing Transportation Data Products from the American Community Survey That Comply with Disclosure Rules*. NCHRP Web-Only Document 180; contractor's final report for NCHRP Project 08-79. Washington, DC: Transportation Research Board.

National Research Council (1995). *Measuring Poverty: A New Approach*. Panel on Poverty and Family Assistance: Concepts, Information Needs, and Measurement Methods, Constance F. Citro and Robert T. Michael, eds., Committee on National Statistics. Washington, DC: National Academy Press.

National Research Council (2006). *Once, Only Once, and in the Right Place: Residence Rules in the Decennial Census*. Panel on Residence Rules in the Decennial Census, Daniel L. Cork and Paul R. Voss, eds., Committee on National Statistics, Division

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

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

National Research Council (2007a). *Tools and Methods for Estimating Population at Risk from Natural Disasters and Complex Humanitarian Crises*. Committee on the Effective Use of Data, Methodologies, and Technologies to Estimate Subnational Populations at Risk. Board on Earth Sciences and Resources, Division on Earth and Life Studies, and Committee on Population, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

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

National Research Council (2009). *Vital Statistics: Summary of a Workshop*. Michael J. Siri and Daniel L. Cork, rapporteurs, Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (2012). *Small Populations, Large Effects: Improving the Measurement of the Group Quarters Population in the American Community Survey*. Panel on Statistical Methods for Measuring the Group Quarters Population in the American Community Survey. Paul R. Voss and Krisztina Marton, eds., Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

New York City Center for Economic Opportunity (2012, April). *The CEO Poverty Measure, 2005–2010*. Working Paper. New York City: New York City Center for Economic Opportunity. Accessed: http://www.nyc.gov/html/ceo/downloads/pdf/CEO_Poverty_Measure_April_16.pdf [July 2012].

Pisarski, A. E. (2006). *Commuting in America III: The Third National Report on Commuting Patterns and Trends*. National Cooperative Highway Safety Program Report 550 and Transit Cooperative Research Program 110. Washington, DC: Transportation Research Board.

Schwede, L. (2003, August 27). *Complex Households and Relationships in the Decennial Census and in Ethnographic Studies of Six Race/Ethnic Groups*. Census 2000 Testing, Experimentation, and Evaluation Program Report. Washington, DC: U.S. Census Bureau.

Social IMPACT Research Center (2011). *2011 Report on Illinois Poverty*. Chicago: Heartland Alliance. Accessed: Interactive form at http://www.ilpovertyreport.org and PDF format at http://ilpovertyreport.org/sites/default/files/uploads/PR11_Report_FINAL2.pdf [July 2012].

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

State Health Access Data Assistance Center (2007). *Bridging the Gap Between Research and Policy: SHADAC, 2001–2006*. Minneapolis: State Health Access Data Assistance Center, University of Minnesota. Accessed: http://www.shadac.org/files/BridgingGapReport_2007.pdf [July 2012].

U.S. Census Bureau (2006, October). *Design and Methodology*. Current Population Survey Technical Paper 66. Washington, DC: U.S. Census Bureau.

U.S. Census Bureau (2008a, November). *A Compass for Understanding and Using American Community Survey Data: What Congress Needs to Know*. Drafted by Terri Ann Lowenthal and Mary Jo Hoeksema. Washington, DC: U.S. Government Printing Office.

U.S. Census Bureau (2008b, December). *A Compass for Understanding and Using American Community Survey Data: What High School Teachers Need to Know*. Drafted by William P. O'Hare, Linda A. Jacobsen, and Mark Mather. Washington, DC: U.S. Government Printing Office.

U.S. Census Bureau (2008c, October). *A Compass for Understanding and Using American Community Survey Data: What the Business Community Needs to Know*. Drafted by Ken Hodges. Washington, DC: U.S. Government Printing Office.

