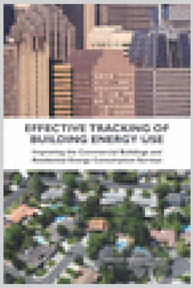


Effective Tracking of Building Energy Use: Improving the Commercial Buildings and Residential Energy Consumption Surveys

ISBN
978-0-309-25401-4

140 pages
6 x 9
PAPERBACK (2012)

William F. Eddy and Krisztina Marton, Editors; Panel on Redesigning the Commercial Buildings and Residential Energy Consumption Surveys of the Energy Information Administration; National Research Council

EFFECTIVE TRACKING OF BUILDING ENERGY USE

Improving the Commercial Buildings and Residential Energy Consumption Surveys

Panel on Redesigning the Commercial Buildings and
Residential Energy Consumption Surveys of the
Energy Information Administration

William F. Eddy and Krisztina Marton, *Editors*

Committee on National Statistics
Division of Behavioral and Social Sciences and Education

Board on Energy and Environmental Systems
Division on Engineering and Physical Sciences

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

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

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

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

This study was supported by contract number DE-DT 0000670 T.O. #30 between the National Academy of Sciences and the U.S. Department of Energy. Support for the work of the Committee on National Statistics is provided by a consortium of federal agencies through a grant from the National Science Foundation (award number SES-0453930). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the organizations or agencies that provided support for this project.

International Standard Book Number-13: 978-0-309-25401-4
International Standard Book Number-10: 0-309-25401-9

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Printed in the United States of America

Suggested citation: National Research Council. (2012). *Effective Tracking of Building Energy Use: Improving the Commercial Buildings and Residential Energy Consumption Surveys.* Panel on Redesigning the Commercial Buildings and Residential Energy Consumption Surveys of the Energy Information Administration. W.F. Eddy and K. Marton, Eds. Committee on National Statistics, Division of Behavioral and Social Sciences and Education, and Board on Energy and Environmental Systems, Division on Engineering and Physical Sciences. Washington, DC: The National Academies Press.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

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

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

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

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

www.national-academies.org

PANEL ON REDESIGNING THE COMMERCIAL BUILDINGS AND RESIDENTIAL ENERGY CONSUMPTION SURVEYS OF THE ENERGY INFORMATION ADMINISTRATION

WILLIAM F. EDDY (*Chair*), Department of Statistics, Carnegie Mellon University
MARILYN A. BROWN, School of Public Policy, Georgia Institute of Technology
FREDERICK CONRAD, Institute for Social Research, University of Michigan
DONALD A. DILLMAN, Social and Economic Sciences Research Center, Washington State University
DWIGHT K. FRENCH, Energy Consumption Division, Energy Information Administration (retired), Silver Spring, Maryland
JACK G. GAMBINO, Household Survey Methods Division, Statistics Canada, Ottawa
CLARK W. GELLINGS, Electric Power Research Institute, Palo Alto, California
JANE F. GENTLEMAN, National Center for Health Statistics, Hyattsville, Maryland
DAVID G. HUNGERFORD, California Energy Commission, Sacramento
PHILLIP S. KOTT,* RTI International, Rockville, Maryland
NINA S.-N. LAM, Department of Environmental Sciences, Louisiana State University
ALAN K. MEIER, Lawrence Berkeley National Laboratory, Berkeley, California
MICHAEL M. MEYER, Google, Seattle, Washington
MAXINE L. SAVITZ,** Honeywell, Inc. (retired), Los Angeles, California

KRISZTINA MARTON, *Study Director*
MICHAEL L. COHEN, *Senior Program Officer*
NANCY J. KIRKENDALL, *Senior Program Officer*
AGNES GASKIN, *Administrative Assistant*

*Resigned from the panel on August 27, 2010.
**Joined the panel on July 1, 2010.

COMMITTEE ON NATIONAL STATISTICS
2010–2012

LAWRENCE D. BROWN (*Chair*), Department of Statistics, The Wharton School, University of Pennsylvania
JOHN M. ABOWD, School of Industrial and Labor Relations, Cornell University
ALICIA CARRIQUIRY, Department of Statistics, Iowa State University
WILLIAM DuMOUCHEL, Oracle Health Sciences, Tucson, AZ
V. JOSEPH HOTZ, Department of Economics, Duke University
MICHAEL HOUT, Survey Research Center, University of California, Berkeley
KAREN KAFADAR, Department of Statistics, Indiana University
SALLIE KELLER, Science and Technology Policy Institute, Institute for Defense Analyses, Washington, DC
LISA LYNCH, The Heller School for Social Policy and Management, Brandeis University
SALLY C. MORTON, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh
JOSEPH P. NEWHOUSE, Division of Health Policy Research and Education, Harvard University
RUTH PETERSON, Criminal Justice Research Center, Ohio State University
HAL S. STERN, Donald Bren School of Computer and Information Sciences, University of California, Irvine
JOHN H. THOMPSON, National Opinion Research Center, University of Chicago
ROGER TOURANGEAU, Westat, Rockville, MD
ALAN M. ZASLAVSKY, Department of Health Care Policy, Harvard University Medical School

CONSTANCE F. CITRO, *Director*

BOARD ON ENERGY AND ENVIRONMENTAL SYSTEMS
2010–2011

ANDREW BROWN, JR. (*Chair*), Delphi Corporation, Troy, Michigan
RAKESH AGRAWAL, Department of Chemical Engineering, Purdue University
WILLIAM BANHOLZER, Dow Chemical Company, Midland, Michigan
MARILYN A. BROWN, School of Public Policy, Georgia Institute of Technology
MICHAEL CORRADINI, Department of Engineering Physics, University of Wisconsin–Madison
PAUL A. DeCOTIS, Long Island Power Authority, Albany, New York
CHRISTINE EHLIG-ECONOMIDES, Department of Petroleum Engineering, Texas A&M University
SHERRI GOODMAN, CNA, Alexandria, Virginia
NARAIN HINGORANI, Consultant, Los Altos Hills, California
ROBERT J. HUGGETT, Department of Environmental Sciences, College of William and Mary
DEBBIE A. NIEMEIER, Department of Civil and Environmental Engineering, University of California, Davis
DANIEL NOCERA, Department of Chemistry, Massachusetts Institute of Technology
MICHAEL OPPENHEIMER, Departments of Geosciences and International Affairs, Princeton University
DAN REICHER, Steyer-Taylor Center for Energy Policy and Finance, Stanford University
BERNARD ROBERTSON, DaimlerChrysler Corporation (retired), Bloomfield, Michigan
ALISON SILVERSTEIN, Consultant, Pflugerville, Texas
MARK H. THIEMENS, Division of Physical Sciences, University of California, San Diego
RICHARD WHITE, Oppenheimer & Company, New York, New York

JAMES J. ZUCCHETTO, *Director*

Acknowledgments

We wish to thank everyone who contributed their time and expertise to assist the work of the Panel on Redesigning the Commercial Buildings and Residential Energy Consumption Surveys of the Energy Information Administration. These generous contributions made this report possible.

We would like to thank the sponsor of the study, the U.S. Energy Information Administration (EIA)—under the leadership of then administrator Richard Newell and acting administrator Howard Gruenspecht—for the support provided as the panel reviewed background materials and information about the challenges faced by the Commercial Buildings Energy Consumption Survey (CBECS) and the Residential Energy Consumption Survey (RECS). Communication between the panel and EIA was greatly facilitated by technical liaisons Stephanie Battles, Stephanie Brown, and later Thomas Leckey. A number of EIA staff made very informative presentations and provided useful materials that helped the panel become thoroughly familiar with the surveys. We would especially like to thank James Berry, Eugene Burns, Joelle Michaels, and Eileen O'Brien for their assistance.

Lynda Carlson of the National Science Foundation (retired) and former head of the Energy End Use Division within EIA provided valuable background information about the evolution of the CBECS and RECS during the first two decades of the surveys' existence. EIA's data collection contractors were very helpful in getting the panel up to speed on details related to recent sample designs and data collection operations. We would like to thank Rachel Harter and Susan Hinkins from the National Opinion Research Center and David Judkins from Westat for their presentations to the panel.

Staff from other statistical agencies also took time from their busy schedules to share their experiences with similar data collections and discuss with the panel the challenges faced by the CBECS and RECS. We thank Mark Miller from the National Agricultural Statistics Service and David Vandenbroucke from the U.S. Department of Housing and Urban Development for their input. The panel also greatly benefited from learning about the experiences of researchers who collect energy consumption data in other countries. We are indebted to Sylvie Gauthier from Statistics Canada, Shane Holt from the Australian Department of the Environment, and Shigeki Kametani from Tokyo University of Marine Science and Technology.

A fundamental aspect of the panel's work was to obtain a thorough understanding of data user needs. We thank all CBECS and RECS data users who directly or indirectly provided input to the panel. This report greatly benefited from their candid contributions. We also greatly appreciated the in-person presentations delivered by Jennifer Amann from the American Council for an Energy-Efficient Economy, Erin Boedecker from the U.S. Energy Information Administration, Miriam Goldberg from KEMA, Piljae Im from the Oak Ridge National Laboratory, Elicia John from Energy Star, Richard Levy from the National Multi Housing Council, Leon Litow from the U.S. Department of Health and Human Services, Joe Loper from ITRON, Paul Mathew from Lawrence Berkeley National Laboratory, Harvey Sachs from the American Council for an Energy-Efficient Economy, Marla Sanchez from the Lawrence Berkeley National Laboratory, Anant Sudarshan from Stanford University, and Michael Zatz from Energy Star.

We would also like to thank the National Research Council (NRC) staff who contributed to this study. In particular, the panel could not have done its work in such a thorough and thoughtful manner without the skills of Krisztina Marton, who served as the panel's study director. We are also grateful for the guidance and support received from Constance Citro, director of the Committee on National Statistics (CNSTAT), for the input provided by CNSTAT senior program officers Michael Cohen and Nancy Kirkendall, and for the contributions to the study of James Zucchetto, director of the Board on Energy and Environmental Systems. Robert Pool provided editorial help with the report, and Kirsten Sampson Snyder shepherded the report through the review process. Administrative assistance was provided by Agnes Gaskin.

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 NRC. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report: Paul A. DeCotis, Power Markets, Long Island Power Authority; Miriam L. Goldberg, Sustainable Use, KEMA Inc.; Andy Kohut, Manufacturing and Energy Division, Statistics Canada; Nancy L. Leach, consultant; H. Scott Matthews, Green Design Institute, Carnegie Mellon University; Stephen M. Miller, Bureau of Labor Statistics, U.S. Department of Labor; Steven Nadal, American Council for an Energy-Efficient Economy; and Sarah Nusser, Department of Statistics, Iowa State University.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report before its release. The review of this report was overseen by Elisabeth Drake, Energy Laboratory, Massachusetts Institute of Technology (retired), and Joel Horowitz, Department of Economics, Northwestern University. Appointed by the NRC, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring panel and the institution.

Finally, we recognize the many federal agencies that support CNSTAT directly and through a grant from the National Science Foundation. Without their support and their commitment to improving the national statistical system, the committee work that is the basis of this report would not have been possible.

William F. Eddy, *Chair*
Panel on Redesigning the Commercial Buildings and
Residential Energy Consumption Surveys of the
Energy Information Administration

Contents

Appendixes

Summary

At the request of the U.S. Energy Information Administration (EIA), a panel of the National Research Council conducted a study to review the designs of two of EIA's energy consumption surveys: the Commercial Buildings Energy Consumption Survey (CBECS) and the Residential Energy Consumption Survey (RECS). The panel's charge was to recommend updates to these surveys based on current and expected future data needs. The CBECS and RECS are used by Congress, the executive branch, state governments, local governments, energy producers, energy providers, investors, researchers, the media, and the public. As the energy demands associated with commercial and residential buildings grow and change, the need for reliable, valid, and increasingly more sophisticated data is likely to grow, making state-of-the-art energy consumption surveys even more necessary.

The designs of the CBECS and RECS are fundamentally similar, but there are certain differences that result from the differing characteristics of their two populations of interest. The CBECS is a survey of commercial buildings, and, as such, it faces some unique challenges. The RECS is a household survey with a more conventional—albeit also challenging—data collection approach. Both surveys collect data about the complex, highly technical topic of energy consumption and its determinants, and both are the surveys of record on this topic for researchers and policy makers in the United States. The CBECS, in particular, is the only national data source of detailed characteristics and energy use of commercial buildings (Michaels, 2010).

Based on the panel's research—and, in particular, on discussions with data users—two topics emerged as priority areas for both surveys: (1) the timeliness and frequency of the data and (2) data gaps. The panel concluded that the frequency of the surveys does not meet data user needs well and that the length of time it takes to process and prepare the data for release is a particular concern for many data users. The panel also identified a variety of data gaps, some of which are related to EIA's inability to release all of the data collected because the sample sizes are too small to produce data that meet quality and confidentiality thresholds. In some cases, data users would like more detailed information about building characteristics and energy consumption than what the surveys currently collect or would like new measures added to the questionnaires to capture new topics of interest, such as new end uses.

The panel's main recommendations, which are briefly discussed below, focus on addressing the two priority areas mentioned, in the context of additional important considerations for the surveys, such as data quality and costs. The main recommendations, some of which involve major changes to the survey designs, are described first, followed by additional recommendations that address more specific, technical aspects of the data collections.

COMMERCIAL BUILDINGS ENERGY CONSUMPTION SURVEY

Timeliness and Frequency of the CBECS Data

The panel's research, which included discussions with data users, found that one of the biggest concerns related to the surveys is the amount of time it takes to collect, process, and release the data. One option for addressing data users' need for quicker access to data is to implement a rotating sample design by, for example, dividing the full sample into four subsamples and collecting a quarter of the data each year over a four-year period instead of collecting data from the entire sample once every four years. This change would enable annual data releases for a subset of the sample.

> **Recommendation CBECS-1:** EIA should evaluate the usefulness of implementing a rotating sample design for the CBECS to improve the timeliness of the data.

Another way to shorten the time required for the survey would be to move some of the CBECS data collection from face-to-face interviews to the web. For at least a subset of the sample, this will reduce the amount of time needed to complete the data collection. When transitioning the survey to the web, it will also be important to eliminate and simplify some of the editing procedures in order to reduce the associated delays.

Recommendation CBECS-3: Informed by a methodological research program, EIA should begin developing procedures for a multimode approach and should begin moving some of the CBECS data collection to the web.

Recommendation CBECS-4: EIA should investigate strategies for releasing the CBECS data faster, for example, by revising the editing procedures or by completing the editing for a subset of the variables and releasing these data prior to completing the editing for the full data set.

CBECS Data Gaps

The current CBECS sample design is best suited for producing descriptive statistics for larger geographical divisions, such as the entire country or census division levels. The relatively small sample sizes, in combination with strict quality control and confidentiality protections, severely limit the amount of data that can be released from the survey. This in turn limits, in terms of both geography and complexity, the analyses that can be conducted based on the data.

Recommendation CBECS-5: As part of its efforts to address the needs of data users, EIA should make it a priority to identify opportunities for increasing the sample size in order to enable the release of more of the CBECS data that are currently being collected.

Recommendation CBECS-6: EIA should consider establishing a research data center (RDC) or evaluate the option of using an existing RDC maintained by another organization to provide data users with secure access to CBECS data that are currently not publicly released.

Other CBECS Recommendations

In addition to the core recommendations highlighted above, the report includes additional recommendations that are more technical and specific in nature and that are focused on supporting the large-scale revisions and on further improving and updating the surveys.

While the CBECS cannot meet all data user needs, the panel makes several recommendations aimed at ensuring that the questionnaire remains responsive to changes in the energy landscape.

Recommendation CBECS-2: EIA should consider integrating a longitudinal element into the CBECS sample design to obtain better estimates of change.

Recommendation CBECS-7: EIA should prepare for the more widespread availability of smart meter data in the future by evaluating potential uses of such data, strategies for collecting them, and ways of addressing new confidentiality challenges.

Recommendation CBECS-8: EIA should develop a process for the periodic review of new energy end uses or end uses that are becoming more widespread in the commercial sector and which may need to be included on the CBECS questionnaire. The process should also identify end uses that are becoming obsolete and that can be removed from the questionnaire.

Recommendation CBECS-9: To accommodate data user needs for more detailed information, EIA should evaluate the possibility of administering a short-form and a long-form CBECS questionnaire.

Recommendation CBECS-15: EIA should invest in periodic reviews of the CBECS questionnaire content and wording. This should be understood as a routine updating of the instruments, separate from the concept of a major redesign of the survey.

Given the challenges associated with developing and maintaining a sampling frame based on buildings, EIA should revisit the possibility of using establishment-based lists as the primary source of the CBECS sampling frame.

Recommendation CBECS-10: EIA should conduct research to evaluate the advantages and costs associated with using an establishment-based list for the CBECS sampling frame.

The panel also recommends several changes to the CBECS data collection procedures with the goals of increasing efficiency and overcoming some of the challenges associated with collecting highly technical information from respondents who may find some of the questions challenging. This includes researching alternative ways of collecting some of the data that may be particularly difficult to collect from building respondents or that could be more efficiently collected from other sources. For example, EIA could explore collecting some of the data from the headquarters of major centralized accounts (such as Walmart or McDonald's) and relying more on energy suppliers. Alternatives sources of data, such as existing administrative records databases (for example, property tax databases), should also be evaluated.

Recommendation CBECS-11: To increase CBECS data collection efficiency, EIA should explore the possibility of contacting the headquarters of major centralized accounts and collecting data about all of the buildings in the sample from these centralized sources.

Recommendation CBECS-12: EIA should evaluate whether working more closely with the energy suppliers of commercial buildings could lead to efficiencies in the data collection process.

Recommendation CBECS-13: EIA should conduct an ongoing evaluation of administrative records as potential sources for substantive data for the CBECS.

To address challenges associated with collecting highly technical information from respondents who are not necessarily experts on topics such as building characteristic and energy consumption, EIA should consider evaluating how involving professional energy auditors in the data collection process would affect costs and data quality. Debriefings with field interviewers could also be useful in identifying questionnaire items that respondents find particularly difficult to answer and in developing possible improvements to these items.

Recommendation CBECS-14: EIA should test the use of professional energy auditors on a small scale to better understand the costs and benefits related to having experts collect data for a subset of the CBECS sample.

Recommendation CBECS-16: Interviewer debriefings should become an integral part of the CBECS data collection process in order to identify problems with the questionnaires and procedures and to serve as a source of ideas for increased efficiencies.

RESIDENTIAL ENERGY CONSUMPTION SURVEY

Timeliness and Frequency of the RECS Data

As with the CBECS, one of the biggest concerns related to the RECS is the amount of time it takes to collect, process, and release the data.

Recommendation RECS-1: EIA should evaluate the usefulness of implementing a rotating sample design for the RECS to improve the timeliness of the data.

Another way of speeding up the survey operations would be to move some of the RECS data collection from face-to-face interviews to web interviews and to eliminate and simplify some of the editing procedures to reduce the associated delays.

Recommendation RECS-3: Informed by a methodological research program, EIA should begin developing procedures for a multimode approach and should begin moving some of the RECS data collection to the web.