U.S. Census Bureau (2009, February). *A Compass for Understanding and Using American Community Survey Data: What State and Local Governments Need to Know*. Drafted by Joseph J. Salvo, Arun Peter Lobo, and Joel A. Alvarez. Washington, DC: U.S. Government Printing Office.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

Appendixes

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

– A –

Workshop Agenda and Participant List

AGENDA

Workshop on the Benefits (and Burdens) of the American Community Survey
June 14–15, 2012 • Washington, DC
Room 100, Keck Center of the National Academies

DAY 1: Thursday, June 14, 2012

1:00pm Call to Order
Linda Gage, *Co-Chair; California Department of Finance*
Ken Hodges, *Co-Chair; Nielsen, Ithaca, NY*
Brief Introduction of Workshop and Participants

1:15pm A. How Will the American Community Survey (ACS) Help Us Plan Health Care and Transportation in the Future?
Moderator/Organizer: Joan Naymark, *Independent Consultant, Minneapolis, MN*

1:15 Kathleen Thiede Call, *School of Public Health and State Health Access Data Assistance Center, University of Minnesota*

1:30 James Stark, *New York City Department of Health and Mental Hygiene*

1:45 Richard Thomas, *Center for Population Studies, University of Mississippi and Health and Performance Resources, Memphis, TN*

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

2:00 Beth Jarosz, *San Diego Association of Governments*
2:15 Vincent Sanders, *Metropolitan Transit Authority, Houston, TX*

2:30pm *Break*

2:45pm B. How Will the ACS Help Us Respond to Disasters and Plan Social Services in the Future?
Moderator/Organizer: Richard Rathge, *North Dakota State University*
2:45 Linda Giannarelli, *Urban Institute, Washington, DC*
3:00 Amy Terpstra, *Social IMPACT Research Center at Heartland Alliance, Chicago*
3:15 Allison Plyer, *Greater New Orleans Community Data Center*
3:30 Russ Paulsen, *American Red Cross National Headquarters, Washington, DC*

3:45pm C. ACS and the Media: Communicating Estimates (and Uncertainty) to the Public
Moderator/Organizer: Ken Hodges, *Nielsen, Ithaca, NY*
3:45 Haya El Nasser, *USA TODAY*
4:05 Ronald Campbell, *The Orange County Register*
4:25 Ford Fessenden, *The New York Times*

4:45pm Floor Discussion for Day 1 Sessions

5:00pm *Adjourn Day 1*

DAY 2: Friday, June 15, 2012

8:45am Call to Order and Reintroduction to the Workshop

9:00am D. The Burdens of the ACS: A Panel Discussion of Respondent Burden, Privacy/Confidentiality Concerns, and User Burden
("Opening Statements" of up to 10 minutes, plus discussion)
Moderator/Organizer: Linda Jacobsen, *Population Reference Bureau*
Ronald Fecso, *U.S. Government Accountability Office, Washington, DC*
Warren Brown, *Cornell Institute for Social and Economic Research, Cornell University*
Barry Steinhardt, *Friends of Privacy USA, Washington, DC*

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Kenneth Darga, *Michigan Department of Technology, Management, and Budget, and Michigan State Demographer*
Stephen Tordella, *Decision Demographics, Arlington, VA*

10:15am *Break*

10:30am E. State/Local/Tribal and Urban/Rural Uses of American Community Survey (ACS) Data
Moderator/Organizer: Richard Rathge, *North Dakota State University*
10:30 Susan Brower, *Minnesota State Data Center, Department of Administration, St. Paul*
10:45 Kathleen Miller, *Rural Policy Research Institute, University of Missouri, Columbia*
11:00 Steven Romalewski, *CUNY Mapping Service at the Center for Urban Research, City University of New York*
11:15 Lester Tsosie, *Division of Economic Development, The Navajo Nation, Window Rock, AZ*
11:30 *Discussion Leader:* Jacqueline Byers, *National Association of Counties*

11:45am *Working Lunch* to continue discussion of presentations; third floor cafeteria

1:00pm F. Business, Economic Development, and Data Aggregator Uses of ACS Data
Moderator/Organizer: Patrick Jankowski, *Greater Houston Partnership*
1:00 Andrew Conrad, *Business and Community Services and the Institute for Decision Making, University of Northern Iowa*
1:15 Katie Genadek, *IPUMS Project, University of Minnesota*
1:30 Matthew Christenson, *Acxiom Corporation, Little Rock, AR*
1:45 Gad Levanon, *The Conference Board, New York, NY*
2:00 Cheryl Hayes, *AIR Worldwide, Boston, MA*
2:15 *Discussion Leader:* David Crowe, *National Association of Home Builders, Washington, DC*

2:30pm *Break*

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

3:00pm G. Legal/Political and Social Equity Uses of ACS Data
Moderator/Organizer: Linda Gage, *California Department of Finance*
3:00 Kimball Brace, *Election Data Services, Manassas, VA*[1]
3:15 Joe Salvo and Peter Lobo, *New York City Department of City Planning*
3:30 Jeanne Gobalet, *Lapkoff and Gobalet Demographic Research, Inc., Saratoga, CA*
3:45 Andrew Beveridge, *Social Explorer, Inc., Bronxville, NY, and Queens College and Graduate Center, Flushing, NY*
4:00 *Discussion Leader:* Terri Ann Lowenthal, *Funders Census Initiative and The Census Project*

4:15pm H. Wrapping Up: The Future of ACS Data, and Floor Discussion of Workshop Themes
Moderators: Linda Gage, *California Department of Finance* and Ken Hodges, *Nielsen, Ithaca, NY*
Steve Murdock, *Department of Sociology and Hobby Center for the Study of Texas, Rice University*

5:00pm *Planned Adjournment*

[1]Mr. Brace was unable to attend the workshop; accordingly, additional time was allotted to the discussion portion of the session.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

CONTRIBUTORS TO ACCOMPANYING VOLUME OF CASE STUDIES AND USER PROFILES

PATRICIA C. BECKER, APB Associates, Inc.
CHRIS BENNER, Geography Graduate Group and Department of Human and Community Development, University of California, Davis
ANDREW G. BIGGS, American Enterprise Institute
THOMAS BRYAN, Bryan Geodemographics
RONALD CAMPBELL, *The Orange County Register*
ROSA CASTRO, Metropolitan Water District of Southern California
WON KIM COOK, Asian & Pacific Islander American Health Forum
A. RUPA DATTA, NORC at the University of Chicago
GENEVIEVE DUPUIS, Population Reference Bureau
NED ENGLISH, NORC at the University of Chicago
OWEN FURUSETH, University of North Carolina at Charlotte
NANCY GEMIGNANI, California State Census Data Center, Demographic Research Unit
ROBERT GEORGE, Chapin Hall Center for Children
MARK GOLDSTEIN, Maryland Department of Planning
RANDY GUSTAFSON, Tennessee State Data Center
JEFF HARDCASTLE, Nevada State Demographer
LESLEY HIRSCH, New York City Labor Market Information Service
GEORGE C. HOUGH, JR., Rice University
PATRICK JANKOWSKI, Greater Houston Partnership
BARBARA JOHN, University of North Carolina at Charlotte
DANIEL KASPRZYK, NORC at the University of Chicago
THOMAS LUDDEN, University of North Carolina at Charlotte
MARK MATHER, Population Reference Bureau
MEG MERRICK, Portland State University
PETER A. MORRISON, RAND (retired)
KORY NORTHROP, Graduate Student in Environmental Studies, University of Oregon
BECKY PETTIT, University of Washington
ROSELLA PICADO, Parsons Brinckerhoff
PHILLIP REESE, *The Sacramento Bee*
PAUL REIM, Boston Region Metropolitan Planning Organization
IDANIA R. RODRÍGUEZ AYUSO, Puerto Rico Institute of Statistics
GUY ROUSSEAU, Atlanta Regional Commission
DAVID A. SWANSON, University of California, Riverside
BRYAN SYKES, DePaul University
JANE TRAYNHAM, Maryland State Data Center
LESTER TSOSIE, The Navajo Nation
QINGFANG WANG, University of North Carolina at Charlotte