Recommendation RECS-4: EIA should investigate strategies for releasing the RECS data faster, for example, by revising the editing procedures or by completing the editing for a subset of the variables and releasing these data prior to completing the editing for the full data set.

RECS Data Gaps

The current RECS sample design is best suited for producing descriptive statistics for larger geographic areas, such as the entire country and census divisions. State-level data are available for only 16 states. The small sample sizes, in combination with strict quality control and confidentiality protections, severely limit the amount of data that can be released from the survey. This in turn limits the complexity and geographic resolution of the analyses that can be conducted based on the data.

Recommendation RECS-5: As part of its efforts to address the needs of data users, EIA should make it a priority to identify opportunities for increasing the sample size in order to enable the release of more of the RECS data that are currently being collected.

Recommendation RECS-6: EIA should consider establishing a research data center or evaluate the option of using an existing RDC maintained by another organization to provide data users with secure access to RECS data that are currently not publicly released.

Other RECS Recommendations

In addition to the core recommendations described above, the report contains additional recommendations that are more technical and specific in nature and that are focused on supporting the large-scale revisions discussed as well as on further improving and updating the survey.

While it is important to be mindful of the fact that the RECS cannot meet all data user needs, the panel makes several recommendations to assure that the questionnaire remains responsive to changes in the energy landscape.

Recommendation RECS-2: EIA should consider integrating a longitudinal element into the RECS sample design to obtain better estimates of change.

Recommendation RECS-7: EIA should prepare for the more widespread availability of smart meter data in the future by evaluating potential uses of such data, strategies for collecting them, and ways of addressing new confidentiality challenges.

Recommendation RECS-8: In anticipation of the spread of electric vehicles, EIA should prepare for RECS to collect more information about these, especially more detail about the capacity to charge electric vehicles.

Recommendation RECS-9: EIA should develop a process for the periodic review of new energy end uses or end uses that are becoming more widespread in the residential sector and which may need to be included on the RECS questionnaire. The process should also identify end uses that are becoming obsolete and that can be removed from the questionnaire.

Recommendation RECS-10: EIA should consider implementing a "whole building" supplement to the RECS to address the data gap related to multiunit residential buildings.

Recommendation RECS-11: To accommodate data user needs for more detailed information, EIA should evaluate the possibility of administering a short-form and a long-form RECS questionnaire.

Recommendation RECS-16: EIA should invest in periodic reviews of the RECS questionnaire content and wording. This should be understood as a routine updating of the instruments, separate from the concept of a major redesign of the survey.

Given the costs associated with area probability samples, combined with the recent advances made in the field of address-based sampling, the panel recommends evaluating address-based sampling for the RECS.

Recommendation RECS-12: Research should be conducted to evaluate the possibility of replacing the RECS area probability sample with address-based sampling, using an address list developed by commercial vendors based on the U.S. Postal Service delivery sequence file.

The panel also recommends several changes to the RECS data collection procedures that are aimed at increasing efficiency and overcoming some of the challenges associated with collecting highly technical information from respondents who may find some of the questions challenging.

Recommendation RECS-13: EIA should conduct an ongoing evaluation of administrative records as potential sources for substantive data for the RECS, especially for square footage data.

Recommendation RECS-14: EIA should evaluate whether working more closely with the energy suppliers could lead to efficiencies in the data collection process.

Recommendation RECS-15: EIA should test the use of professional energy auditors on a small scale to better understand the costs and benefits related to having experts collect data for a subset of the RECS sample.

Recommendation RECS-17: Interviewer debriefings should become an integral part of the RECS data collection process in order to identify problems with the questionnaires and procedures and to serve as a source of ideas for increased efficiencies.

A STATE-OF-THE-ART ENERGY CONSUMPTION DATA COLLECTION PROGRAM

The panel's recommendations in this report are based on a careful balance between what the data needs are and what is most realistic to implement in the short term. The last chapter of the report outlines some additional ideas that could be considered as part of envisioning a state-of-the-art data collection program for the energy surveys of the future.

A comprehensive approach to collecting building energy consumption information could be based on a sampling frame that includes all buildings, regardless of type. Advances in computer technology and satellite mapping are greatly increasing the feasibility of implementing such a design, and are changing how much of the work can be performed by a computer program prior to, or instead of, field work. Other advances in technology are increasing opportunities for enhancing features of the data collection. For example, it is becoming cheaper and easier to integrate digital images that can help reduce ambiguities surrounding building characteristics or equipment. Increasingly more sophisticated interactive online tools are becoming available and can be used to engage sample members and encourage participation in the surveys.

In terms of possibilities for growth and meeting data needs, EIA could explore opportunities to partner with other government statistical agencies, organizations, or energy suppliers to collect data on specialized topics, without increasing the burden on the current CBECS or RECS. EIA could consider serving as the coordinator or data center for research on a number of special topics or special populations or performing additional functions that involve more active dissemination and closer collaboration with the research community.

1

Introduction

The United States is responsible for nearly one-fifth of the world's energy consumption (U.S. Energy Information Administration, 2011a), or approximately 100 quadrillion Btu of energy, which is the equivalent of over $1 trillion in energy expenditures every year (U.S. Energy Information Administration, 2011b). Population growth, and the associated growth in housing, commercial floor space, transportation, goods, and services, is expected to cause a 0.7 percent annual increase in energy demand for the foreseeable future (U.S. Energy Information Administration, 2011c). The scale of these numbers, combined with increasing concerns about U.S. economic competitiveness, energy security, national security, and climate change, underscores the importance of making sound energy policy and investment decisions for the future. In the words of U.S. Secretary of Energy Steven Chu, "Any path close to 'business as usual' will imperil future generations with dangerous and unacceptable economic, social, and environmental risks" (U.S. Department of Energy, 2011a).

The primary federal government authority charged with collecting, analyzing, and disseminating energy data to inform these decisions is the U.S. Department of Energy's (DOE's) statistical agency, the U.S. Energy Information Administration (EIA). EIA was established by the Department of Energy Organization Act of 1977, and its activities are funded through an annual appropriation from Congress. Although the data produced by EIA play an extremely important role in informing energy policy decisions, over the years the agency has struggled with budget limitations that have made meeting its mandate difficult—and sometimes impossible. Figure 1-1 shows EIA's funding levels between 1978 and 2011.

EIA's portfolio of data collections includes three surveys of energy-consuming end use sectors: the Commercial Buildings Energy Consumption Survey (CBECS), the Residential Energy Consumption Survey (RECS), and the Manufacturing Energy Consumption Survey. Prior to 1994, EIA also conducted a transportation energy use survey, the Residential Transportation Energy Consumption Survey, but budget cuts forced this data collection to be discontinued after 1994.

EIA asked the National Research Council's (NRC's) Committee on National Statistics, with input from its Board on Energy and Environmental Systems, to convene a panel to study two of these surveys, the CBECS and RECS, which cover the commercial and residential end use sectors. The panel's charge was to evaluate the designs of the two surveys and recommend updates based on current and expected future data needs. (See Box 1-1 for the exact wording of the panel's charge.)

The CBECS and RECS were established in response to the energy shocks of the 1970s, which highlighted the need for coordinating the nation's energy functions and for a comprehensive federal energy plan. At the time policy makers had accurate data regarding the sources of our energy (coal, petroleum, etc.) but had only crude guesses about how that energy was used to carry out basic functions such as space heating, lighting, and refrigeration. While the 1970s energy crisis is a distant memory to many, threats to energy security persist and are likely to persist for the foreseeable future. Mitigating these threats requires reliable information on how energy is used.

Today the energy used by the commercial and residential sectors represents approximately 40 percent of the nation's total energy consumption, and the share of these two sectors is expected to increase in the future, in large part because of the 5.8 quadrillion Btu growth expected over the next 25 years in the commercial sector, shown in Figure 1-2 (U.S. Energy Information Administration, 2011c). On the other hand, various federal and local policies have been put into place with the goal of reversing this trend and reducing overall energy use to a more sustainable level. The importance of a thorough understanding of how commercial and residential buildings use energy has therefore never been greater. Only with good quality data can policy makers, industry, and consumers respond effectively to a variety of challenges, ranging from the mundane to potentially catastrophic scenarios.

The energy consumption surveys of EIA are the most relevant sources

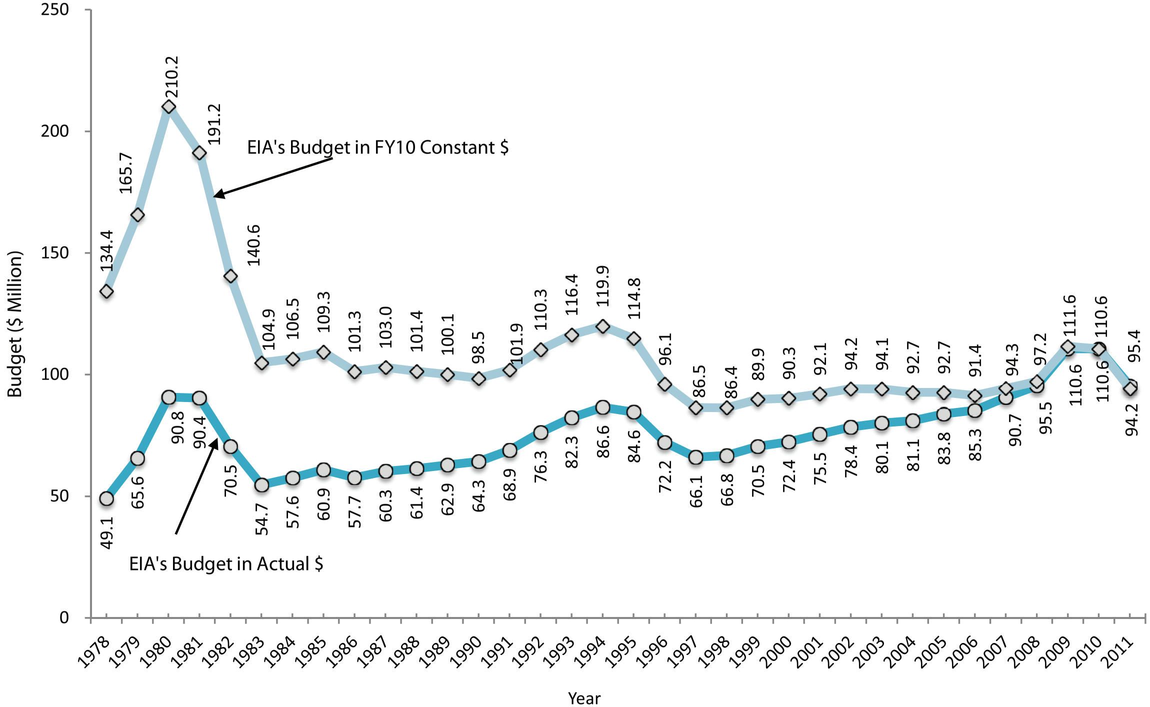


FIGURE 1-1 EIA's appropriation history, 1978–2011.
NOTE: Gross Domestic Product deflators from EIA's 2011 Annual Energy Outlook (U.S. Energy Information Administration, 2011c).
SOURCE: U.S. Energy Information Administration, Office of Resource and Technology Management.

BOX 1-1
Panel Charge

The Committee on National Statistics will convene a panel to conduct a comprehensive 30-month study of the U.S. Energy Information Administration's Commercial Buildings Energy Consumption Survey (CBECS) and the Residential Energy Consumption Survey (RECS). The panel will consider for these two surveys how to improve data quality, geographic coverage, timeliness of data releases, and relevance of the data for meeting user needs for energy end use information in the next decade and beyond. The panel will review survey design, frequency, and scope options, survey practice and operations, and the role that auxiliary data could play in improving survey coverage and editing and imputation methods. The panel will issue a letter report by spring 2010 that comments on design and data collection options for the 2010 CBECS to enable it to support U.S. Department of Energy program information needs, reduce respondent burden, and increase the quality and timeliness of the data.* The panel will issue a final report at the conclusion of a 24-month study that makes recommendations for the design and conduct of CBECS and RECS and the dissemination of CBECS and RECS data for the next decade and beyond, including consideration of the level of resources likely to be required in comparison with the current survey program. The last 6 months of the study will be used for dissemination.

———

*See Appendix F for the panel's letter report.

of data available to policy makers seeking to make informed decisions and plan for the future by, for example, ensuring that energy production is in line with needs and that savings opportunities are identified wherever possible. Congress and the executive branch relied on EIA data to inform analyses evaluating the impact of the American Power Act of 2010, which was proposed to regulate emissions of greenhouse gases (U.S. Department of Energy, 2011b). EIA data also informed an analysis of several provisions of the discussion draft of the Domestic Manufacturing and Energy Jobs Act of 2010 (U.S. Department of Energy, 2011b).

Government investments in energy programs have long shaped in-

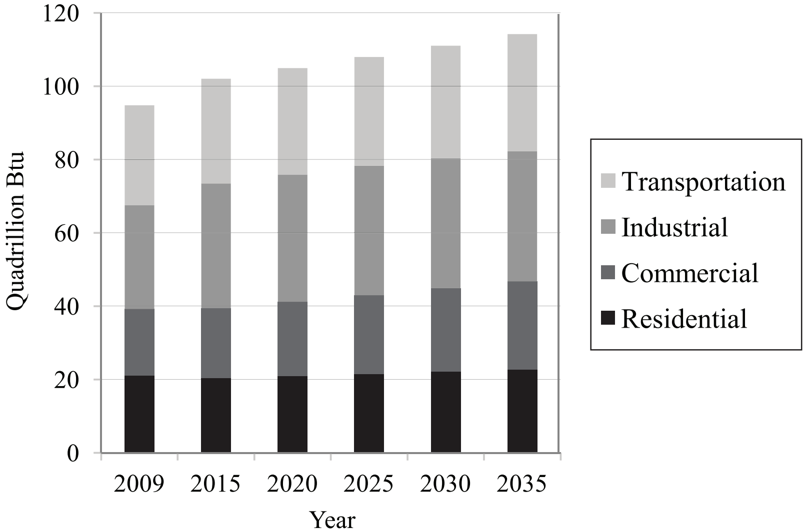


FIGURE 1-2 Primary energy use by end use sector, 2009–2035 (quadrillion Btu). SOURCE: Annual Energy Outlook 2011 (U.S. Energy Information Administration, 2011c).

novation and growth in America, but the appropriate size and scope of these subsidies has always been a topic of debate among policy makers and the public, and subsidies for emerging fuel types have been particularly controversial (Pfund and Healey, 2011). Historical data on federal energy subsidies show that there has been strong support for oil and gas over the years and that subsidies for nuclear energy during the early days of support for this technology dwarfed all other energy subsidies (Pfund and Healey, 2011). At present there is great interest in renewable forms of energy, although subsidies for renewable energy have so far been significantly lower than the subsidies for other forms of energy during the early days of each subsidy's life (Pfund and Healey, 2011).

As priorities shift, it is important to develop an increasingly clear understanding of energy demand and consumption patterns in order to evaluate the economic and environmental implications of subsidies, credits, and programs across the energy sector and to avoid investments in the wrong programs, standards, or technologies, which could be irresponsible and wasteful. Making wise choices regarding government support for the energy sector is most likely if those choices are supported by valid and reliable data.

The rapid pace of changes in technology and lifestyles, combined with ongoing building renovations and stock turnover makes it vital to perform periodic data collections that can identify trends and emerging issues. Data from the CBECS and RECS guide the development of building energy-rating tools, such as the HERS index, a scoring system established by the Residential Energy Services Network to quantify the energy performance of residential buildings, as well as simulation programs, such as EnergyPlus developed by DOE to model energy use and optimize building design in both the commercial and residential sectors. "Deep retrofit" projects that involve major renovations to an existing structure undertaken to increase energy efficiency are of increasing interest, and their effectiveness is being researched. "Net-zero energy" buildings that produce as much energy as they use are another emerging area of interest for which energy consumption data are necessary.

Furthermore, with the increase in the number of electrical devices in homes and workplaces, the role of appliance and commercial equipment standards that manufacturers must adhere to is increasingly important. Data from the CBECS and RECS are used to gather information about the quantity of Energy Star products in the market and about their energy consumption. The surveys are also used to develop and revise guidelines for product specifications to increase efficiency.

Over the past few decades miscellaneous energy use has grown rapidly, and researchers have been warning about the lack of adequate data concerning this area of energy consumption (Meier et al., 1992). Miscellaneous, or residual, energy use is generally defined as energy use other than the traditional end use categories, such as space heating and cooling, water heating, clothes drying, refrigeration, cooking, and lighting. In other words, this category includes a wide range of end uses, such as televisions, computers, electronic picture frames, ceiling fans, and so on. Residential miscellaneous energy consumption today accounts for more energy use than any other single end use in homes, and some devices consume particularly large amounts of energy (Roth et al., 2008). The Electric Power Research Institute estimates that if every household in the United States had a digital picture frame, five new natural gas power plants would be needed to operate these (Mansoor, 2008). In the commercial sector, miscellaneous end uses (such as medical devices and video displays) are expected to account for approximately 40 percent of all electricity consumption (U.S. Energy Information Administration, 2011c).

Inadequate measures of miscellaneous energy use can result in under-estimates of the energy consumed by the rapidly growing use of consumer electronics and devices and, consequently, an overestimate of the energy use associated with more traditional activities such as heating, cooling, and lighting. This makes it difficult to evaluate the success—or lack thereof—of energy efficiency programs and could result in costly overinvestments in electricity generation and infrastructure, similar to what happened in the 1980s in the United States (Gold and Elliott, 2010).

An NRC committee argued that several recent trends in the energy industry, such as industry consolidation, a greater reliance on automation, and increasingly congested transmission corridors, have increased the vulnerability of the energy supply to terrorist attacks (National Research Council, 2002). The energy infrastructure is also vulnerable to natural disasters. When Hurricane Katrina struck in the summer of 2005, for instance, it hit a region critical to U.S. oil production and distribution, and sent prices surging. In addition to these risks at home, the U.S. dependence on foreign energy sources also contributes to instability. For example, about half of the petroleum consumed by the country in 2010 was imported (U.S. Energy Information Administration, 2011d). The growing vulnerability of the energy supply system to terrorist attacks and natural disasters is another factor that makes it important to develop a thorough understanding of the country's energy needs.

The Energy Independence and Security Act of 2007 (EISA)[1] directed EIA to "establish a 5-year plan to enhance the quality and scope of the data collection necessary to ensure the scope, accuracy and timeliness of the information needed for efficient functioning of energy markets and related financial operations." (Appendix A includes the relevant section of the legislation.) The law highlighted the need to assess the gaps in the data obtained and compiled by EIA. Many data users have reported that such gaps represent the biggest shortcoming of EIA energy consumption surveys because in most cases the small sample sizes resulting from the low budgets make the data inadequate, particularly for state-level analyses. The CBECS and RECS sample sizes are approximately 6,000 cases nationally, although with funds from the American Recovery and Reinvestment Act of 2009, the most recent RECS was expanded to include 12,083 households, which made it possible to release data individually for 16 states for the first time since the survey was launched.

[1]Public Law 110-140, 110th Congress (December 19, 2007). The quotation is from § 805(a)(1).