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

KIRK WOLTER, NORC at the University of Chicago
DAVID W. WONG, George Mason University
TING YAN, NORC at the University of Chicago
LAWRENCE YUN, National Association of Realtors

• • •

We note with sadness the untimely passing of Randy Gustafson, director of the Tennessee State Data Center, on June 9, 2012—shortly after submitting his entry to our case study set and shortly before the workshop itself. He was an early and enthusiastic supporter of this workshop—among the first to reply to the steering committee's canvass for interested ACS users and the first to agree to contribute a study.

Copyright © National Academy of Sciences. All rights reserved.

– B –

Biographical Sketches of Steering Committee Members and Presenters

MEMBERS OF THE STEERING COMMITTEE

Linda Gage (*Co-Chair*) retired in July 2012 as senior demographer in the State of California's Demographic Research Unit in the Department of Finance. In that capacity, her primary objective was to improve the currency, completeness, and accuracy of official state and federal demographic data that portray the people of California. She was actively involved in producing and evaluating intercensal population estimates for California and assessing data from the American Community Survey, and represented the state's demographic program and interests in federal and professional forums and evaluation of the effect of various demographic and statistical programs on the state. Previously chair of the steering committee of the Census Bureau's Federal-State Cooperative Program for Population Estimates, she remains active in professional associations. She serves on the Population Association of America (PAA) Public Affairs Committee and Committee on Population Statistics, and represents PAA on the Census Advisory Committee of Professional Associations. She served as the Governor's Liaison for Census 2000 and represented the State Data Center network and the Population Association of America on the U.S. Secretary of Commerce's Decennial Census Advisory Committee. At the National Research Council, she served on the Panel on the Design of the 2010 Census Program of Evaluations and Experiments. She has B.A. and M.A. degrees in

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

sociology, with emphasis in demography, from the University of California, Davis.

Ken Hodges (*Co-Chair*) is chief demographer at Nielsen. His responsibilities include methodology and evaluation for domestic small-area demographic estimates, the incorporation of U.S. census (including American Community Survey) data in the company's information products, and universe estimates standards for the company's global operations. He is active in the field of applied demography, having served on the 2010 Census Advisory Committee, the Population Association of America's Committee on Population Statistics, and the boards of the Association of Public Data Users and Council of Professional Associations on Federal Statistics. Prior to joining Claritas (now Nielsen) in 1993, he was director of demography at Donnelley Marketing Information Services. He has a Ph.D. in sociology/demography from Cornell University.

Linda Jacobsen is vice president of domestic programs at the Population Reference Bureau (PRB). She directs research projects, develops strategy and tools for disseminating population data and research, and writes PRB reports and web articles. She has extensive research experience with the American Community Survey and decennial census, and served for 6 years on the Census Advisory Committee of Professional Associations. Before joining PRB, she was a senior executive and chief demographer for two leading marketing information companies; research director for *American Demographics* magazine; and faculty member at Cornell University and the University of Iowa. She is chair of the Government and Public Affairs Committee of the Population Association of America, and a member of the board of directors of the Council of Professional Associations on Federal Statistics. For the Census Bureau, she coauthored two handbooks (targeted at general data users and high school teachers, respectively) in the Bureau's "Compass" series. She holds a doctorate in sociology from the University of Wisconsin–Madison.

Patrick Jankowski is vice president of research in the Support Services Division of the Greater Houston Partnership (formerly the Houston Chamber of Commerce), a metropolitan planning organization that works to enhance business activity, economic development, and quality of life in the Houston region. He oversees the Partnership's research department, which supports information gathering, data analysis, economic forecasting, and mapping for the Partnership's operating divisions. Prior to joining the Partnership in 1981, he was a business writer, editor, researcher, and economic analyst. He is a member of the Texas Economic and Demographic Association, the Council for Community and Economic Research, the National Association for Business Economics, and the Commercial Real Estate Research Forum. He has a bachelor's degree in economics from the University of Texas at Austin.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Joan Naymark retired in March 2012 as director of market analytics and planning at Target Corporation. Prior to joining Target Corporation in 1982, she held research positions at the Minnesota State Demographer's Office and the Upper Midwest Council. She has testified before Congress on business uses of census and American Community Survey data and represented the U.S. Chamber of Commerce on the Census Bureau's Census Advisory Committee. She is currently a member of the board of directors of the Association of Public Data Users and chair of the North American Research Task Force of the International Council of Shopping Centers. She coauthored (with Ken Hodges) articles on the use of census data in the private sector in the first and second editions of the *Encyclopedia of the U.S. Census*. She holds B.S. and M.A. degrees in sociology and demography from Western Washington University.

Richard Rathge is professor of sociology in the Departments of Sociology and Anthropology and Agribusiness and Applied Economics at North Dakota State University (NDSU). For 30 years, from 1981 through 2011, he served as director of the North Dakota State Data Center; in 2012, the State Data Center functions will transition from NDSU to a new North Dakota Census Office within the state commerce department. As director of the State Data Center and state demographer, he served as North Dakota's representative to the Federal-State Cooperative Programs for Population Estimates and Population Projections. He has also served as policy analyst for the North Dakota Kids Count program, a research center that monitors the health and well-being of children. He holds B.A. and M.A. degrees from New Mexico State University and a Ph.D. from Michigan State University.

WORKSHOP PRESENTERS AND DISCUSSION LEADERS

Andrew Beveridge is president and CEO of Social Explorer, Inc., in Bronxville, New York, and professor of sociology at Queens College and the Graduate School and University Center of the City University of New York. Since 1993, he has been a consultant on census data to *The New York Times*, and specializes in using geographic information systems techniques to analyze demographic data. He has previously taught in the sociology department at Columbia University. He holds a B.A. in economics and M.Phil. and Ph.D. degrees in sociology, all from Yale University.

Susan Brower became Minnesota's State Demographer in February 2012. Prior to that, she worked as a demographic researcher in nonprofit and academic settings. She has a master's degree in public policy from the University

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

of Minnesota and is completing a Ph.D. in sociology-demography at the University of Michigan.

Warren Brown is senior research associate at the Cornell Institute for Social and Economic Research at Cornell University. He returned to Cornell (where he previously directed the Program on Applied Demographics) after serving as public service associate and director of the applied demography program at the Carl Vinson Institute of Government at the University of Georgia. Accordingly, he has been responsible for producing population estimates and projections for both New York and Georgia. A member of the Population Association of America's Committee on Population Statistics and Committee on Applied Demography, he currently serves on the National Research Council's Panel to Review the 2010 Census. He received a B.A. in religious studies from the University of Virginia, an M.A. in sociology from the New School for Social Research, and a Ph.D. in development sociology from Cornell University.

Jacqueline Byers retired in fall 2012 as director of research and outreach at the National Association of Counties (NACo), a post she held since October 1996. Prior to joining NACo, she directed the research center at the Georgia Department of Community Affairs. She has been a member of the Decennial Census Advisory Committee since 1996 and is a member of the District of Columbia Bar. She has a B.A. from Bloomfield College, Bloomfield, New Jersey, and a J.D. from Catholic University, Columbus School of Law.

Kathleen Thiede Call is a professor in the University of Minnesota's School of Public Health, Division of Health Policy and Management. She is an investigator at the State Health Access Data Assistance Center (SHADAC), funded by the Robert Wood Johnson Foundation, which assists other states in monitoring rates of coverage and using data to inform policy and improve access. She created a survey of health insurance coverage for use by states and has developed a stream of research concerning the complexities of measuring and estimating rates of insurance coverage. Her research interests include disparities in access to health care and health insurance as well as developing community-driven solutions to health care barriers.