While the increase in the RECS sample size was a promising development, the fiscal year 2011 budget reduced EIA's funding by 14 percent compared to fiscal year 2010, which caused EIA to suspend work on a large number of products and services, including the 2011 CBECS. This is a concern because the CBECS is the only national source of data on energy consumption in commercial buildings (Michaels, 2010) and because for many purposes CBECS is the only source of data on other building characteristics as well. As such, policy decisions that will shape the future depend on it.

The panel believes that high-quality, valid, and reliable data are essential to informed policy making and that they enable rapid response to changing needs, support efficient markets for end use devices, and spur innovation by guiding cutting edge research. Extended lags or gaps in the collection of energy consumption information are of a particular concern because most missing data cannot be filled in retrospectively. Compromises in the design of the surveys, such as sample sizes that are too small to be adequate and shortcuts that affect the data collection approach, are more subtle problems that also affect the quality of the data available for research and decision making. Estimates based on surveys that are conducted with insufficient resources are subject to large uncertainties, which in turn affect the confidence in conclusions based on analyses that depend on the data.

The need is especially high for information about energy demand, which is as important for energy security as information about energy supply is. While information about current energy supplies and energy reserves is available from a variety of sources, including other EIA surveys, our understanding of how energy is used, where it goes, and the services it provides remains woefully inadequate. In the case of energy supply, there are programs and mechanisms in place to collect data about energy resources, and there are a limited number of energy types that must be considered (for example, petroleum, coal, nuclear, and renewables). On the other hand, the energy consumption community is much more diffuse and does not have significant data collection assets, while at the same time it has to consider a large and growing number of end uses.

ISSUES FOR THE PANEL

In 2009, EIA asked NRC to conduct a comprehensive review to determine how the CBECS and RECS can best take advantage of recent developments in survey methods and to ensure the relevance of the data for meeting increased user needs in a world characterized by a rapidly changing energy landscape. Many of the design and operational procedures for the CBECS and RECS were developed in the 1970s and 1980s, and resource limitations during much of the time since then have prevented EIA from making significant changes to its survey methodology and operations. This study is an opportunity to focus on the future of the two surveys, and the study charge directs the panel to make recommendations "for the next decade and beyond."

In response to this charge, the panel appointed by NRC undertook a range of activities over the course of approximately two years. The panel met with EIA on several occasions to learn as much as possible about the designs of the surveys and the challenges encountered. The panel also met with EIA's data collection contractors to understand the specifics of the activities they performed as part of their contractual agreements with EIA and to obtain their perspectives on the surveys. The panel also discussed data collection strategies and methodological considerations with researchers conducting their own surveys in the energy sector, as well as with researchers conducting surveys on other topics with similar challenges.

An important aspect of the panel's work was to gain a thorough understanding of data user needs. A data user workshop brought together researchers and other data users from the government, commercial, and academic sectors and provided the panel with an opportunity to hear about the most pressing data needs.[2] The panel also received both formal and informal input from other researchers and organizations who did not participate in the workshop. For a list of data users who provided input see Appendix B (the panel reached out to several additional data users who failed to provide feedback). While publishing a written report of the data user input was not within the panel's charge, data user considerations were at the heart of the panel's deliberations. The common themes that emerged are woven throughout the discussion in this report and played an important role in shaping the recommendations.

The panel's research and deliberations made clear the need for an energy consumption data collection program that can accurately characterize changes in technology and identify trends. Doing so will require maintaining data consistency over time, but in order to meet user needs the data collection program will also have to be nimble, flexible, responsive, and timely.

[2] Presentation slides from this workshop are available on the Committee on National Statistics website at http://www7.nationalacademies.org/cnstat/Presentations%20Main%20Page.html [December 2011].

In addition to its role in informing policy making, the data collected by the surveys should also serve as catalysts for innovation. Specifically, in an ideal world the United States would have energy consumption data collection programs characterized by the following:

- Annual data releases with minimal lags between the time of the data collection and the time of publication.
- Sample sizes that are large enough to enable valid state-level analysis for all states.
- An integrated survey design with all energy-consuming units (buildings, housing units, and so on) represented in the sample.
- Questionnaires and data collection forms that capture all important uses of energy.
- A longitudinal component of the study design to inform a better understanding of trends.

The panel, of course, understands that bringing the surveys closer to the ideal would come at a significant cost, and it is realistic about the need to find a balance between data needs and budget constraints. Indeed, the panel's charge calls for "consideration of the level of resources likely to be required in comparison with the current survey program," which the panel understood as the need for recommendations that are the result of careful consideration of these trade-offs.

When possible, the recommendations in this report are accompanied by a brief discussion of their cost implications, although the actual costs can vary greatly depending on the method of implementation selected. For example, many of the recommendations described in the individual sections are related and would likely be most efficient if implemented together, which would also lead to cost savings compared to the costs associated with any particular recommendation carried out independently of the others. Some of the panel's suggestions are recommendations to evaluate the relative costs and benefits of a change to the design of the surveys or operational procedures. The reader should also note that although some of the recommendations involve an increase in survey costs in the short term, they were proposed by the panel based on the assumption that these increases will translate into more efficient surveys in the longer term.

On some issues the panel recommends alternatives for consideration by EIA, and the relative advantages of one option compared to another may depend on the amount of funding available and expected for subsequent years. Survey design is a careful balance between survey errors and costs. The precision of the survey statistics is highly dependent on the resources that are invested into the collection and production of the data. For a given budget, survey design options can be evaluated to identify the option that leads to the smallest total survey error (Groves and Lyberg, 2010). Although the panel was not asked to base its recommendations on specific assumptions about the budget, its deliberations were centered around the need to be cost effective and achieve an optimal balance between data quality and costs. The last chapter of the report takes a slightly different perspective, revisiting some of the ideal program characteristics highlighted above and discussing some ideas that may be less feasible in the short term but which may help EIA envision the energy consumption data collections of the future.

The report discusses some overarching, transformational recommendations that, if implemented, would represent major changes in the designs of the surveys. Other recommendations are more technical and incremental in nature, and are focused on smaller-scale revisions that could improve and update the surveys. The two surveys will be discussed separately, which results in some duplication in the text and overlap in the recommendations.

2

Historical Background

The U.S. Energy Information Administration (EIA) is the primary government agency responsible for collecting, analyzing, and disseminating energy statistics. A section of the 1977 Department of Energy Organization Act established EIA as the single federal government authority for energy information. The law states:

> The Administrator shall be responsible for carrying out a central, comprehensive, and unified energy data and information program which will collect, evaluate, assemble, analyze, and disseminate data and information which is relevant to energy resource reserves, energy production, demand, and technology, and related economic and statistical information, or which is relevant to the adequacy of energy resources to meet demands in the near and longer term future for the Nation's economic and social needs.[1]

Under EIA's current organizational structure, the energy consumption surveys are housed in the Office of Energy Consumption and Efficiency Statistics, under the assistant administrator for energy statistics. (See http://www.eia.gov/about/organization_chart.cfm [December 2011] for EIA's current organization structure.) Other units in the energy statistics division are responsible for collecting and disseminating information related to energy supply. These units are organized around energy types—the Office of Electricity, Renewables, and Uranium Statistics; the Office of Oil, Gas and Coal Supply Statistics; and the Office of Petroleum and Biofuels Statistics.

[1] P.L. 95-91, 42 U.S.C.A. § 7135(a)(1).

Energy consumption and efficiency analysis activities are part of a separate office, under the assistant administrator for energy analysis.

Although the EIA organizational structure has undergone several changes since the establishment of the agency, the primary mission of the unit responsible for the consumption surveys has always been to collect and disseminate information about the demand and consumption of energy (French, 2007). The energy consumption and efficiency statistics unit organizes its work around energy-use sectors, including commercial, residential, and manufacturing, as described below:

Commercial: The Commercial Buildings Energy Consumption Survey (CBECS) covers energy consumption in the commercial sector. For its consumption surveys, EIA defines a *building* as a structure that is completely enclosed, with walls that extend from the foundation to the roof, and a *commercial building* as a building intended for human occupancy, in which at least half of the floor space is used for a purpose that is not residential, industrial, or agricultural. Based on the definition used by EIA, institutional and organizational buildings such as schools, libraries, correctional institutions, and houses of worship are included in the CBECS. On the other hand, unenclosed energy-consuming structures, such as street lights, pumping stations, and billboards, are not included. Some enclosed structures not intended for human occupancy are also excluded, such as cooling towers and monuments.

Residential: The Residential Energy Consumption Survey (RECS) covers energy consumption in occupied, primary housing units. Vacant and seasonal housing is not included, nor are group quarters, such as prisons, nursing homes, and college dormitories.

Manufacturing: The Manufacturing Energy Consumption Survey (MECS) accounts for energy consumption by manufacturing establishments with five or more employees. Energy consumption by non-manufacturing industrial sectors such as agriculture, mining, and construction is not captured by the MECS.

From 1983 to 1994 EIA's Residential Transportation Energy Consumption Survey (RTECS) collected data on energy consumption by household highway vehicles, which account for nearly a third of domestic energy consumption, but a lack of funding caused this survey to be discontinued despite the importance of the sector. Another energy-use sector not currently covered by any of EIA's energy consumption surveys is energy suppliers.

EIA's funding is based on annual appropriations from Congress, with the President's budget request for EIA falling under the authority of the House and Senate appropriations subcommittees on energy and water development. The 1992 Energy Policy Act (EPACT, P.L. 102-486) introduced a requirement to conduct surveys of residential and commercial energy consumption at least every three years, and a survey of manufacturing energy consumption every two years, but EIA never had the funding necessary to follow the mandated frequencies (French, 2007). In the mid 1990s the CBECS, RECS, and MECS were moved to a quadrennial schedule, and as discussed, the RTECS was discontinued.

The panel's charge for this study was to review EIA's energy consumption surveys in the commercial and residential sectors. The panel did not evaluate the availability of energy consumption data in any of the other sectors. However, data users voiced concerns related to the lack of adequate energy consumption data in the transportation field in particular. It appears that although there are efforts to fill the gap (for example, by the Oak Ridge National Laboratory and by states that produce their own estimates), there is a need for national, standardized data on energy consumption in the transportation sector, comparable to the data produced by the CBECS and RECS in the commercial and residential sectors.

EIA interprets the EPACT reference to the collection of residential energy consumption data as a directive to conduct the RECS. However, an external review of the surveys conducted in 2006 argued that the directive includes the collection of household transportation data—in other words, the RTECS as well (French, 2007). The 2009 RECS incorporated a small number of new questions on household transportation, but it is difficult to imagine that the already lengthy RECS could accommodate more questions on this topic.

The year 2011 was especially challenging for EIA, and particularly for the CBECS. After a lengthy delay in the processing of the 2007 CBECS, which had been based on a new design that was expected to be more cost effective, EIA announced, "Because the data do not meet EIA standards for quality, credible energy information, neither data tables nor a public use file will be released" (U.S. Energy Information Administration, 2011e). A second blow came in the form of a significant budget cut, when despite the President's budget request of $129 million, the fiscal year 2011 budget provided only $95.4 million for EIA, a reduction of 14 percent from the previous year (U.S. Department of Energy, 2011b). The budget cut led to the unfortunate suspension of work on the 2011 CBECS, which had just been awarded to a new data collection contractor.[2] The panel assumes that reinstating the CBECS will be a priority in fiscal year 2012, given the lack of alternative sources of information about commercial buildings in the United States.

[2] EIA's energy consumption surveys are conducted by data collection contractors. EIA's involvement and the division of responsibilities between the organizations have varied over the years. Typically, the data collection contractor performs the sample design and selection activities, hires and trains interviewers, collects the data, and prepares the cleaned and weighted files for EIA. For the CBECS, EIA currently does the editing, imputation, estimation, and modeling in house. For the RECS, EIA does only the modeling in house.

3

Commercial Buildings Energy Consumption Survey Program History and Design

The primary purpose of the Commercial Buildings Energy Consumption Survey (CBECS) is to provide comprehensive information about commercial-sector energy use in the United States, and it does this by collecting data from a national sample of commercial buildings. Both the scope of the survey and the approach to data collection have evolved over the years, driven mostly by practical considerations. Table 3-1 summarizes the history of the survey, which started out with a design that aimed to represent all nonresidential buildings in the four census regions and has moved to a more focused approach that includes only buildings that are predominantly commercial in nature (50 percent or more of the floor space) and excludes certain commercial buildings that are difficult to survey and that do not represent a large proportion of commercial energy use.

The smallest level of geographic detail for which the CBECS data are available is the census division, which is a relatively large geographic area, covering many different climates and, therefore, different energy consumption characteristics. For example, the Mountain West division includes eight states, from New Mexico to Montana (see Appendix C). This limits the analysis that can be conducted at the subnational level using data from the CBECS.

In addition to the energy consumption and expenditure data, the CBECS also collects various details about building characteristics, such as a building's physical structure, activities performed, equipment used, and energy source. Typically, the CBECS is administered in two data collection

TABLE 3-1 CBECS Data Collection Years and Survey Characteristics

Year	Completed Interviews	Survey Characteristics
1979	6,221	The survey was called the Nonresidential Buildings Energy Consumption Survey (NBECS). The sample was based on a nonresidential building sample that had been developed by the data collection contractor for another survey. The design included nonresidential buildings of all sizes as well as buildings that were predominantly residential or industrial if commercial activity was present. In addition to the building interview, information about energy consumption and expenditures was also collected from the building's energy suppliers. The smallest level of geographic detail for which data were available was the census region.
1983	7,140	The survey was a follow-up to the 1979 survey. The same buildings were surveyed again, along with a sample of new construction. The smallest level of geographic detail for which data were available was the census region.
1986	6,072	The survey was renamed the Commercial Buildings Energy Consumption Survey (CBECS). The sampling frame was redesigned and based on an area-based listing approach, in which the area sample was supplemented with lists of large and special buildings. The new approach excluded buildings of 1,000 square feet or less and limited the sample to buildings that were predominantly commercial. The smallest level of geographic detail for which data were available was the census division.
1989	5,876	No major changes to the sample design. A sample of new construction was added. The smallest level of geographic detail for which data were available was the census division.
1992	4,806	No major changes to the sample design. A sample of new construction was added. The smallest level of geographic detail for which data were available was the census division.
1995	5,766	Commercial buildings that were part of industrial facilities and parking garages were dropped from scope. A sample of new construction was added. The survey was moved from paper-and-pencil personal interview to computer-assisted personal interview. The smallest level of geographic detail for which data were available was the census division.

TABLE 3-1 Continued

Year	Completed Interviews	Survey Characteristics
1999	5,430	The survey was conducted by computer-assisted telephone interview, using the same buildings selected in 1995. A sample of new construction of buildings over 10,000 square feet was added (and data for smaller new buildings were imputed). Starting this year, energy consumption and expenditures data were collected from the building respondents, and energy suppliers were contacted only if the data provided by the building respondent were inadequate. The smallest level of geographic detail for which data were available was the census division.
2003	5,215	The sampling frame was redesigned and a new area frame listing was created for the first time since 1986. The area frame was supplemented with lists of large and special buildings. New data collection procedures were implemented for shopping malls, universities, and hospitals. The smallest level of geographic detail for which data were available was the census division.
2007		The area frame was updated based on a commercial version of the USPS delivery sequence file. No data were released.
2011		Data collection currently suspended (as of December 2011).

SOURCE: U.S. Energy Information Administration (http://www.eia.gov/emeu/cbecs/contents.html [December 2011]).

stages: a building characteristics survey and an energy suppliers survey. During the first stage of the data collection, interviewers visit the buildings selected into the sample and ask a representative, such as the building's owner, manager, or other knowledgeable person, to complete the survey. During the second stage of the data collection, the energy suppliers (e.g. utilities) of the buildings for which inadequate information was provided in the first stage are contacted to obtain usage and expenditure data for those buildings from the supplier's records. This usually happens if the building respondent is not able to report energy usage and cost information or if the information provided falls outside the range of "likely" energy consumption, determined based on regression models that take into account responses from previous years. In a typical CBECS, the energy suppliers for about half of the buildings in the sample need to be contacted (Michaels, 2010). Unlike the building survey responses, the suppliers' responses are mandatory under the Federal Energy Administration Act. Box 3-1 shows the topics included

BOX 3-1
2003 CBECS Building Questionnaire Main Topics

Building size and age
Square footage
Building structure
Floors
Escalators and elevators
Year of construction

Principal building activity
Building activity or activities
Specific building activity
Building capacities
Multibuilding complex information

Occupancy and operating hours
Building ownership
Number of occupants
Vacant space
Hours in use
Employees

Energy sources, end uses, and equipment
Energy end uses
Energy sources
Heating sources
Heating equipment
Cooling sources
Cooling equipment
Heating/cooling system conservation features
Water heating sources
Water heating equipment
Cooking sources
Manufacturing sources
Electricity generation
Electricity/natural gas purchasing
Bottled gas/LPG/propane usage
Wood usage

Miscellaneous equipment
Specialized space use
Specialized equipment
Refrigeration
Computer equipment
Office equipment
Lighting
Lighting/other conservation features

Electricity usage
Electricity consumption
Electricity expenditures
Electricity supplier(s)

Natural gas usage
Natural gas consumption
Natural gas expenditures
Natural gas supplier(s)

Fuel oil usage
Fuel oil consumption
Fuel oil expenditures
Fuel oil supplier(s)

District steam usage
District steam consumption
District steam expenditures
District steam supplier(s)

District hot water usage
District hot water consumption
District hot water expenditures
District hot water supplier(s)

SOURCE: U.S. Energy Information Administration (http://www.eia.gov/emeu/cbecs/forms.html [December 2011]).

on the 2003 CBECS building questionnaire. For more information on the CBECS data collection instruments, including the shopping mall protocols and energy supplier surveys, see http://www.eia.gov/emeu/cbecs/forms.html [December 2011].

The basic unit of data collection and analysis for the CBECS is the building. The U.S. Energy Information Administration (EIA) originally considered using establishments instead of buildings as the energy-consuming unit that would best explain energy use in the commercial sector, but this approach was abandoned during the initial planning of the survey program for two reasons: (1) energy use in the service industry (which many commercial buildings belong to) is more closely associated with the characteristics of a building than with the characteristics of the establishments that occupy the building (for example, an office building is different from a warehouse, regardless of the profile of the companies that may be sharing the space within the same building); and (2) if several establishments share space in the same building, energy data are more likely to be available at the building level than at the establishment level (French, 2007). It has always been recognized, however, that no matter what unit of data collection is used, there will always be cases in which energy data are not available for the level of the data collection unit and there may also be differences by energy type within the same data collection unit.

Exceptions to the building-level energy-supply model can and do cause coverage problems, for example, in cases where tenants in a shopping mall receive their power individually or cases where energy is delivered to an entity larger than a building, such as a college campus. This led EIA to experiment with slightly different data collection procedures for shopping malls and college campuses and to customize questionnaires for establishments within a mall as well as for a mall manager who can provide information about the building overall.

Today the CBECS is the only national survey of the characteristics and energy use of commercial buildings in the United States, which presents a particularly difficult challenge for EIA. Without another nationwide source of information about the stock of commercial buildings in the United States, there is no comprehensive list of buildings that could help in the construction of a sampling frame for the CBECS. Thus EIA periodically builds a new area probability sampling frame for the CBECS. Full area frame listings have been created twice, in 1986 and 2003.

By definition, field listings are resource intensive, but relying on sources that are not comprehensive to update the sampling frame leads to coverage problems. As a result, the CBECS sample design has undergone numerous revisions over the years, as EIA has attempted to address both the cost and the coverage issues. Generally speaking, most rounds of the CBECS have relied on a combination of an area frame and a list frame, merging existing lists of commercial buildings from a variety of sources and adding them at the second stage of the area frame sample.

The area frame is based on field listings of commercial buildings within specified geographic areas. Over the years, the sample typically included between 100 and 130 counties or groups of counties, which served as primary sampling units (PSUs). Within the PSUs smaller geographic areas were randomly selected, and then all commercial buildings within these areas were listed and stratified. In the final step a sample of buildings was randomly selected from each stratum. The sampling frame is updated from one data collection to the next in order to account for changes, such as new construction. The updates are typically based on information from local sources or databases of commercial projects, and new construction is usually sampled at a higher rate to enable EIA to produce estimates for this subpopulation.

Given the variations in building size and in the intensity of energy use among commercial buildings, the sampling rates for large buildings that use large amounts of energy must be higher than would be feasible to accomplish through the area-probability sampling alone. To assure that these types of buildings are adequately represented in the sample, the area frame is supplemented with a list frame of special buildings, such as very large buildings, hospitals, and government buildings. The two frames are then matched, and duplicates are eliminated.

Over the years, EIA has considered a variety of lists to use as the primary source for the sampling frame, including tax records, mail delivery points, insurance lists, customer information from utilities, Federal Emergency Management Agency records, and aerial and satellite photographs (French, 2007). Some of these alternatives have been evaluated empirically, while others have not. One feasibility study, which was conducted in 1981 in the service areas of Seattle Power and Light Company and Portland Electric, was performed to evaluate the use of electric utility customer information in building a sampling frame. The study concluded that differences between the way that records are kept at different utilities and the utilities' varying degree of motivation to respond to an energy survey would represent significant challenges (French, 2007).

For the last round of the CBECS in 2007, EIA's data collection contractor, the National Opinion Research Center (NORC), proposed a new sample-updating approach that was expected to be cost efficient. After supplementation with lists of buildings that were over 200,000 square feet in size, including federal buildings, colleges, hospitals, and airports, the 2003 area frame was updated using a commercial version of the U.S. Postal Service (USPS) delivery sequence file (DSF), purchased from a vendor licensed by USPS. The DSF is the USPS's list of all delivery points in the United States. To use the DSF for updating, the list had to be matched to the addresses in the second-stage area frame and the duplicates removed. NORC reported that the removing of duplicates turned out to be a major challenge, in part because of imprecise address records. The use of this new approach to update the survey frame, along with implementation errors made by the contractor, ultimately led to the data quality problems that prevented EIA from being able to release the 2007 CBECS data (U.S. Energy Information Administration, 2011e).

The CBECS was initially conducted in the form of in-person interviews, using paper and pencil. Since 1995, however, it has been conducted using computer-assisted personal interviewing. The 1999 data collection was an exception because, in an effort to control costs, it was done by computer-assisted telephone interviewing. This was made possible by reinterviewing the sample from the 1995 CBECS. The supplier survey has always been conducted by mail.

4

Residential Energy Consumption Survey Program History and Design

The goal of the Residential Energy Consumption Survey (RECS) is to collect comprehensive information about residential energy use in the United States. Unlike the Commercial Buildings Energy Consumption Survey (CBECS), the RECS is not a building survey, but rather it is a survey of a national sample of housing units. In 2009 the sample represented approximately 113.6 million housing units in the United States. The design of the RECS is more straightforward than that of the CBECS, and the U.S. Energy Information Administration (EIA) is able to benefit from the vast experience of the federal statistical system and of other research organizations with nationwide household data collections. The scope of the survey has not changed significantly since its inception except for the EIA's efforts to meet an increasing demand for state-level data. The 1993 RECS was the first survey that produced state-level data for the four most populous states in the United States: California, Texas, New York, and Florida. In 2009, EIA was able to triple the RECS sample size and produce state-level data for 16 states, which together account for approximately 65 percent of total household energy consumption in the United States. This expansion also enabled EIA to produce estimates for groups of states that together represent a smaller geographic entity than a census division. Table 4-1 summarizes the evolution of the RECS design.

The RECS collects information about the physical characteristics and energy consumption of the surveyed housing units as well as the demographic characteristics of the households. As is the case with the CBECS,

TABLE 4-1 RECS Data Collection Years and Survey Characteristics

Year	Completed Household Interviews	Survey Characteristics
1978	4,081	The first survey was called the National Interim Energy Consumption Survey (NIECS). The sample was based on a household survey sample developed by the data collection contractor for a previous survey.
1979	4,033	The Household Screener Survey (HSS) was conducted as a follow-up to the NIECS, with the same sample being recontacted.
1980	6,051	A new multistage area probability sample was developed based on the 1970 decennial census. The name of the survey was changed to the Residential Energy Consumption Survey (RECS). A rotating panel component was introduced. The smallest level of geographic detail for which data were available was the census division.
1981	6,269	No major changes to the sample design. The smallest level of geographic detail for which data were available was the census division.
1982	4,724	No major changes to the sample design. The smallest level of geographic detail for which data were available was the census division.
1984	5,682	A new sampling frame was developed based on the 1980 decennial census. The smallest level of geographic detail for which data were available was the census division.
1987	6,229	No major changes to the sample design. The reference period was changed from April through March to the calendar year. The smallest level of geographic detail for which data were available was the census division.
1990	5,095	No major changes to the sample design. The smallest level of geographic detail for which data were available was the census division.
1993	7,111	A new multistage area probability sample was developed based on the 1990 decennial census. The rotating panel component was discontinued. State-level data were available for the four most populous states (California, Texas, New York, and Florida) in addition to the data by census division.

TABLE 4-1 Continued

Year	Completed Household Interviews	Survey Characteristics
1997	5,902	No major changes to the sample design. State-level data were available for the four most populous states.
2001	4,822	No major changes to the sample design. State-level data were available for the four most populous states.
2005	4,382	A new sampling frame was developed based on the 2000 decennial census. State-level data were available for the four most populous states in addition to the data by census division.
2009	12,083	Most of the sampling frame was based on a commercial version of the USPS delivery sequence file. Most of the segments were new, but some segments from the 2005 sampling frame were updated for 2009. State-level data were available for 16 states in addition to the data by census division.

SOURCE: U.S. Energy Information Administration (http://www.eia.gov/consumption/residential/index.cfm [December 2011]).

the RECS is conducted in two data collection stages: a housing unit survey and an energy supplier survey. In the case of some rental units, information is collected from a rental agent in addition to the residents, using a modified version of the questionnaire. And, as is the case with the CBECS, only the energy supplier survey is mandatory. Box 4-1 shows the topics included on the RECS household survey. For more information on the RECS data collection instruments, including the full list of questions included in each of the questionnaires, see http://205.254.135.7/consumption/residential/surveyforms.cfm [December 2011].

The RECS uses a multistage area probability sample of occupied housing units that serve as the primary residence for a household for at least 6 months in a particular year. Under the current design the first step in the sample selection for the RECS involves randomly choosing counties to serve as primary sampling units (PSUs). Over the years, the sample has included anywhere from 103 to 441 PSUs. Typically the counties selected are then subdivided into segments composed of census blocks, and a sample of these segments is randomly selected. Within the selected segments, field staff members list all housing units in order to create a housing unit frame.

BOX 4-1
2009 RECS Household Questionnaire Main Topics

Housing unit characteristics
Kitchen appliances
Home appliances and electronics
Space heating
Water heating
Air conditioning
Miscellaneous topics
Fuels used
Housing unit measurements
Fuel bills
Residential transportation
Household characteristics
Energy assistance (supplement)

SOURCE: U.S. Energy Information Administration (http://205.254.135.7/consumption/residential/surveyforms.cfm [December 2011]).

In 2009 a commercially available version of the U.S. Postal Service (USPS) delivery sequence file (DSF) was used to perform the listing, where feasible. A traditional field listing was used when DSF-based listing was not possible, as was the case in many rural segments. The sampling frame is typically updated from one data collection to the next in order to account for such changes as new construction.

In the early days, the household portion of the RECS was conducted in the form of in-person interviews, using paper and pencil. Beginning with the 1997 data collection, the survey was moved to computer-assisted personal interviewing. The supplier survey was a mail survey before 2009, when the web became the primary mode of data collection.

5

Redesigning the Commercial Buildings Energy Consumption Survey

This section details the panel's recommendations for redesigning the Commercial Buildings Energy Consumption Survey (CBECS). During the panel's deliberations two areas emerged as the top priorities for revision: (1) the timeliness and frequency of the CBECS and (2) data gaps. Thus a discussion of these two topics follows, along with the panel's recommendations for changes in these areas. Some of the recommendations involve major changes in the way the survey is administered; ideally, these changes would all be implemented together, so as to maximize the benefits. The major recommendations include implementing a rotating sample design with a longitudinal element for part of the sample and transitioning to a multimode data collection. The second half of this chapter offers additional suggestions for updating the CBECS with changes that are more incremental in nature.

TIMELINESS AND FREQUENCY OF THE CBECS DATA

The panel's research, including discussions with users of the CBECS data, showed that one of the biggest concerns about the CBECS is the delay that often occurs between the time the data are collected and the time they are released to the public. For example, the public-use microdata file for the 2003 CBECS was released in November 2008. A related issue is the survey's quadrennial schedule, which does not meet the legislative requirements and is considered inadequate by at least some of the data users. An important aspect of this concern is simply the face validity of data that have been collected years before publication. In addition, data users argue that some characteristics of building energy consumption change relatively frequently and therefore that energy consumption data would be more useful if they were captured on a more frequent basis.

Just before this report went to print, the data that had been collected for the most recent CBECS were deemed unusable four years after the reference period (2007). Even in the best case, the result would have been an eight-year time lag between one round of the survey and the next round that was usable. However, a decision to suspend the 2011 data collection because of insufficient funding will make the lag even greater. These are unusual circumstances, and they have put stakeholders who rely on the CBECS data for research and decision making in a very difficult situation.

For the purpose of discussion in the sections that follow, the panel assumes that the CBECS data collection will resume on at least a quadrennial schedule in fiscal year 2012.

Rotating Panel Design for the CBECS

One option for addressing data users' need to have quicker access to the data is to implement a rotating sample design, which would make more frequent data releases possible. This could be done by dividing the sample into four subsamples and collecting the data over a four-year period, instead of once every four years. Data could then be released annually, with each release containing one year's worth of new data combined with data from the previous three years in order to achieve a sufficiently large sample size that will enable EIA to release an amount of data that is comparable to what is released under the current design.

This design would have several advantages, besides offering faster access to the data. For example, while the goal would be to have roughly the same number of completed interviews at the end of the four-year cycle as would have been collected with the standard pattern of collecting data once every four years, with self-contained annual samples, the rotating design would not require a commitment to four full years' worth of interviews upfront, which could be an advantage if fluctuations in CBECS funding persist. The continuous sample design would give EIA more options when faced with a substantial budget cut, for example, by making it easier to reduce the sample size without having to sacrifice the entire survey. Furthermore, the new design would also increase EIA's ability to be flexible in terms of testing new data collection approaches because testing could be performed on one year's sample (or a subset of one year's sample), and the procedures could be fully implemented—or revised, if necessary—for the following year's sample. In some years, there will be some overlap in the reference period for the data collected by the three energy consumption surveys, which could also have some analytic advantages.

Although the transition to a new sample design and new data collection operations will involve some temporarily increased costs compared to the typical start-up costs associated with the current design of the survey, the expectation is that, at least for the most straightforward implementation of the rotating design (Option 1 described below), over time the costs would be at least comparable—and possibly lower than—the cost of conducting the survey once every four years. EIA could phase in the first subset of the sample gradually, for example, over the course of a two-year period, instead of aiming to complete a quarter of the interviews during the first year of the implementation.

An operational advantage is that the survey would not have to be resurrected every four years, and there may be some cost savings and data quality improvements associated with uninterrupted operations, since there would be more continuity in the activities performed by staff and, in particular, the field interviewers, who require extensive training. It is possible that the decrease in the number of interviews conducted each year might lead to a decrease in efficiency in terms of ability to provide interviewers with an optimal caseload in their respective geographic areas. However, as discussed later, the panel encourages EIA to explore the possibility of collecting some of the data by web, which would also contribute to a drop in the field interviewer workload and would likely necessitate the restructuring of interviewer assignments. This could be accomplished, for example, by crosstraining interviewers to perform additional tasks.

The rotating design would also integrate well with a longitudinal approach, which would involve following a subset of the sample over time. This would improve estimates of change and would allow researchers to assess how changes in the economy or in incentive programs affect energy consumption patterns. EIA experimented with longitudinal data collection for the CBECS in the 1980s but found that inconsistencies in reporting and other sources of error were sometimes confounded with the actual changes of interest, which limited the usefulness of the longitudinal data (French, 2007). However, techniques for reducing inconsistencies and controlling for confounders in longitudinal studies have become more sophisticated since that time (Lynn, 2009), which could make longitudinal data more useful now.

Data collection costs for the longitudinal subset of the sample are expected to be lower in the long run than collecting the same number of data points on a different sample every year, because information associated from the first contact with the respondent can be used to reduce the costs of the second interview. At the minimum, information about the building and the designated respondent will be useful in reducing operational costs, especially if an email address can be obtained to conduct the second interview by web. It would also be possible to significantly shorten the questionnaire and to only ask respondents to update their responses, as needed, carefully wording the questions to reduce concerns about possible biases related to the repeat interview.

The value of changing the frequency of the data collections and introducing sample overlap between different rounds of the survey will depend on a number of factors, both user-driven (e.g., the importance of estimates of change and the importance of having more frequent estimates) and technical (e.g., the correlation over time of key variables and cost implications due to changes in field operations). Although the panel cannot fully assess these factors as they apply to the EIA energy consumption surveys, it is able to suggest options that should be explored using real data.

It is worth noting that although the discussion in this section focuses on CBECS, similar approaches could be applied to the RECS as well. For simplicity's sake, the discussion in this section will be based on the assumption that CBECS has funding for a sample of 6,000 buildings every four years under the current design. The discussion will refer to the process of collecting data as "interviews" even though, in practice, data may be collected by various other means, such as via the web.

The discussion will focus first on two specific options, one of which involves including some of the buildings in the sample twice. The discussion of the two specific options is followed by a discussion of generalized versions of these options. All options presented are special cases of Option 5, so readers familiar with rotating panel designs may wish to read that section of the text first. Ultimately, EIA will have to evaluate which approach is most suitable for the CBECS, but the panel believes that an overall assessment of the costs and benefits of introducing a rotating sample design should receive priority as part of EIA's planning for the near future.

Option 1: Full Rotation

In Option 1, the data collection that currently takes place every four years is spread evenly over four consecutive years. That is, instead of sampling 6,000 buildings every fourth year, 1,500 buildings would be sampled in Year 1, a different set of 1,500 different buildings would be sampled in Year 2, and so on. An obvious advantage here is that it is possible to produce estimates that are more up to date (for example, by taking a four-year moving average every year). A disadvantage is that a "sharp snapshot" estimate taken once every four years is replaced with a "fuzzy" one taken every year. To be more precise, if θ_t is the parameter of interest, then the annual estimates θ_t, θ_{t+4}, … are replaced by estimates formed by taking averages over four years, such as ¼ × $(\theta_t + \theta_{t-1} + \theta_{t-2} + \theta_{t-3})$. Of course, it would still be possible to estimate θ_t using data from only Year t, but the estimate would be based on only one-quarter of the current sample size. Depending on the details of how the sample is distributed over the four years (e.g., spread within primary sampling units or across them), there may be significant operational disadvantages as well.

Option 2: 50% Rotation

In Option 2, the rotation pattern would be as follows, with numbers in italics denoting buildings that are in the sample for the second time:

Year 1:	1500 +	1500				
Year 2:		*1500* +	1500			
Year 3:			*1500* +	1500		
Year 4:				*1500* +	1500	
Year 5:					*1500* +	1500

.
.
.

Once the pattern has stabilized (i.e., after the first year), each building is in the sample twice. Thus the number of unique new buildings in the sample remains at 6,000 over a four-year period, but the number of interviews doubles to 12,000.

In this approach, as in Option 1, it is possible to use the data for various estimates, including θ_t, ½ × $(\theta_t + \theta_{t-1})$ and ¼ × $(\theta_t + \theta_{t-1} + \theta_{t-2} + \theta_{t-3})$. A major advantage of the rotation pattern in Option 2 is that it makes it possible to use composite estimation to improve the estimates. Composite estimation[1] can be useful in this context because this is one method that can help exploit the overlapping sample (i.e., the fact that, for a given year, half of the sample overlaps with the sample from the previous year) to gain efficiency (reduced variance). The efficiency gain depends on the correlation over time for the variables of interest—the greater the correlation, the greater the gain. Of course, the existence of sample overlap also greatly improves the quality of year-to-year estimates of change. An operational advantage of this option is that most second interviews can be conducted more cheaply (e.g., by telephone or via the web), with personal visits necessary only for nonrespondents. A disadvantage of this option is the increased burden on respondents since they would be interviewed twice rather than just once.

A much less expensive variant of Option 2 would be to keep the number of *interviews* (as opposed to buildings) steady at 6,000 every four years:

Year 1:	750 + 750					
Year 2:		*750* + 750				
Year 3:			*750* + 750			
Year 4:				*750* + 750		
Year 5:					*750* + 750	

.
.
.

In practice, a compromise between the two variants seems reasonable. The gains in statistical efficiency offered by the first variant of Option 2 can be translated into a reduction in sample size for fixed variability. However a 50 percent reduction in sample size would likely be too big for most variables of interest. A suitable reduction in sample size would be determined by deciding which variables are key ones, estimating their variances under different sample size reduction scenarios, and picking a scenario that balances costs and quality.

Option 3: (100/R)% Rotation (Generalization)

Option 3 is a special case of a more general rotation pattern with R rotation groups or panels (R = 2 in Option 2). In general, there would be

[1] For further details on composite estimation, see Appendix E.

n/R buildings in each rotation group, where n is the number of buildings (old and new) sampled in one year. As illustrated below, each year n/R new buildings enter the sample, and an equal number drop out of the sample. The part of the sample that is common between two consecutive years is displayed in italics (and thus nonitalics denotes a fresh sample).

Year 1: n/R + n/R + n/R + … + n/R
Year 2: *n/R* + *n/R* + … + *n/R* + n/R
Year 3: *n/R* + … + *n/R* + *n/R* + n/R
.
.
.
Year R: *n/R* + *n/R* + *n/R* + … + *n/R* + n/R
.
.
.

The year-to-year overlap in sampled buildings would be 100(R – 1)/R%, and each building would be in the sample on R occasions. Therefore, as R increases, so does the proportion of overlap between years, which will result in better estimates of change and better composite estimates. On the other hand, the response burden will also increase since each unit will be subjected to R interviews. Another disadvantage of having a larger value of R is that there is a longer start-up period (R years) before the rotation pattern stabilizes. That is, up until the R^{th} year, there will be some buildings that stay in the sample for fewer than R occasions.