Ronald Campbell is a staff writer at *The Orange County Register* in Southern California. His multipart series on immigrants and the California economy—which he describes in a case study in this agenda book—was recognized in April 2011 with a "Best in Business" award from the Society of American Business Editors and Writers.

Matthew Christenson is a senior research analyst at Acxiom Corporation in Little Rock, Arkansas. As Acxiom's resident expert on census data, he is

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

responsible for building and supporting U.S. and international census data products and services. He also consults with clients and internal stakeholders on appropriate use of geospatial data. Prior to joining Acxiom in 2006, he was a statistician-demographer in the Population Division of the U.S. Census Bureau for more than 8 years. He holds his M.A./Ph.D. in sociology from the Ohio State University and a B.A. in communications from Concordia College, Moorhead, Minnesota.

Andrew Conrad is senior program manager for the Institute for Decision Making and the special projects manager for the Business and Community Services Division at The University of Northern Iowa. His responsibilities include the design, delivery, and follow-up of community and economic development services to communities and regions. A Certified Economic Developer, Drew currently serves on the Council for Community and Economic Research (C2ER) Board of Directors and the Iowa Workforce Development Board. He is a graduate of the Economic Development Institute (EDI) at the University of Oklahoma, and holds a B.A. in American history/pre-law and a master's in public policy.

David Crowe is chief economist and senior vice president at the National Association of Home Builders (NAHB). He is responsible for NAHB's forecast of housing and economic trends, survey research, and analysis of the home building industry and consumer preferences, as well as microeconomic analysis of government policies that affect housing.

Kenneth Darga has served as state demographer for Michigan since 2000. He is a member of the Steering Committee of the Federal-State Cooperative for Population Estimates and is the author of two books on the census.

Haya El Nasser has been with *USA TODAY* for 28 years, first as a business reporter and later as the newspaper's Los Angeles news correspondent. She began covering demographics and the census in 1996 and has reported on the lead-up to the 2000 census and the 2010 census and data findings from both. She has also covered the introduction of the American Community Survey and has mined the data for stories. She is responsible for wading through all Census data and identifying national trends, both on deadline and for enterprise stories. She tries to incorporate demographic data in everything she covers, including planning and development issues which have become part of her beat.

Ronald Fecso retired as chief statistician at the U.S. Government Accountability Office (GAO) shortly after the workshop, having held that position since 2006. Prior to GAO, he served as chief statistician in the Division of Science Resources Statistics at the National Science Foundation (now the National

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Center for Science and Engineering Statistics) and the National Agricultural Statistics Service (NASS) of the U.S. Department of Agriculture. He has a B.A. in mathematics from Rider College and an M.A. in mathematical statistics from the University of Rochester.

Ford Fessenden is a graphics editor at *The New York Times*. He has been a reporter on the metropolitan desk and the investigations desk, specializing in exploiting databases and computer tools for news stories, including a study of rampage killers, and the re-creation of the last 102 minutes of the World Trade Center. He ran the Florida ballot project for a consortium of the *New York Times*, *Washington Post*, and *Wall Street Journal* in 2001, cataloguing the ballots that went uncounted after the Supreme Court stopped the recount. He has worked for *Newsday* and other newspapers.

Katie Genadek is a research assistant on the Integrated Public Use Microdata Series (IPUMS) user support core at the Minnesota Population Center, University of Minnesota. She has been an IPUMS user since 2003 and has worked on the IPUMS projects since 2007. She currently directs IPUMS outreach and dissemination efforts and she gives IPUMS trainings at conferences, universities, and policy institutes. She is finishing a Ph.D. in applied economics at the University of Minnesota; she holds an M.S. in applied economics and a B.S. in economics from Montana State University.

Linda Giannarelli is a senior fellow in the Income and Benefits Policy Center of the Urban Institute. She serves as project director for the maintenance and development of the TRIM3 microsimulation model, and conducts research in the areas of poverty, child care, and low-income benefit programs. In recent years, she has co-led efforts applying the TRIM3 model to ACS data to estimate the impact of state policy choices on poverty.