Note that setting n = 3,000 and R = 2 gives us variant 1 of Option 2, and setting n = 1,500 and R = 2 gives us variant 2 of Option 2. Option 1 is the "degenerate" case with n = 1,500 and R = 1. (To see this, simply erase all the terms in italics for Years 2, 3, … and note that there is only one term in Year 1, namely n/R = 1,500/1 = 1,500.)

Option 4: Preserving the Four-Year Gap Between Surveys

A different approach is to introduce sample rotation while continuing to conduct the survey every four years instead of every year. Any of the previous options can be applied by simply changing years 1, 2, 3, … to years 1, 5, 9, … Thus, the general case (Option 3) becomes:

Year 1: n/R + n/R + n/R + … + n/R
Year 5: *n/R* + *n/R* + … + *n/R* + n/R
Year 9: *n/R* + … + *n/R* + *n/R* + n/R
. .
. .
. .

Having a four-year gap for a rotating panel design would have some disadvantages. One would be the possible attenuation of correlations over time, which could make it impossible to construct a reliable composite estimator for more than a small subset of the variables. In addition, there may be operational impacts caused by the large volume of building changes likely to be observed in the overlap sample following a four-year gap. Building occupants would also be more likely to change during the four years between interviews, complicating the reinterview process and introducing another confounding factor.

Option 5: F-Year Gap Between Surveys, with Rotation

The most general option along these lines would be to have R rotation groups for a survey that is conducted every F years:

Year 1: n/R + n/R + n/R + … + n/R
Year 1+F: *n/R* + *n/R* + … + *n/R* + n/R
Year 1+2F: *n/R* + … + *n/R* + *n/R* + n/R
.
.
.

Here F can represent any number of years, but a small value is preferred since correlations weaken over time. Setting F = 2, for example, would offer a compromise between conducting the survey annually and having a four-year gap between surveys. If buildings were surveyed every two years, the correlations for some key variables might still be high enough to yield good gains using composite estimation. The case R = 2, F = 2 is illustrated below.

Year 1: n/2 + n/2
Year 3: *n/2* + n/2
Year 5: *n/2* + n/2
.
.
.

Thus, with a sample size of n = 6,000, the survey would include 6,000 buildings visited every other year, with 3,000 of these being new buildings and the other 3,000 being second visits (starting in Year 3).

Additional Considerations

The above is by no means a comprehensive list of options. The goal of this section is to show that there are many alternatives to the current design that are worth consideration. For example, an approach inspired by the rotation pattern used by the Current Population Survey of the Bureau of Labor Statistics would be to have a building in the sample one year, then out of the sample for one or more years, and then back in the sample in a later year. Regardless of which approach is chosen, data releases should be accompanied by enough technical documentation about the survey design and data collection procedures that users can understand the limitations of the data and select appropriate methods for analyses.

Recommendation CBECS-1: EIA should evaluate the usefulness of implementing a rotating sample design for the CBECS to improve the timeliness of the data.

Recommendation CBECS-2: EIA should consider integrating a longitudinal element into the CBECS sample design to obtain better estimates of change.

CBECS Multimode Data Collection

Another strategy for increasing the timeliness of the data releases would be to change some of the interviews from face-to-face settings to collecting data via the web. This could shorten the data collection period significantly, and make data processing and release times faster. Nonpersonal interview modes of data collection are likely to be both necessary and very useful when implemented in conjunction with a design based on the recommendations in the previous section. An additional methodological benefit of speeding up data collection would be a shortening of the reference period, which would result in the data capturing a narrower, more precise window in time.

Although EIA has in the past considered using other modes of data collection, at this time CBECS data are still collected primarily by in-person interviewing. One reason is the initial cost of implementing a new data collection mode. Another reason is the complexity of the CBECS data collection procedures and questionnaires. With the current system of face-to-face data collection, the field interviews play a significant role in dealing with these complexities. For example, field interviewers are trained to apply CBECS definitions to determine the boundaries of a building and to identify the most appropriate respondent for the survey. They also mediate an elaborate editing and error-resolution system that is currently built into the computer-assisted personal interviewing (CAPI), and they carry hard-copy "show cards" that list the answer options for specific items and which can be handed to the respondent to assist with answering questions that may be too difficult to remember if they are only read by an interviewer. Interviewers also collect and scan utility bills when these are available. Finally, the involvement of field interviewers can often lead to higher response rates to the survey in general and to individual questions as well. Field interviewers can locate sample units that are difficult to find through other means and they can often persuade a reluctant respondent more effectively than a contact through another method, such as the telephone, mail or web. Field interviewers can also make sure that all of the questions are answered without break-offs, which reduces the levels of missing data for individual questions.

The panel acknowledges the advantages of face-to-face interviewing and that some of these may be lost by implementing other modes of data collection. However, given the costs associated with in-person data collections and the role of the Internet in today's society, it is difficult to imagine that a large-scale nationwide survey based solely on face-to-face data collection will be sustainable in the long run. The panel believes that EIA has to make it a priority to begin preparations for implementing a multimode data collection approach which will have at least a portion of the interviews being conducted online.

Although issues associated with transitioning to a new data collection mode are faced by many other large-scale surveys, some of the challenges and opportunities are more specific to the CBECS. The upfront costs of introducing a new data collection mode will be significant, but some of the costs associated with the transition may be lowered if the sample design is changed to continuous measurement. The new design will allow EIA to test and revise procedures on a subset of the sample, and implement improvements on an annual basis (and often faster). Transitioning at least a subset of the buildings to the web will free up some resources after the initial implementation period, and these resources can then be allocated to the more complex cases and possibly invested into increasing the sample size. Once a web questionnaire has been built and tested, there are virtually no costs associated with the actual individual web interviews, in contrast to face-to-face interviewing, for which cost is directly related to the number of cases. Obviously, introducing a web data collection mode does not mean that follow-up, editing, and other data collection costs can be eliminated, even for the interviews that are conducted over the web.

Collecting technical data of this nature on the web does present some methodological challenges (for example, selecting the appropriate building respondent without a face-to-face visit will be more difficult), but there is a vast and growing methodological literature concerning how best to transition data collections from one data collection mode to another and how best to integrate multimode approaches (see, for example, de Leeuw, 2005; Dillman and Smyth, 2007; Dillman et al., 2008). Based on the findings in this literature, the panel believes that these methodological challenges can be addressed. The panel also strongly encourages EIA to develop its own methodological research program to help inform the transition to web data collection in those cases where the current literature does not have adequate answers.

It is also important to note that there are certain methodological advantages associated with web data collection, including the possibility of embedding visual and other multimedia aids into the questionnaires and also the availability of metadata about each interview, which can be useful in fine tuning the data collection approach.

The panel recognizes that enabling this transition will require that difficult decisions be made in terms of eliminating some of the complexities in the data collection procedures and questionnaires. In particular, the editing and error-resolution system that is embedded into a computer-administered instrument but that relies on an interviewer to navigate will have to be revised. Furthermore, some steps will have to be eliminated altogether in order to avoid situations in which the burden of resolving unusual errors is transferred from the field interviewer to EIA staff, possibly causing further delays in the release of the data. While many of the complex procedures in question were originally introduced to improve data quality, some data users indicated to the panel that they would gladly accept less thoroughly edited data in the interest of faster access. The topic of postsurvey editing will be discussed further in the next section.

One possible approach that should be explored as part of a gradual transition to the web is to divide the sample into two sets: buildings that can be transitioned relatively easily to a web administration and buildings with more complicated characteristics that may benefit from interviewer administration. It may also be necessary to treat large buildings differently from smaller ones. A review of the case histories and interviewer debriefings would be helpful in identifying the building types for which data collection is fairly straightforward. Given that some of the data collection complexities are associated with defining building boundaries, buildings that are in the sample during more than one year and have already been visited during a previous round of the CBECS (typically large buildings) may be especially good candidates to move to the web once one or more in-person visits have been completed and the boundaries have been ascertained. This strategy could be implemented regardless of whether a longitudinal element is introduced in the sample design.

Until reliable auxiliary data sources can be integrated into the data collection process, a first in-person visit to each building may still be useful, especially if it is decided that the first interview in a longitudinal series will be completed in person. During that visit, interviewers should follow a protocol developed by EIA to determine if the building is a good candidate for a web response. Given the concerns related to the definition of a building, the decision whether to transition a case to the web will likely depend in part on whether the definition seems straightforward—as it would be, for example, for a small, standalone building occupied by one tenant. Resources should be invested in testing ways to communicate the definition of a building through a self-administered format, in anticipation of transitioning more and more complex buildings to web administration in the future.

The best way to collect contact information for a web survey should be explored as part of the methodological research program. Possible options include obtaining the information during the first visit or obtaining it by telephone. Alternatively, information on how to access a web survey can sometimes be included in a hard-copy advance letter.

An important consideration is the possible impact of the web administration on response rates. This question has been receiving a great deal of interest as more surveys are moved to the web. Recent research has found that augmenting multiple postal letters with supportive emails can be a successful strategy that combines the advantages of both modes of contact and which leads to web-only response rates that are comparable to mail response rates (Millar and Dillman, 2011). The postal letters help establish the importance and legitimacy of the study, while the emails reduce the burden of responding by web because they contain both the link to the survey and the necessary log-in information. Regardless of the initial strategy used, for the foreseeable future the CBECS will likely require at least one additional mode of contact to allow follow-up with sample members who do not respond by web and to assure that response rates do not drop below current levels.

There is no question that identifying the best respondent for completing the interview is crucial in the case of the CBECS, but ways of accomplishing this without involving an interviewer should be explored. It is possible that a web option could actually result in more interviews being conducted with qualified respondents. In some cases, for example, it may be easier to forward a questionnaire to the right person than to locate him or her in a building and arrange an interview. Furthermore, in-person interviewers may have an incentive to complete an interview as quickly as possible by settling for a willing respondent rather than pursuing the most appropriate one.

If a questionnaire was available on the web, it would be easier for several respondents to collaborate, each completing the sections that he or she is most knowledgeable about. A web option could also result in more complete data because it would give respondents the option to pause the survey while they obtain information to answer questions they are not sure about and then resume the survey later. Naturally, if respondents pause while responding to the survey or forward it to someone else, there is a risk that the questionnaire will not be completed at a later time, so an extensive follow-up effort is likely to be necessary. However, when a topic is too technical for many respondents, such as is the case with the CBECS, this kind of follow-up could make a significant difference in data quality.

Currently, the face-to-face CBECS interviews rely heavily on hard-copy show cards. However, transferring the show cards to the web, especially following a unified-mode construction approach, will have significant methodological advantages. In a web survey, show cards are effectively built into the questionnaire and placed directly into the path respondents follow as they go through the survey. Thus respondents can focus on the questionnaire, as opposed to having one eye on the interviewer and one eye on the show cards.

The use of hard-copy show cards also raises the concern of possible biases introduced by the order in which the response options are listed, especially because many of the show cards contain a large number of answer options, which can make it difficult for respondents to focus equally on all of them. For example, the show card listing the answer options for the primary activity in the building contains 16 items, and a respondent for a building with multiple activities may be tempted to select the first one that is applicable as the "primary" activity instead of carefully reviewing the entire list. A web questionnaire would make it easier to restructure these questions into layered sets of items with fewer answer options, or to reduce the possibility of primacy effects by using various innovative methods, such as the animated presentation of response choices or an eye-catching emphasis on the end of the list.

Testing will be required to determine the best way to ask questions that EIA has identified as being challenging because of their technical nature. For example, different approaches to obtaining the square footage information from respondents can be tested in split-sample experiments. Deconstructing this kind of an item into a series of questions would introduce complex skip patterns, but the changes would be easy to implement on the web without increasing the cognitive burden on respondents. Furthermore, a web-based survey could integrate various aids and tools for respondents, such as definitions or diagrams that can pop up if a respondent requests help or seems to be having trouble with a question. Interviewer debriefings would be particularly useful in pinpointing specific questions that could benefit from a different approach.

As is always the case with self-administered surveys, it will be valuable to provide respondents with an email address and toll-free telephone number they can use if they have questions. The staff members who respond to such inquiries should be able to provide assistance related to the technical topics in the questionnaire and to answer questions specific to the web administration.

Finally, when evaluating the implications of transitioning to a mixed-mode administration, it will be important to consider the options available for collecting the utility bills that are currently collected during the interview. Some respondents may be able to upload electronic copies of their bills through the questionnaire website, and this option should be explored. Asking respondents to mail a copy of their utility bills would likely require extensive follow-up, but it is unclear whether this approach would still be more cost effective than multiple in-person visits to a building. Another possible solution is to increase reliance on supplier data, as discussed later.

> **Recommendation CBECS-3:** Informed by a methodological research program, EIA should begin developing procedures for a multimode approach and should begin moving some of the CBECS data collection to the web.

Revised Editing Procedures for the CBECS

CBECS data undergo a complex series of edits to locate and correct errors in the recorded responses in order to increase accuracy. Data errors are inevitable even in surveys on straightforward topics, but they are a particular problem in surveys like the CBECS because of the technical nature and difficulty of the questions. According to EIA, the items in the CBECS that require the most editing are the questions about building equipment.

Postsurvey editing is particularly resource intensive, and it is a major contributor to the time required to prepare the data for public release. The editing standards and, in particular, the optimal extent of the data editing have been the subject of debate among survey researchers and analysts, with some arguing that federal government surveys tend to suffer from "overediting" (Weisberg, 2005).

For a typical CBECS, the editing process consists of CAPI edits that are built into the computer-assisted interview software and occur in real time, while the respondent is completing the survey, and postsurvey edits that occur after the interview is completed. The CAPI edits include

- Warnings about possible inconsistencies in the reporting (for example, if a respondent reports that a building has six floors, and answers "No" to the question about whether there are any elevators, a question pops up that asks the respondent to verify that the answers recorded to both of these questions are correct). Interviewers can revise the responses or "suppress" the warning if the answers recorded were correct.
- Checks for blatant errors (for example, if the building activities reported do not add up to 100 percent, an error message pops up pointing this out). These types of error messages cannot be sup-

pressed. The responses have to be corrected before the interview can proceed.

- Automatic recodings, including responses derived from answers to previous questions, intended to minimize response error and burden (for example, a library building on a college campus is considered a public assembly for the purposes of the CBECS, and it is recoded as such if the respondent describes it as an education building). These types of edits happen automatically and do not require an action from the respondent or interviewer.

Because the CAPI edits are built into the interview program, they are not as resource- or time-intensive as the postsurvey editing steps. While some of the postsurvey checks are automated, the editing is almost always based on manual review, and sometimes requires a callback to the building. The types of edits that take place after the interview include

- Editing for critical items that cannot be wrong, missing, or machine imputed. These items include, for example, square footage; number of workers; year constructed; operating hours; energy uses; energy sources; and percent heated, cooled, or lit. With some exceptions, the editing rules require a callback to the building if the answer to a critical item is missing or appears to be incorrect based on plausibility tests run after the interview (for example, if the square foot per worker and square foot per floor were both higher than the 95th percentile of those ratios from the previous CBECS). Because a callback to the building is often required, the timing of these edits is important and was a concern in the 2007 CBECS when the data collection contractor failed to perform these tasks once the interviews were completed.
- Miscellaneous edits (for example, recoding of "other, specify" answers; verification that buildings coded as enclosed malls are indeed enclosed malls; checking possible inconsistent responses such as a report about a building that has a high-density computer room but no servers, or a government-owned building that does not have any government entities occupying space in the building). Each of the cases in this category is reviewed individually.
- Comprehensive edits of the final CAPI, critical or miscellaneous edit failures. After the data are delivered by the contractor, EIA manually reviews every case that met one or more of the following

> criteria: failed selected CAPI edits (with some modifications) that were not resolved during the interview, failed a critical edit, failed a miscellaneous edit, or failed an edit specific to the consumption and expenditure data.

EIA also manually reviews cases that contain an interviewer remark or responses to "open-ended" questions as well as other questions that accept verbatim responses instead of a precoded answer category. This includes, for example, other building activities; other types of heating equipment; other types of separate computer areas; and the explanation from a respondent who reports only a boiler, only a furnace, or only district heat as the heating equipment but reports more than one energy source.

Although it is not, strictly speaking, an editing procedure, imputation, or the filling in of missing values that still remain after the editing, is another important step that takes place after the survey is completed. This is done using a hot deck imputation method and is limited to the items that are considered important enough that they cannot be left blank. The most frequently imputed CBECS items include square footage and year of construction. For the 2003 data, the imputation rates for these two items were both around 10 percent.

EIA has been planning to revise the CBECS editing procedures because it believes that these procedures could be streamlined and improved. The panel agrees that an evaluation of the editing procedures is necessary to determine whether the data are "overedited" and whether changes could be made to accelerate the processing of the data. As discussed in the previous section, emphasis should be placed on a shift to editing procedures that can be implemented in a self-administered setting to facilitate the integration of a web data collection mode. This could include revising the warning and error messages that pop up during the interview to make them suitable for self-administration and evaluating whether more of the CAPI edits that currently involve warnings could be handled automatically. The manual edits should also be reviewed with an eye toward simplifying the process in order to facilitate faster processing of the data. The review should identify any editing rules that are particularly resource intensive but that contribute relatively little to increasing data quality overall, for example, because of the small number of cases involved. In addition, it may be possible to transfer some of the postsurvey editing that is currently performed manually to automatic processing if a review indicates that there are standard decisions that tend to be applied to resolving some of the problems. Finally, to the extent some of the edits will continue to require a callback to the building, a review of the processes in place to assure that these are performed promptly by the data collection contractor may be necessary.

To determine which aspects of the editing process can be simplified, automated, or eliminated, EIA should conduct an analysis to determine the impact of the current editing procedures on data quality. Comparing data from the previous CBECS from before and after the editing process should be useful in identifying specific editing rules that may not improve the data sufficiently to justify the time and resources invested into pursuing the revisions. In some cases, the conclusion may be that a certain amount of error is an acceptable consequence of the technical nature of the survey and that editing does not eliminate this particular problem.

In addition to a possible reduction in the time required to process the data, revised editing procedures could have the additional benefit of substantial reducing the staff time that is currently dedicated to this process. If some of the staff time could be freed up by eliminating some of the editing steps, EIA would be able to invest these resources into the start-up work required to implement a rotating sample design and web data collection. Staff would also have more time to research and evaluate the costs and benefits of some of the panel's other recommendations, discussed later.

As an alternative to, or in combination with, the above changes, the possibility of releasing a preliminary data set with only a subset of the variables edited, prior to the release of a fully edited data set, also needs to be evaluated. Because many data users have very focused interests and typically use only a subset of the data, releasing some of the variables before the full data set is edited could provide some researchers with earlier access to all of the variables that they need. To the extent that the editing steps for some of the variables are interconnected, care will have to be taken to avoid releasing data that later may have to be further edited based on checks performed on other variables.

Recommendation CBECS-4: EIA should investigate strategies for releasing the CBECS data faster, for example, by revising the editing procedures or by completing the editing for a subset of the variables and releasing these data prior to completing the editing for the full data set.