Jeanne Gobalet has been an applied demographer for many years. Since 1992, she has been a principal of Lapkoff & Gobalet Demographic Research, Inc., located in California. Much of her recent work utilizes geographic information systems techniques and ACS data. She has served as an expert witness in cases involving political redistricting, racially polarized voting, environmental justice, jury selection, and discrimination in housing, employment, and mortgage lending. Her consulting practice includes helping clients meet federal and California Voting Rights Act requirements. She earned four degrees from Stanford University, including a Ph.D. in sociology, and was a postdoctoral fellow in demography at the University of California, Berkeley.

Cheryl Hayes is the senior manager of the Exposures team in the Research and Modeling Group at AIR Worldwide. She is responsible for leading the

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

team in the development of comprehensive, high-resolution industry exposure databases for more than 90 countries worldwide.

Beth Jarosz is senior demographer at the San Diego Association of Governments (SANDAG) and has more than a decade of experience in demographic estimation, forecasting, and analysis. At SANDAG, she manages the annual population estimates and long-range forecast programs, and she has worked on a variety of topics ranging from transportation planning to public health. She also teaches sociology at Pensacola State College. She holds a B.S. in applied economics from the University of Rhode Island and an M.A. in demographic and social analysis from the University of California, Irvine.

Gad Levanon is director of macroeconomic research at The Conference Board, where he also leads the labor markets program. He created The Conference Board Employment Trends Index™, a widely used measure that fills the need for a leading index of employment. His research focuses on compensation growth, retirement trends, labor shortages, labor productivity, and future trends in employment and unemployment. Levanon received his Ph.D. in economics from Princeton University, and he holds undergraduate and master's degrees from Tel Aviv University in Israel.

Terri Ann Lowenthal, a former staff director of the House Subcommittee on Census and Population, advises nonprofits, foundations, and the business sector on issues related to the census and federal statistics generally. She is a co-director of The Census Project, a non-partisan coalition of census and ACS stakeholders; consultant to the Funders Census Initiative, a collaborative of foundations interested in an accurate and comprehensive census and ACS, and writes The Census Project Blog.

Steve H. Murdock is the Allyn R. and Gladys M. Cline Professor of Sociology and director of the Hobby Center for the Study of Texas at Rice University. He served as director of the U.S. Census Bureau from December 2007 until the change of administrations in January 2009. The first official state demographer of Texas, he previously held faculty appointments at the University of Texas at San Antonio and Texas A&M University.

Kathleen Miller joined the Rural Policy Research Institute (RUPRI) at the University of Missouri in January 2000 as program director. Miller is responsible for the coordination of the RUPRI program of work, which encompasses researchers and practitioners across the country analyzing the rural impacts of public policies and programs in health care, human services, entrepreneurship, and regional and community development. Her research interests include understanding how the definitions of rural geographies impact policy outcomes,

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

and she has published several RUPRI policy briefs on this topic. She received an M.S. in agricultural economics from Penn State University, where she also received her B.S. in agricultural economics and rural sociology.

Russ Paulsen is executive director for community preparedness and resilience at the national headquarters of the American Red Cross in Washington, in which capacity he leads Red Cross efforts to build public preparedness for disasters at the neighborhood level. Prior to his current position, he directed the Red Cross hurricane recovery unit and led the organization's long-term recovery efforts after Hurricane Katrina.

Allison Plyer is deputy director of the Greater New Orleans Community Data Center. She is co-author of The New Orleans Index series, developed in collaboration with the Brookings Institution to analyze the state of the New Orleans recovery. She is recognized as an international expert in post-Katrina demographics and New Orleans recovery trends.

Steven Romalewski directs the CUNY Mapping Service at the Graduate Center of the City University of New York. For the past 20 years he has leveraged geographic information system (GIS) tools and concepts to help organizations in diverse fields understand and act on the spatial dimensions of their work. Since 2006 he has also taught GIS to graduate planning students at Pratt Institute.

Joseph Salvo is director of the Population Division at the New York City Department of City Planning. He is the former president of the Association of Public Data Users and a member of the Population Association of America's Committee on Population Statistics, and is active in research on the evaluation and use of American Community Survey data and on the demography of immigration in the New York area. He has served on several National Research Council panels related to the census and the ACS, including a recently completed Panel on Statistical Methods for Measuring the Group Quarters Population in the ACS. A fellow of the American Statistical Association, he holds M.A. and Ph.D. degrees in sociology from Fordham University.

Vincent L. Sanders is a lead transportation systems planner for the Metropolitan Transit Authority of Harris County, Texas (METRO). He joined METRO's Service Design & Development/Travel Demand Forecasting & Analysis Division in 1998. He is primarily responsible for managing patronage forecasts, and also manages the socioeconomic and travel databases that generate reports on trip patterns, demographic profiles, and ridership estimates for proposed transit service and regional options and facilities. He has worked with other agencies to help coordinate joint development and commerce studies relative to

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

employment concentrations and worker travel patterns, and serves on several interagency committees. Prior to joining METRO, he worked as a transportation consultant with RSM Services. Sanders earned a B.S. in business administration from Wayne State University and an M.S. in transportation planning and management from Texas Southern University.

James Stark is an epidemiologist for the New York City Department of Health and Mental Hygiene. He works in the Bureau of Epidemiology Services and provides methodological consultation to programs throughout the agency. In addition, he collaborates with the Built Environment and Health group at Columbia University's Mailman School of Public Health.

Barry Steinhardt is founder of Friends of Privacy USA and a senior advisor and trustee of Privacy International. He founded Friends of Privacy USA after retiring in 2009 following nearly 30 years with the American Civil Liberties Union (ACLU), including service as associate director of the ACLU and director of ACLU's Program on Technology and Liberty. He has also served as president of the Electronic Frontier Foundation. During his time with the ACLU, he was a member of the Decennial Census Advisory Committee. He is currently a member of the Data Privacy and Integrity Committee of the U.S. Department of Homeland Security.

Amy Terpstra is the associate director of the Social IMPACT Research Center at Heartland Alliance. Her work focuses on researching and disseminating knowledge on poverty and income-related issues. She holds a master's degree in social work from the Jane Addams College of Social Work at the University of Illinois at Chicago.

Richard K. Thomas is a Memphis-based health care consultant with more than 40 years experience in health services research, planning, and development. He provides consultation to hospitals, health systems, clinics and other health care organizations. He is on the faculties of the University of Mississippi and the University of Tennessee Health Science Center and has authored numerous publications on health care and demographics.

Stephen Tordella, president of Decision Demographics, Inc., in Arlington, Virginia, is an applied demographer with more than 35 years of public- and private-sector consulting experience. He was a state-level demographer for 10 years, and then worked in the commercial demographics industry for 3 years. Since 1987 he has run a small consulting business. He holds an M.A. in demography from Brown University.

Copyright © National Academy of Sciences. All rights reserved.

Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop

Lester Tsosie has been working for the Navajo Nation Division of Economic Development as a demographer. He has used U.S. Census and ACS data and information to create data products for Navajo Nation leaders, local elected leaders, practitioners, and residents of local Navajo chapters.

Copyright © National Academy of Sciences. All rights reserved.

COMMITTEE ON NATIONAL STATISTICS

The Committee on National Statistics was established in 1972 at the National Academies to improve the statistical methods and information on which public policy decisions are based. The committee carries out studies, workshops, and other activities to foster better measures and fuller understanding of the economy, the environment, public health, crime, education, immigration, poverty, welfare, and other public policy issues. It also evaluates ongoing statistical programs and tracks the statistical policy and coordinating activities of the federal government, serving a unique role at the intersection of statistics and public policy. The committee's work is supported by a consortium of federal agencies through a National Science Foundation grant.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.