CBECS DATA GAPS

The current CBECS sample design is best suited for producing descriptive statistics at the national, census region, and census division levels. However, each building in the CBECS is also assigned a climate zone, based on the 30-year average of heating degree-days and cooling degree-days from the National Oceanic and Atmospheric Administration (NOAA) climate division in which the weather station closest to the sampled building is located (see Appendix D). NOAA climate divisions are regions within a state that are climatically homogeneous.

One of the major limitations of the CBECS is a lack of granularity in the data it provides. The sample design—and, in particular, the sample size—limits the analyses that can be conducted on the data, not only geographically but also in terms of the level of complexity. It is especially difficult, using the data that are currently available, to perform multivariate analyses to assess the impact of policy decisions or to explain changes in energy consumption.

A CBECS sample size of approximately 6,000 cases represents nearly 5 million commercial buildings. While this is small in an absolute sense, the size limitations are exacerbated by other factors, such as the inherent variability among buildings of different sizes and activity types, high relative standard error (RSE) thresholds, and the strict confidentiality protections that are designed to prevent survey data from being linked to a specific building (and thereby tenant), and which limit the data that EIA is able to release even further. In other words, some of the most pressing "data gaps" identified by data users are in fact limitations imposed by factors associated with the sample size rather than shortcomings associated with the content of the questionnaire.

Researchers could benefit greatly from having survey information that is broken up into finer geographical divisions. This need is also emphasized by a recent assessment of state energy data needs conducted by EIA (U.S. Energy Information Administration, 2009). State-level data would be more useful than data divided by census region or census division because of the differences that exist among states in a variety of characteristics, ranging from climate to policy. At a minimum, researchers would like to have more information about a building's geographic location—specifically, the American Society of Heating, Refrigerating and Air-Conditioning Engineers climate zone—as well as state information associated with the summary statistics. The fact that this information is not currently available from the survey severely limits the analyses that can be performed and thus the usefulness of the data.

Perhaps even more important for researchers is information about the principal building activity (which EIA defines as the activity or function occupying the most floor space in a building), given that this plays a major role in understanding building energy consumption. EIA groups principal building activities into categories of activities that have similar patterns of energy consumption (for example, office, health care, lodging, and mercantile and service). Since 1999 the CBECS has been classifying building activities into more than 100 different categories, but small sample sizes and the resulting statistical and confidentiality limitations make it necessary to combine these fine-grained categories into 16 broad categories when the data are published. Specific building types that are of interest to researchers include data centers, laboratories, convention centers, and arenas—and there are probably others as well—and these needs have to be evaluated and prioritized.

One addition to the survey that could help data users better understand trends in the building sector would be to provide oversamples of new construction for use in assessing the degree of compliance with building energy codes. Data on renovations and retrofits would also be of value, particularly when the remodeling has been done to incorporate energy efficient designs, systems, and equipment. Researchers could also use larger samples of those buildings that have received leadership in energy and environmental design (LEED) or Energy Star certification. If these designations could be made available as part of the data released, it would be possible to compare the characteristics and energy use of buildings that have an energy efficiency certification with those that do not. As discussed in the panel's letter report (see Appendix F), an oversample of buildings for which some data are available from sources other than the CBECS interviews would allow researchers to make comparisons between the two data sources and thus achieve a better understanding of possible data quality issues with the survey. In the case of LEED and Energy Star buildings, some of the data submitted as part of the certification process are publicly available. An alternative to oversampling specific populations would be to incorporate some of these populations as a "take all" stratum in the sample, in other words, to include all known instances of these populations in the sample.

Another request from data users related to the sample and population of interest for the CBECS is to include buildings that contribute to energy consumption in the commercial sector but that are currently not in scope due primarily to practical considerations. For example, an argument has been made for resuming data collection in buildings that are less than 1,000 square feet because even though they do not represent a large portion of the commercial sector energy consumption, there are a large number of buildings in this category.

Subsequent sections of the report will discuss a variety of data needs that cannot be met by the CBECS because of considerations related to the length of the survey and respondent burden. Given the difficulties related to addressing that challenge, it is especially ironic that, because of the limitations imposed on the data that can be released resulting from a small sample size and the associated quality and confidentiality rules, EIA is unable to at least maximize the usefulness of the data that *have* been collected.

One change that would make it possible to release more of the data that are currently collected, and to meet more data user needs without increasing the length of the questionnaire is to increase the sample size. A substantial increase in sample size would require a substantial increase in funding. However, this is an important factor in EIA's ability to meet data user needs.

> **Recommendation CBECS-5:** As part of its efforts to address the needs of data users, EIA should make it a priority to identify opportunities for increasing the sample size in order to enable the release of more of the CBECS data that are currently being collected.

Another option for making more data available to data users would be to establish a research data center (RDC) that would allow researchers to analyze restricted microdata either at secure sites maintained by a statistical agency or remotely, for example, by submitting computer programs to be run by the RDC. The safeguards that could be implemented through an RDC to protect respondent confidentiality would enable EIA to make accessible some of the data that are currently collected but that are restricted from use due to confidentiality concerns. Typically, data access to the research community as part of an RDC is provided on a project-specific basis. Researchers must submit a proposal describing why analysis of a specified set of restricted data is necessary to answer a particular research question, and the proposal is reviewed and approved only if it is determined that the project would not breach confidentiality. Statistical output that is generated during the use of the RDC is reviewed for disclosure risk before it is released to the researcher.

EIA might either establish its own RDC or collaborate with an organization that already has an RDC, such as the Census Bureau, which has RDCs housed at several universities and research organizations (http://www.census.gov/ces/rdcresearch/rdcnetwork.html [December 2011]), or the Center for Health Statistics, which has research data centers in Hyattsville, Maryland, and Atlanta, Georgia (http://www.cdc.gov/rdc/ [December 2011]). Exploring collaborations with existing RDCs may be the more cost-effective approach, because the start-up costs can be significant. A collaboration with an existing RDC will also require a dedicated budget, which EIA may have to secure through a separate funding request. However, wide dissemination of the data collected should be viewed as central to any statistical agency's mission because it maximizes the data collection's benefits to society (National Research Council, 2009).

Recommendation CBECS-6: EIA should consider establishing a research data center or evaluate the option of using an existing RDC maintained by another organization to provide data users with secure access to CBECS data that are currently not publicly released.

Increasing the sample size and establishing a research data center are recommendations aimed at maximizing the use of the data that can be collected with the current CBECS questionnaire. However, many data needs go beyond what is currently included in the survey. A CBECS interview takes on average between 30 and 45 minutes to administer, and it is a compromise between research needs and respondent burden. Respondent burden is an issue not only in terms of the time required to complete the survey, but also in terms of the difficulty of the questions. Any revisions to the questionnaire would have to take these factors into consideration, in addition to evaluating changing needs due to a changing energy landscape. Although the CBECS cannot—and probably should not—be expected to explain all aspects of energy use by any one building in the sample, it is important to periodically re-evaluate the questionnaire to make sure that any new uses that contribute significantly to a building's energy consumption, or which have the potential to do so in the future, are captured by the survey. Setting practical considerations aside, if significant data gaps exist related to commercial building energy consumption, it would be most useful to the research community and policy makers if the gaps could be filled by EIA rather than by other organizations with even fewer resources or a narrower geographic scope. However, EIA is in the difficult position of having to assess the mission of the CBECS and prioritize among a large number of competing needs.

Below, we list some of the data that researchers would like to see included in the CBECS. The list is by no means exhaustive, but it provides an illustration of the broad range of interests and possibilities for the survey.

- More information about building characteristics (for example, building orientation, wall insulation, length, width, height, floor-to-floor height).
- More information about roofs (for example, roof orientation, area, material composition, reflectivity, and insulation).
- More information about windows (number, window-to-wall ratio, orientation, and overall heat transfer coefficient).
- More information about building systems characteristics (for example, type of fuel used by each of the main building systems, sizes of heating and air-conditioning equipment, and the type of chiller system in the case of buildings using central chiller systems).
- Data about emerging end uses, such as data centers and plug loads, and existing end uses that are growing rapidly and that could become more important in the future (for example, because of climate legislation).
- More information about how much energy is used by specific end uses.
- Information about the efficiency ratings of end use equipment (ratings of fluorescent lighting of different sizes, such as T8 versus T12, energy efficiency ratio, seasonal energy efficiency ratio, annual fuel utilization efficiency ratings, and so on).
- Whole building energy performance data using national consensus building energy performance metrics.
- Information about the available outside parking area, which will be important to understand the potential space available for photovoltaic panels, especially for use in charging electric vehicles.
- More information about building operations (for example, the number of transactions in service and sales buildings, the percent occupancy in offices and hotels, the number of licensed beds in hotels, the number of beds in dormitories).
- Questions about education, awareness, attitude, and behavior.
- Better information about who owns the building and who pays the energy bills.

- Energy consumption data metered in the field and connected to building system characteristics.
- Time-of-use energy consumption data.
- Billing data on a monthly, instead of annual, basis.
- Information about utility rate structures and incentives, along with marginal and average prices at the account level.

Given the wide range of topics that are of interest and would be useful to researchers, it will be important to identify those topics that offer the best balance between value and practicality. That is, the value of new questions to improve end use estimation and provide better information for policy analyses should be weighed against what is feasible and what is within the scope of the existing survey. Some potentially useful data, for example, might require collecting additional technical details about buildings and systems and place an unrealistic burden on both interviewers and respondents. In such cases, however, there may be alternatives for collecting the data. Some data might, for instance, be collected as part of a special study involving energy auditors, and it might be more efficient to collect information on rate structures from the energy utilities rather than from the sample buildings. Such options for collecting additional data will be discussed further in subsequent sections.

For data that are best collected from the building respondent, EIA has to evaluate whether adding topics or expanding on the current ones is in line with the mission of the CBECS; if it is decided that it is, then EIA must prioritize the various needs in the context of the data that are currently collected and that perhaps may have become less useful over the years. The need to manage the length of the questionnaire requires that if any new questions are added, some of the old ones must be dropped.

A particular type of data that may be worth exploring is time-of-use data. Given the expected growth in the number of smart meters, it would seem to make sense to add time-of-use data to the CBECS to take advantage of the wealth and accuracy of information that should become available through this technology. Smart meter data could greatly increase analytic capabilities while reducing the burden on respondents and possibly also reducing interviewing costs. Since a variety of complexities are likely to emerge with respect to the collection, handling, and processing of this type of data, including new confidentiality concerns, it is important to begin now to evaluate the consumption surveys' potential use of this technology. EIA would need to allocate some resources, primarily in the form of staff time, to assess the options and evaluate procedures for collecting this type of data, but the return on the investment would be substantial.

There are a number of potential approaches to collecting this sort of data. As a first step, EIA should contact suppliers that have smart metering in place to evaluate what level of data frequency is available for what customer groups, and to assess options for accessing the data. Options then could be evaluated for collecting smart meter data from a random sample of the suppliers contacted for a follow-up interview, all suppliers who are contacted for a follow-up interview, or from a random sample of the suppliers for buildings for which interviews were also conducted.

Recommendation CBECS-7: EIA should prepare for the more widespread availability of smart meter data in the future by evaluating potential uses of such data, strategies for collecting them, and ways of addressing new confidentiality challenges.

Another topic that deserves attention is the ongoing rapid change in end uses. As new end uses appear and existing end uses become more widespread, they need to be evaluated periodically for potential inclusion in the CBECS. Conversely, end uses that are becoming obsolete should be dropped from the questionnaire. The CBECS staff has been keeping track of these changes, but it may be helpful to develop a formal process for making such updates. Formalizing the process should require only minimal investment of additional staff time, although pretesting any new questions through cognitive interviews or focus groups is always worthwhile, even if it requires additional resources. When evaluating the data needs related to particular end uses, it is important to keep in mind that miscellaneous end uses—and electronics use in particular—do not vary substantially by climate or geography, which means that collecting meaningful data could be accomplished without having to increase the sample size due to concerns about regional variations.

Recommendation CBECS-8: EIA should develop a process for the periodic review of new energy end uses or end uses that are becoming more widespread in the commercial sector and which may need to be included on the CBECS questionnaire. The process should also identify end uses that are becoming obsolete and that can be removed from the questionnaire.

Given the burden that the questionnaire already imposes on respondents, EIA should evaluate creative alternatives for collecting additional data without increasing the overall burden on all CBECS respondents. One possible way to collect additional details or new topics would be to field two versions of the questionnaire—a short form and a long form. EIA would have to evaluate the optimal distribution of content between the two forms as well as the size of the sample for each, taking into consideration both costs and analytic needs. To complete the evaluation, EIA staff may need input from its data collection contractor on the cost of the different options. One option would be to design a short form that collects only basic information from all buildings in the sample and a long form that collects additional details, further customized by building type, from a subset of the sample. This would mean that for the full sample less information would be available than what is available now, but for a smaller sample more information would be available than what is available now. The question is whether the same amount of resources could be allocated differently to meet different types of analytic needs. It is possible that overall this would satisfy fewer data user needs than the current approach does, but the concept is worth considering in order to place data use requests for additional detail in perspective.

Recommendation CBECS-9: To accommodate data user needs for more detailed information, EIA should evaluate the possibility of administering a short-form and a long-form CBECS questionnaire.

REVISIONS TO THE CBECS SAMPLE DESIGN AND DATA COLLECTION PROCEDURES

In this section, we discuss a number of additional changes that could be made to update the CBECS, including revisions aimed at increasing the efficiency of the data collection and at making the survey more useful to researchers and policy makers.

CBECS Sample Design

As discussed above, the CBECS is based on an area probability sample of commercial buildings, supplemented with lists of special buildings. This is a thorough and well-respected method for developing a sampling frame, but it is expensive. The frame is used only once every four years, and updating it for each new round of the survey is costly and imperfect. EIA has tried a number of updating approaches over the years, and it has found the challenges to be substantial, particularly in terms of deduplication. For the 2007 CBECS, difficulties related to updating the sampling frame contributed to the problems that eventually rendered the data unusable.

Given the challenges associated with developing and maintaining a sampling frame of buildings for the CBECS, it is worth taking a second look at the possibility of using establishments instead of buildings as the primary basis of the sampling frame. Commercial lists of establishments are more readily available than lists of buildings, and it may be possible to integrate these lists into the CBECS sample-development process while preserving the survey's focus on buildings as the central unit of interest for the data collection. Because establishment surveys are more common than building surveys, the use of establishments as the basis of the sampling frame would allow the CBECS to benefit from procedures, lists, and data that already exist in connection with these data collections.

As a first step in this process, EIA staff time would be required to evaluate the establishment-based list options that are available for use in sampling frame development and to assess their implications for the efficiency of the data collection. The coverage rates associated with relying on establishment-based lists is an important consideration. If it is necessary to rely on more than one list to achieve adequate coverage rates, the issue of deduplication between the sampling frames should receive careful thought, although deduplicating several lists is likely to be less error prone than prior efforts (especially the 2007 sampling frame development) that involved deduplicating an area-based sample and several lists. Cost and timeliness are two other factors that will also have to be assessed in comparison with the current approach. It is likely that there are lessons to be learned from the experience with the 2007 CBECS sample design, and as EIA continues to evaluate that experience the conclusions could underscore the need for a careful evaluation prior to implementing this change as well as for close monitoring of the implementation once the sampling plan is developed. After a careful evaluation, a staggered launch of the fieldwork may be useful to identify and eliminate problems that could emerge. This could involve releasing a small subset of the cases prior to beginning work on the full sample.

> **Recommendation CBECS-10:** EIA should conduct research to evaluate the advantages and costs associated with using an establishment-based list for the CBECS sampling frame.

CBECS Data Collection Procedures

A major part of the expense of any survey is the cost of following up with respondents to complete the interviews, and this is particularly true for in-person data collection. One way to make the process more efficient would be to take advantage of the fact that in the commercial sector there are many establishments with large numbers of buildings across the country (for example, large companies such as McDonald's or Walmart) and to collect data about these buildings from a centralized location instead of from the individual buildings. This could work especially well if the survey were to be transitioned to an establishment-based frame and multimode data collection, but, even if those changes are not made, EIA staff should carry out a small research project to evaluate the option of contacting the headquarters of major centralized accounts to inquire about collecting as much data as possible from them. The evaluation should determine whether there are centralized accounts that can provide all of the information necessary for the CBECS for all of their buildings in the sample, and assess how much variability can be expected between the accounts in terms of the data available centrally.

> **Recommendation CBECS-11:** To increase CBECS data collection efficiency, EIA should explore the possibility of contacting the headquarters of major centralized accounts and collecting data about all of the buildings in the sample from these centralized sources.

Data from Energy Suppliers

The CBECS includes an energy supplier survey for about half of the CBECS buildings in the sample. The survey is initiated in cases where the energy usage and cost information cannot be obtained through a building interview or if the data obtained through the building interview are flagged as out of the expected range. EIA should research the possibility of working more closely with energy suppliers. Although differences in the ways different energy suppliers keep their records represents a challenge, EIA is already working with a large number of utilities to obtain consumption and cost data for buildings that cannot supply this information. It would be worth evaluating whether the supplier survey could become the first step in the data collection process, with individual building respondents only being asked the remaining questions. By significantly reducing the burden on building respondents and possibly also improving data quality for the consumption and cost data, efficiencies could be gained that could offset the increased effort associated with requesting utility data for more buildings. It is possible that energy suppliers may find it easier to provide a data file with information for all of their customers as opposed to providing information for individual customers that need follow-up because they are in the CBECS sample and were not able to provide the data themselves. If EIA could request data for all of a utility's customers, and then extract the data for the buildings that are in the sample, this could reduce the burden on the utilities sufficiently to facilitate collaboration.

As part of evaluating the costs and benefits of collecting supplier data before the building interviews, it may also be worth assessing possibilities related to using energy supplier records to assist with building a sampling frame. EIA considered this option early on, but identified a number of challenges, beyond the variability in the ways utility records are kept. For example, in the sample of commercial accounts examined for a feasibility study, about a quarter were not buildings, but other structures (French, 2007). However, if reliance on energy suppliers were to be increased for the data collection, relying more on utilities for the sampling frame development as well could become more cost effective. Another consideration is that establishing a framework for closer collaboration with utilities would also be useful in paving the way for greater reliance on smart meter data.

> **Recommendation CBECS-12:** EIA should evaluate whether working more closely with the energy suppliers of commercial buildings could lead to efficiencies in the data collection process.

Auxiliary Data

As face-to-face data collection becomes more expensive, and maintaining high response rates through most traditional modes of interviewing becomes more difficult, the survey research community is becoming increasingly interested in tapping into existing data sources, such as administrative records. The panel understands that EIA has considered the use of administrative records in the past, and the panel also acknowledges that gathering and combining data from a variety of sources can be resource intensive. Nonetheless, it seems inevitable that the CBECS will eventually have to move toward the use of a multimode data collection approach, and the potential of auxiliary data sources should be revisited periodically as the costs and data quality benefits associated with integrating such sources into the data collection evolve. As more records become available online, it is likely to be increasingly more efficient to use these auxiliary data sources.

A successful integration of auxiliary data sources into the energy consumption surveys could have some important advantages. Given the technical nature of the questions asked, for example, there are concerns about respondents' ability to provide accurate answers, which lead in turn to concerns about the accuracy and reliability of the data. The data quality could be improved by finding one or more alternatives to self-reports for some of the information collected, such as square footage and intensity of use. Being able to drop some of these questions from the questionnaire would also free up resources that could be allocated to meeting other data user needs by, for instance, using a larger sample size or collecting data about different topics.

Local government databases, such as county property tax databases, are often available online, and some of them include information on square footage and heat sources, albeit based on definitions that are not always consistent with those used by EIA. Some of these databases also include photographs that can provide useful information on building characteristics. There should be regular evaluation of the availability and accuracy of these databases around the country in order to assess what proportion of cases would have associated records that lend themselves to relatively straightforward integration into the CBECS data. For example, knowing the proportion of counties with property tax records that present no major challenges for use would help determine whether it was cost effective to integrate these sources of data into the survey. It will also be important to evaluate how integrating administrative data could affect the data collection timeline, to assure that adding new sources of data does not adversely affect the time required to process and publish the CBECS data.

The panel's letter report (see Appendix F) contains suggestions for a more comprehensive research plan for evaluating auxiliary sources of data, if funding for this research can be secured. Collaborations with one or more of the Department of Energy's research labs could be one productive way of pursuing this type of research without putting a strain on EIA staff time. At the minimum, EIA staff members should stay abreast of what potential new sources of building data are available and should monitor trends indicating what data may become available online in the near future.

Recommendation CBECS-13: EIA should conduct an ongoing evaluation of administrative records as potential sources for substantive data for the CBECS.

Energy Audits

EIA has considered hiring professional energy auditors to collect building data instead of relying on interviewers, but it has never had sufficient funding to do it. The 2011 CBECS data collection plan included a small-scale test with both auditors and interviewers collecting data from the same buildings in order to evaluate the differences between the two approaches, but the test was put on hold along with the survey.

The panel considers this an important test to complete because it could have a large number of benefits. Professional energy auditors are trained and certified to assess building energy use, and for observable building characteristics the data collected by them could be treated as the "gold standard" for use in calibrating the responses gathered by interviewers and evaluating the quality of data from other sources, such as administrative records. The data collected by auditors would also be useful for evaluating some of the current back-end procedures, such as data editing or the regression model used to identify outliers and to initiate a supplier follow-up survey.

While hiring professional auditors would cost more than hiring interviewers, these costs should be evaluated in the context of how much it would cost to provide additional training to interviewers. The use of auditors could reduce some other costs as well. For example, it is possible that building respondents may be more willing to trust a professional auditor than a survey interviewer because auditors may be perceived as more highly qualified on the topic and possibly more "legitimate." Editing costs could drop if the use of auditors helped alleviate some of the privacy concerns on the part of the building respondents, causing them to be more likely to answer some questions that they would have been reluctant to answer before. Respondents might also be more willing to provide access to equipment and areas of the building to professional auditors that interviewers would find difficult to gain access to. If privacy concerns are reduced, this could also reduce the resources needed to follow up with reluctant respondents. Evaluating these factors and their impact on the costs and benefits of involving energy auditors could be accomplished in the form of a small pilot test, which could be an additional task performed by the data collection contractor.

Recommendation CBECS-14: EIA should test the use of professional energy auditors on a small scale to better understand the costs and benefits related to having experts collect data for a subset of the CBECS sample.

Data Collection Instruments

The CBECS questionnaire should be reviewed on a routine basis to make sure that it is not becoming out of date. EIA keeps track of the content of the data collection instruments, but it is also important to develop a formalized process for carrying out a periodic review. This review should be separate in concept from a major redesign, but could be integrated with the periodic review conducted to assess the relevance of end uses measured by the questionnaire (see Recommendation CBECS-8). The resources required for these activities will depend on the research techniques employed, which should be guided by types of changes that are considered.

The techniques used to evaluate the questions could involve examining the distribution of the responses to each of the questions to identify answer options that no longer reflect the distribution of building characteristics of interest as well as answer categories that may have become obsolete. Cognitive interviews could be conducted to better understand how respondents relate to the question and answer options and to evaluate whether updates are needed. The cognitive interviews should also examine the strategies used by respondents to come up with answers to questions that are suspected of being especially difficult to answer and, thus, of possibly producing less accurate data.

When a questionnaire is updated, it is important to take into account the continuity of the time series. The analysis conducted will have to evaluate the implications of any changes in the context of the risks posed by the modification, while at the same time performing the evaluation with an eye on the future and on what content is expected to be most relevant going forward.

As mentioned, one important aspect of the questionnaire review process is the identification of items that have become outdated or are no longer useful enough to warrant inclusion in the CBECS. Some questions must be removed if new questions are to be added.

Recommendation CBECS-15: EIA should invest in periodic reviews of the CBECS questionnaire content and wording. This should be understood as a routine updating of the instruments, separate from the concept of a major redesign of the survey.

Greater EIA Involvement in the Data Collection Process

While the panel's understanding is that EIA staff members participate in all interviewer trainings, more active involvement may be necessary to share the study's goals and communicate how the quality of the data determines its usefulness. Furthermore, EIA staff members are best qualified to conduct training on any topics and concepts that are complicated due to a long institutional history, such as the definitions of a building and of a qualified respondent. Increased involvement could be accomplished with a minimal investment of time, and the benefits should be noticeable.

Additional resources should be invested in monitoring and analyzing the characteristics of the field operations and in identifying opportunities for increased efficiency. EIA should review any information available from the data collection contractor regarding the amount of time spent on cases of various types (such as buildings with different characteristics or respondents with different backgrounds). If the case level contact history is not recorded in sufficient detail, efforts should be made to make sure this information is captured in the future. EIA should also ask the data collection contractor to schedule debriefings with the interviewers soon after the beginning of the field period, and EIA staff should attend these debriefings to better understand how interviewers spend their time in the field and what types of cases are presenting the biggest challenges and why. A detailed analysis of the time allocation should reveal whether there are subsets of cases that require a disproportionately large amount of time to complete and whether the effort is justified in the context of data needs and statistical techniques available to compensate for missing information.

Interviewers can also be a good source of background and contextual information on questions that are difficult to administer, especially on whether particular questions are causing concerns about data quality. Discussions with the interviewers could be the first step in a close examination of a questionnaire that has evolved with a face-to-face administration in mind and that may need revising or simplifying to accommodate different future modes of data collection.

EIA should also work closely with the data collection contractor to review the procedures used to select the best respondent for the building interviews and to identify opportunities to streamline this process. Again, debriefings with interviewers can provide invaluable feedback that can help fine-tune the process and contribute to the development of new interviewer protocols. More efficient procedures for identifying a qualified respondent can not only reduce costs but also address some of the concerns related to the technical nature of the questions. Qualitative feedback from the interviewers can then be further examined with an analysis of the quantitative responses by respondent type in order to identify possible differences in data quality. In other words, it is possible that most of the questions are not "too technical" if posed to the right respondent.

> **Recommendation CBECS-16:** Interviewer debriefings should become an integral part of the CBECS data collection process in order to identify problems with the questionnaires and procedures and to serve as a source of ideas for increased efficiencies.

6

Redesigning the Residential Energy Consumption Survey

This section discusses the panel's recommendations for redesigning the Residential Energy Consumption Survey (RECS). Although the RECS is methodologically a more straightforward survey than the Commercial Buildings Energy Consumption Survey (CBECS), the main goals and many of the survey's main uses are comparable to those of the CBECS. Thus the priority areas identified by the panel are also the same as for the CBECS: (1) the timeliness and frequency of the survey and (2) data gaps. We will first discuss these two topics and the panel's recommendations for addressing them, then provide recommendations for more incremental revisions for updating the RECS. Given that there is significant overlap between the recommendations for the two surveys, we will not go into the same level of detail here concerning the background of the survey or overall cost implications as we did in Chapter 5, where much of the information relevant to the RECS was first discussed in the context of the CBECS. However, those recommendations that are specific to the RECS and that appear for the first time in this chapter will be given a more thorough discussion.

TIMELINESS AND FREQUENCY OF THE RECS DATA

By law the RECS should be collected every three years, but a lack of funding has limited the survey to a quadrennial schedule since 1993. The delay between the time the data are collected and the time they are released to the public is a concern for the RECS, as it is for the CBECS, although the RECS has not experienced the same extraordinary challenges in recent years as the CBECS has, and the U.S. Energy Information Administration (EIA) was able to release the data from the 2009 RECS much faster than in previous years, in part by releasing a preliminary public-use microdata file—with a limited number of variables—in October 2011.

Rotating Panel Design for the RECS

Just as with the CBECS, a rotating sample design would be a suitable option for addressing data users' need for quicker and more frequent access to data from the RECS. Rotating design options were discussed in detail in Chapter 5. Briefly, the new approach would involve collecting the data over a four-year period instead of once every four years. This would enable annual data releases, with the new data released each year combining with data from the previous three years to add up to the same total sample size as is now provided by RECS once every four years. Ideally, a subset of the sample would be followed over time to create a longitudinal data set that would lead to an improved capability to analyze trends and changes. Specific options for a sample design are discussed in Chapter 5 in connection with the CBECS, and the approach most suitable for the RECS would have to be evaluated. An overall assessment of the costs and benefits of introducing a rotating sample design should receive priority as part of EIA's near future plans for the RECS as well.

> **Recommendation RECS-1:** EIA should evaluate the usefulness of implementing a rotating sample design for the RECS to improve the timeliness of the data.

> **Recommendation RECS-2:** EIA should consider integrating a longitudinal element into the RECS sample design to obtain better estimates of change.

RECS Multimode Data Collection

As with the CBECS, the RECS should be gradually transitioned to a multimode data collection effort, in part to enable faster processing and release times. The panel acknowledges that many of the existing procedures will have to be modified and simplified in order to make it possible to have a more rapid turnaround. The current set of editing steps is one example of a procedure that will need to be revised; this is discussed further in the next section. Despite the effort required to move to multimode collection, the panel believes the change is necessary because of the high and steadily increasing costs of data collections that are based exclusively on face-to-face interviewing and the particularly high costs associated with nonresponse follow-up in the RECS.

The transition to multimode data collection will be just as complex as with the CBECS, but with an additional complication: In the RECS, square footage is measured by field interviewers (although respondents are also asked), and EIA is especially concerned about shifting the collection of this data to respondent self-reports. The panel understands the importance of square footage data in statistical models, but it remains unconvinced that it would not be feasible to collect adequate information from the respondents. The CBECS already relies on self-reported data on square footage, and with thorough testing it should be possible to develop a series of questions that can produce reliable information on square footage through a web survey as well. As in the case of the CBECS, square footage data is one of the items that require the most editing in the RECS, so it is not clear that the presence of interviewers sufficiently mitigates the challenges related to obtaining this type of information as part of a survey. EIA will need to conduct methodological research to determine the best methods and likely trade-offs when collecting self-reported data on square footage.

In terms of operational considerations, sending an advance letter with log-in instructions to the sampled address may be a feasible first step in the data collection, especially given that EIA already sends a postcard and a letter to the address prior to the interviewer's visit. In most cases, this letter would eventually have to be followed up with another data collection mode for nonrespondents, but the number of surveys completed over the web is expected to increase over time. The multimode approach will be discussed further in later sections of this chapter, in combination with discussions of the possibility of implementing an address-based sampling (ABS) approach for the RECS.

> **Recommendation RECS-3:** Informed by a methodological research program, EIA should begin developing procedures for a multimode approach and should begin moving some of the RECS data collection to the web.

Revised Editing Procedures for the RECS

As with the CBECS, the RECS relies on a series of complex edits to increase the accuracy of the data collected, and these edits contribute to the length of time required to release the data. The most edited items in the RECS tend to be the information about square footage and the main source of space heating. Supplier data also require significant editing.

The most common edits built into the interview are signals that warn about inconsistencies in the reporting. Examples include the lack of a water heater, a heat pump reported as used for cooling, but not for heating, and a single-family detached house with fewer than 2 bedrooms or more than 10.

The types of edits that take place after the interview include

- Critical edits for items that cannot be wrong, missing, or machine-imputed.
- Reviewing "other, specify" responses for possible recoding.
- Reviewing selected interviewer comments for possible recoding.
- Identifying and resolving potentially out-of-scope cases.

The data collection contractor for the RECS typically performs most of the editing work. After the contractor completes the above editing steps, EIA performs additional editing, focusing on the following:

- Key items (for example, cooling not used in home, but more than zero rooms cooled).
- Problems with square footage data, such as:
  - o An unusually small single-family detached house (less than 800 square feet overall, less than 294 square feet per person, or less than 187 square feet per room).
  - o An unusually large single-family detached house (more than 6,045 square feet).
  - o The respondent-estimated square footage is more than twice or less than half of calculated square footage.
- Comparing household and rental agent responses.

EIA reported that it is also resource intensive to edit the supplier survey data. In 2009, in addition to errors that were in the data received from the suppliers, the data collection contractor introduced a number of additional errors. These included (1) scanning errors; (2) keying errors, including errors interpreting the data received from the suppliers; (3) partial expenditure data; (4) a floating decimal point problem in the data entry program; and (5) dates entered incorrectly.

To address problems associated with missing data and to resolve some misreported items, EIA uses a hot deck imputation method, similar to the technique used by the CBECS. The items that most often need imputing include respondent-estimated square footage, measured square footage, the age of the home, and household income.

While the editing procedures developed by EIA undoubtedly enhance the quality of the data, revisions to the RECS editing procedures will have to sacrifice some of these rules and simplify the overall process, to facilitate faster processing of the data. As with the CBECS, emphasis should be placed on a shift to editing procedures that can be implemented in a self-administered setting, and an analysis should be conducted to evaluate whether there are editing rules that do not improve the data quality sufficiently to justify the resources invested and which could possibly be eliminated. In addition, it may be necessary to carry out a review of the processes in place to assure that the editing steps performed by the contractor are completed in a timely manner.

The possibility of releasing a subset of the data before the edits are completed for all of the variables should also be evaluated as an alternative to the above changes or as something done in combination with them. Given that many data users analyze only a subset of the data, releasing some of the variables before the full data set is edited could allow some researchers earlier access to all of the variables they need.

Recommendation RECS-4: EIA should investigate strategies for releasing the RECS data faster, for example, by revising the editing procedures altogether or by completing the editing for a subset of the variables and releasing these data prior to completing the editing for the full data set.

RECS DATA GAPS

The design of the RECS makes it most suitable for producing summary statistics for larger geographical areas (the entire country, census regions, and census divisions) and for selected states. Between 1993 and 2005, the

RECS released data for the four most-populated states; and in 2009, it released data for 16 states.

As is the case with the CBECS, the lack of granularity in the data available from the survey is one of the concerns raised most often by data uses, and, as is the case with the CBECS, which data can be released is largely dependent on the sample size. While the recent significant increase in the sample size (from 4,382 cases in 2005 to 12,083 in 2009) greatly improved EIA's ability to provide data for subnational geographies, the need remains to have data available for each state in order to assess, for example, the effect of state-level programs on energy consumption. Insufficient sample sizes also hinder the ability of researchers to conduct multivariate analysis, making it difficult, for example, to evaluate changes resulting from the adoption of new technologies and energy efficiency measures.

> **Recommendation RECS-5:** As part of its efforts to address the needs of data users, EIA should make it a priority to identify opportunities for increasing the sample size in order to enable the release of more of the RECS data that are currently being collected.

As discussed in Chapter 5, another way to make more of the data collected available to researchers is the use of a research data center (RDC).

> **Recommendation RECS-6:** EIA should consider establishing a research data center or evaluate the option of using an existing RDC maintained by another organization to provide data users with secure access to RECS data that are currently not publicly released.

Like the CBECS, the RECS is in the difficult position of being a lengthy, high-burden survey whose design must take into account and prioritize a wide range of requests to collect data on additional topics and in more detail. The list below provides examples of the types of data that researchers would like to see included in the RECS. The list is by no means exhaustive, but it provides an illustration of the broad range of interests and possibilities for the survey.

- More information about building characteristics (for example, building orientation, wall insulation, length, width, height, floor-to-floor height, and location of ductwork).

- More information about roofs (for example, roof orientation, area, material composition, reflectivity, insulation).
- More information about windows (for example, number, window-to-wall ratio, orientation, overall heat transfer coefficient).
- More information about building systems characteristics (for example, location of equipment in an attic versus a basement or crawlspace, more detail on space heating equipment, and additional categories of water heating equipment).
- More detail on electric appliances and miscellaneous electric loads, such as data on saturation (devices per home) and duty cycles (or time spent in an active state), for a variety of individual technologies, including televisions, computers, set-top boxes, game consoles, monitors, audio-visual equipment, imaging equipment, and chargers for portable devices.
- Information about the efficiency ratings of end use equipment (for example, energy efficiency ratio, seasonal energy efficiency ratio, annual fuel utilization efficiency ratings).
- More detail about vehicle use and fuel use.
- Better data on multiunit buildings.
- Better measures of energy conservation behavior and indicators of the degree of success of energy efficiency programs.
- Thermostat set-point and set-back information based on direct observation of the thermostats instead of self-report.
- Energy consumption data metered in the field and connected to building system characteristics.
- Time-of-use energy consumption data.
- Monthly billing data instead of annual.
- Information about utility rate structures and incentives, along with marginal and average prices at the household level.
- Energy arrearages and bad debt related to the household's energy suppliers.
- Better household income measures.

It is obviously impossible for the RECS to meet everyone's needs completely, and, as is the case with the CBECS, EIA has to weigh the potential of a new measure to advance research and inform policy against the practical limitations and scope of the survey. If any new questions are added, every effort should be made to identify questions that can be dropped. Data needs that involve collecting additional technical details about buildings and systems would probably impose an unrealistic burden on both interviewers and respondents, but they could possibly be accommodated as part of a special study involving energy auditors. Some types of information, such as rate structures, would perhaps be more efficient to collect from the energy utilities than from the households.

One topic area that the panel would like to see pursued as part of the RECS is the integration of smart meter data. Given that this topic is also relevant to the CBECS, a joint effort could be mounted to evaluate the consumption surveys' potential use of this technology.

> **Recommendation RECS-7:** EIA should prepare for the more widespread availability of smart meter data in the future by evaluating potential uses of such data, strategies for collecting them, and ways of addressing new confidentiality challenges.

Another topic that deserves careful consideration is the possibility of collecting household transportation data as part of the RECS. As mentioned previously, EIA's consumption survey portfolio has not included a transportation survey since the 1990s, when the Residential Transportation Energy Consumption Survey was discontinued. In 2009, a small section of transportation-related questions was added to the RECS; the decision to do so was based in part on the argument that household vehicles are a part of household energy consumption. With the growing presence of electric vehicles, which are generally charged at home, this discussion becomes especially relevant. Given that electric vehicles could, by the next round of the RECS data collection, become an end use with a significant effect on residential energy consumption, the panel encourages EIA to be ready to collect more information about this topic.

> **Recommendation RECS-8:** In anticipation of the spread of electric vehicles, EIA should prepare for the RECS to collect more information about these, especially more detail about the capacity to charge electric vehicles.

EIA should also develop a process for evaluating changes in end uses more generally in order to determine if the RECS needs to be updated with new questions and whether any of the existing questions can be dropped. As noted previously, miscellaneous end uses—and electronics use, in particular—do not vary substantially by climate or geography, which implies that collecting meaningful data can be accomplished without having to increase the sample size.

> **Recommendation RECS-9:** EIA should develop a process for the periodic review of new energy end uses or end uses that are becoming more widespread in the residential sector and which may need to be included on the RECS questionnaire. The process should also identify end uses that are becoming obsolete and that can be removed from the questionnaire.

In considering updates to the RECS questionnaire, EIA finds itself in a challenging position because of the already very high burden imposed on respondents by the RECS questionnaire. The panel acknowledges that it is easier to add new questions than to truly revise a questionnaire, but nevertheless it encourages EIA to consider alternative ways of accommodating new topics instead of increasing the burden of the RECS instrument on respondents. For example, EIA could play a leading role in the collection of data on specialized topics in the context of a variety of different arrangements that do not strictly speaking involve the RECS.

One topic that would lend itself to such a specialized study is the energy consumption of multiunit residential buildings. These types of buildings and their energy use are not covered adequately by the RECS, which gathers information only on individual apartments in multiunit buildings. A number of researchers would be interested in whole building energy consumption and additional details about these buildings, such as the number of floors above ground, the number of units, the percent of the building devoted to common space, whether there are central or in-unit laundry facilities, the configuration of utility meters, whether the tenant or owner pays for the utilities, and whether the rent paid is "affordable" or "market-rate." Collecting this type of information would be particularly challenging in the case of buildings that are part residential and part commercial. EIA will have to decide if and how to collect the commercial component of the consumption in these types of buildings and perhaps develop a modeling approach to partition the residential and commercial components. Various organizations have expressed interest in collaborations to fill the data gaps on this topic, and EIA should consider taking on a more active role in addressing this need.

Recommendation RECS-10: EIA should consider implementing a "whole building" supplement to the RECS to address the data gap related to multiunit residential buildings.

A possible option, focused on accommodating questions that the RECS is ideally suited to collect, would be to redesign the data collection approach in a way that involves a longer and a shorter questionnaire. EIA would have to evaluate the optimal distribution of content between the two forms and the size of the sample for each, taking into consideration both costs and analysis needs. One possibility is to design a short form that collects only basic information from all buildings in the sample, and a long form that collects additional details, further customized by building type, from a subset of the sample.

Recommendation RECS-11: To accommodate data user needs for more detailed information, EIA should evaluate the possibility of administering a short-form and a long-form RECS questionnaire.

REVISIONS TO THE RECS SAMPLE DESIGN AND DATA COLLECTION PROCEDURES

In this section, we discuss additional potential changes to update the RECS, including revisions focused on increasing the efficiency of the data collection, and making the survey more useful to researchers and policy makers.

RECS Sample Design

The RECS is based on a stratified, multistage area probability sample of occupied housing units. This is a resource-intensive method, and the sampling frame requires updating every four years. As with the CBECS, EIA has tried a variety of updating methods over the years but difficulties remain.

In recent years, concerns about declining response rates of surveys in general and about the declining coverage rates of random-digit dial surveys due to the spread of cell-phone-only households in particular have led to significant interest in address-based sampling (ABS) methods. The use of the U.S. Postal Service (USPS) delivery sequence file (DSF) as a source for an ABS frame appears to be a particularly promising approach for conducting household surveys of the general public (Link et al., 2008).[1]

Using the DSF to perform address-based sampling provides a very high coverage rate of residential households, and the lists can offer a variety of auxiliary information matched to the addresses, ranging from geocodes to landline telephone numbers and even Spanish-surname indicators for the head of household, all of which can help researchers enhance a survey design and develop a data collection plan (American Association for Public Opinion Research Cell Phone Task Force, 2010).

Researchers have also started evaluating the potential use of DSF and postal mail to collect data on the web (Messer and Dillman, 2011; Smyth et al., 2010). While relying solely on the web is far from realistic for a survey that aims to be representative of the U.S. population, studies have shown that web responses can be obtained by mailing initial requests via postal mail, which is then followed by another data collection mode (Messer and Dillman, 2011; Smyth et al., 2010).

EIA experimented with the use of the DSF for portions of the 2009 sample, and it should further evaluate the possibility of using ABS as the primary method for constructing the sampling frame for the RECS. This change should be specifically assessed in the context of the need to gradually transition some of the interviews to the web and the opportunities represented by a combination of these two methods. EIA should review the lessons learned from the 2007 CBECS and work with the new data collection contractor to better understand the cost implications of this approach to sampling.

> **Recommendation RECS-12:** Research should be conducted to evaluate the possibility of replacing the RECS area probability sample with address-based sampling, using an address list developed by commercial vendors based on the U.S. Postal Service delivery sequence file.

---

[1] The USPS licenses a variety of address products and services to companies that repackage these and then license them in turn for commercial and research applications. The services include access to a weekly snapshot of the USPS computerized delivery sequence (CDS) file, which contains all delivery point addresses serviced by the postal service, with the exception of general delivery. Although researchers do not have direct access to the USPS delivery sequence file, for the sake of simplicity we will refer to lists based on this file as DSF.

RECS Data Collection Procedures

Auxiliary Data

Administrative records for buildings in the residential sector are becoming more readily available online. Although these records are kept in a large variety of formats, and it is not always possible to know how accurate they are, EIA should continue to assess the potential of such data as an alternative to an exclusively survey-based data collection that could help address various challenges related to interview costs and nonresponse. The use of alternatives to self-reported data is especially worth pursuing for square-footage data and for the development of intensity-of-use measures, in other words for measures of the energy consumed per units of interest, such as energy consumed per square foot.

Recommendation RECS-13: EIA should conduct an ongoing evaluation of administrative records as potential sources for substantive data for the RECS, especially for square footage data.

Data from Energy Suppliers

EIA should evaluate whether working more closely with energy suppliers to obtain RECS data would be feasible to increase data collection efficiency and reduce the burden on respondents in the RECS sample. Specifically, collecting energy consumption and cost data from energy suppliers prior to the housing unit interview should be explored. Establishing closer collaborations with utilities may also be necessary to begin developing procedures for obtaining smart meter data.

Recommendation RECS-14: EIA should evaluate whether working more closely with the energy suppliers could lead to efficiencies in the data collection process.

Energy Audits

The chapter on CBECS describes a number of potential advantages of involving professional energy auditors in the data collection. A test to evaluate the use of auditors should also be conducted for the RECS because the relative advantages may be different in the case of a household population.

Recommendation RECS-15: EIA should test the use of professional energy auditors on a small scale to better understand the costs and benefits related to having experts collect data for a subset of the RECS sample.

Data Collection Instruments

The RECS questionnaire should be reviewed on a routine basis to assure that it is not becoming out of date. Although EIA has revised some of the RECS questions recently, it is important to develop a formalized process for a periodic review. This review should be separate in concept from a major redesign, and it should involve a variety of techniques to evaluate the questions, as appropriate.

Integrating cognitive interviews into the questionnaire-updating process may be particularly important to better understand how respondents relate to the questions and the answer options and to evaluate whether updates are needed. The cognitive interviews should also examine the strategies used by respondents to come up with answers to those questions that are suspected of being especially difficult to answer and of possibly producing less accurate data.

For example, the RECS questions about televisions in the home have recently been revised, but the question about the type of the television still lists "standard tube" as the first answer option on the show card. As tube televisions become less common, presumably fewer people will be familiar with this terminology. This, combined with the use of the word "standard" and the placement of the answer option first on the list could lead to overreporting of these types of devices. The survey research literature indicates that respondents attribute more meaning to the range and ordering of answer options provided than researchers would prefer (Schwarz, 1996), and testing may reveal that a reordering of the answer options is needed.

Another example of a question in the RECS that may need updating is the one about wireless access in the home. This question seems to assume that a yes answer implies the presence of a wireless router. However, with the spread of smart phone devices, this question could have a different meaning for some of the respondents.

An important trend in energy consumption is the increase in the number of appliances, tools, and electronics that use energy in a typical home. The current RECS questions simply give respondents a few examples of rechargeable devices and ask for the number used in the home. This approach could become increasingly inadequate as the number of rechargeable devices increases, making it more difficult for respondents to perform a quick but comprehensive mental calculation. An alternative question format, such as a list of devices that prompts respondents to answer yes or no to each item and then provide the number applicable to that category, may have to be considered (along with an "other, specify" answer category).

The panel cannot recommend specific wording changes without further testing. The examples above are instead intended to illustrate the types of issues that should be considered. The panel also acknowledges that the continuity of the time series is an important consideration, but it believes that questions should still be evaluated with an eye on the future and considering which content is expected to be most relevant going forward.

As discussed, one important aspect of the questionnaire review process is identifying items that have become outdated or are no longer useful enough to warrant inclusion in the RECS. Removing some questions is imperative if new questions are to be added.

> **Recommendation RECS-16:** EIA should invest in periodic reviews of the RECS questionnaire content and wording. This should be understood as a routine updating of the instruments, separate from the concept of a major redesign of the survey.

Greater EIA Involvement in the Data Collection Process

The panel recommends greater EIA involvement in the tasks that are performed by data collection contractors. As long as these highly technical data continue to be collected by interviewers, interviewer training deserves particular attention. EIA should also work with the data collection contractor to schedule periodic interviewer debriefings in order to obtain feedback about the fieldwork.

> **Recommendation RECS-17:** Interviewer debriefings should become an integral part of the RECS data collection process in order to identify problems with the questionnaires and procedures and to serve as a source of ideas for increased efficiencies.

7

A State-of-the-Art Energy Consumption Data Collection Program

The panel's recommendations in this report are based on a careful balance between what the data needs are and what is most realistic to implement in the short term. This chapter offers some additional ideas that may be less feasible in the short term or are perhaps less critical but that could help U.S. Energy Information Administration (EIA) as it plans for the future.

A COMPREHENSIVE SURVEY OF ALL ENERGY-CONSUMING PREMISES

Due to practical considerations, the three current energy consumption surveys, the Commercial Buildings Energy Consumption Survey (CBECS), the Residential Energy Consumption Survey (RECS), and the Manufacturing Energy Consumption Survey impose somewhat arbitrary divisions between the energy-consuming sectors, particularly as they apply to premises on the boundaries between the sectors. Furthermore, some energy-consuming premises are treated as out of scope by all three of the existing surveys. This is not unusual or inconsistent with other federal surveys. Still, it is possible to envision a comprehensive design for collecting building energy consumption based on a sampling frame that includes all buildings, regardless of type. Buildings would be sampled from this frame and would then be classified into categories such as commercial, residential, or industrial, instead of being divided into categories first and then sampled.

A variation on the area sampling approach that is becoming more feasible with today's technology is to develop a grid with small enough cell sizes to capture buildings and then lay that grid over a map of the United States. A computer program can perform clustering and determine whether a given cell contains any buildings by, for example, using satellite mapping to identify buildings. The panel anticipates that most buildings could be accurately classified by EIA without the need to rely on field visits. Field procedures would still have to be developed for handling cases that cannot be classified or that are misclassified, but it would be possible to implement an approach of this type in a significantly more cost-effective way than current designs that rely primarily on field listing.

Drawing a random sample of buildings based on satellite maps may be another way to approach a design of this type. For example, the street-view function of programs such as Google maps could help in assigning many of the selected buildings to an energy-consuming sector, although it would not work in all cases, and there are some areas for which the street-view function is not available. The technology and the availability of features are constantly changing, and this should also be an important consideration.

INTERACTIVE ONLINE TOOLS

Few government surveys have embraced the use of interactive online tools as a means to build awareness about a data collection and to encourage survey participation. As it becomes more and more difficult to maintain high response rates, agencies are being forced to invest more and more resources into nonrespondent follow-up, but most agencies rely on a limited number of techniques to gain respondent cooperation. In the case of a survey such as the RECS, in particular, innovative strategies involving interactive online tools may be able to engage sample members who are interested in learning about their homes' energy consumption. Fostering a sense of involvement and reciprocity around the data collection programs will change the dynamics and could help EIA maintain high-response rates at least among a specific segment of the population.

There are a number of possibilities for integrating interactive online tools. A relatively simple approach would be to start an online community for sample members and offer online calculators or tools that would allow users to compare their homes' energy use and features to those of an average home or to those of homes in their community. Such tools would not only engage users but would also produce data that could be "scraped" and analyzed by EIA, for example, to track trends in the behavior of the users.

A more involved approach that would work particularly well in combination with a web survey would be to provide each respondent with a brief analysis of his or her responses once the questionnaire is completed. This is a concept similar to one used by the National Health and Nutrition Examination Survey (NHANES), which provides survey participants with the results of their physical examination as an incentive to participate.[1] The results provided to the RECS respondents might compare their homes' energy use to energy efficiency standards or to the energy use of similar houses in the neighborhood, state, or climate zone. The most compelling, although also most resource-intensive way to implement this feature would be to generate the analysis "live" at the end of the survey. Alternatively, or in combination with the live online feedback, a process could be developed for mailing a respondent the report after the interview is completed.

One concern that arises is whether techniques such as these may be disproportionately more likely to engage sample members who are particularly interested in the topics measured by the survey—for example, in the subject of energy efficiency—and who therefore are likely to be different from the rest of the sample. This concern can probably be mitigated by aiming to maintain high overall response rates. In other words, even if the interactive tools attract a subset of the population which differs from the rest of the sample, intensive nonresponse follow-ups, including through different modes, should assure that any bias introduced by disproportionate interest in the incentives among a subset of the population is minimized after follow-up.

The question of how such feedback could affect participation in the proposed longitudinal survey deserves particular attention. An interactive feature could serve as an especially strong incentive if a longitudinal design is introduced. However, it is possible that the feedback provided will alter the behavior of respondents in the longitudinal sample, not only in terms of their likelihood of participating in a second survey, but also in terms of their energy consumption. The benefits and associated risks would have to be evaluated.

[1] For an example of the Final Report of Findings prepared for NHANES respondents, see http://www.cdc.gov/nchs/data/nhanes/nhanes_07_08/ROF_07_08_eng.pdf [December 2011].

DIGITAL PHOTOGRAPHS

EIA has considered the use of digital photographs in the past, and this idea could be revisited, at least as a one-time research effort. Photos may provide valuable information to supplement the interviews conducted in face-to-face administration. They could also be useful in evaluating the quality of self-reported data in web-administered questionnaires. Pictures can help headquarters staff determine whether the definition of a commercial building was applied correctly in the field, or to better understand the layout of residential buildings. Pictures of critical equipment, including nameplates, when it is possible to obtain these, may be useful in the data cleaning and editing process (for example, for reconciling ambiguous or questionable entries). Using cameras in this way, however, would likely involve privacy and confidentiality considerations, and these would have to be researched before the use of cameras could be implemented.

SURVEYS ON SPECIAL TOPICS AND OTHER OPPORTUNITIES FOR EIA

Given the burden that the questionnaire already imposes on respondents, EIA should evaluate creative alternatives for collecting additional data without increasing the overall burden on respondents. For example, EIA could collect data on some specialized topics, such as knowledge, attitude, and behavior related to energy consumption, in the form of a separate survey, independent of the current CBECS or RECS. EIA has added specialized modules to individual cycles of the survey in the past, and it should consider reinstating this practice. There are many possible ways that such specialized data collections could be funded, but one possibility to consider would be a collaboration with other government agencies or organizations that have interest in the data. EIA might even consider serving as the coordinator or data center for research on a number of special topics or special populations.

Collaborations with energy suppliers could also be pursued beyond what is required to obtain data for nonrespondents. Many utilities offer energy efficiency incentives to their customers and sometimes commission studies to evaluate the effectiveness of these programs. EIA could explore opportunities for data collections partially funded by energy suppliers if data could be produced that is helpful for the utilities. If recommendations to collect energy consumption data about all of a utility's customers (instead of just sample members who do not provide adequate data) can be implemented, providing utilities with some data could also become more feasible. A collaboration of this type could serve as an incentive that encourages energy suppliers to standardize their reporting in ways that enables EIA to more efficiently integrate the data received from them.

Additional opportunities for more active dissemination and closer collaboration with the research community include working with other groups collecting extensive data in limited regions or building types. Collaborations with stakeholders can not only be useful in the short term but can also help EIA plan for the future. Providing cost-recoverable custom analysis to the public may be another area to explore.

The panel expects that interest in energy consumption data will continue to grow. Rapid changes in the energy landscape combined with changes in the survey-taking environment require data collection approaches that are continuously updated. Investments into state-of-the-art data collections are necessary to assure that EIA remains at the forefront of energy consumption research.

References

American Association for Public Opinion Research Cell Phone Task Force. (2010). *New Considerations for Survey Researchers When Planning and Conducting RDD Telephone Surveys in the U.S. with Respondents Reached Via Cell Phone Numbers.* Available: http://aapor.org/AM/Template.cfm?Section=Cell_Phone_Task_Force&Template=/CM/ContentDisplay.cfm&ContentID=2818 [September 2011].

de Leeuw, E.D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics 21*(2):233–255.

Dillman, D.A., and J.D. Smyth. (2007). Design effects in the transition to web-based surveys. *American Journal of Preventive Medicine 32*(5):S90–S96.

Dillman, D.A., J.D. Smyth, and L.M. Christian. (2008). *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method, Third Edition.* Hoboken, NJ: Wiley.

French, D.K. (2007). The Energy Consumption Data Program in EIA: A Historical Summary. Unpublished manuscript.

Gold, R., and R.N. Elliot. (2010). *Where Have All the Data Gone? The Crisis of Missing Energy Efficiency Data.* Washington, DC: American Council For An Energy-Efficient Economy. Available: http://www.aceee.org/research-report/e101 [October 2011].

Groves, R.M. and L. Lyberg. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly 74*(5):849–879.

Link, M.W., M. Battaglia, M.R. Frankel, L. Osborn, and A.H. Mokdad. (2008). A comparison of address-based sampling (ABS) versus random-digit dialing (RDD) for general population surveys. *Public Opinion Quarterly 72*(1):6–27.

Lynn, P., ed. (2009). *Methodology of Longitudinal Surveys.* Hoboken, NJ: Wiley.

Mansoor, A. (2008). *Energy Efficiency: Expanded Scale of Opportunities.* Electrical Power Research Institute Seminar Slides. Available: http://mydocs.epri.com/docs/SummerSeminar08/Presentations/S2C%20Mansoor.ppt [September 2011].

Meier, A., L. Rainer, and S. Greenberg. (1992). Miscellaneous electrical energy use in homes. *Energy 17*(5):509–518.

Messer, B.L., and D.A. Dillman. (2011). Surveying the general public over the Internet using address-based sampling and mail contact procedures. *Public Opinion Quarterly 75*(3):429–457.

Michaels, J. (2010). *The Commercial Buildings Energy Consumption Survey: An Overview.* Presentation to the Panel on Redesigning the Commercial Buildings and Residential Energy Consumption Surveys of the Energy Information Administration, February 1, National Research Council, Washington, DC.

Millar, M.M., and D.A. Dillman. (2011). Improving response to web and mixed-mode surveys. *Public Opinion Quarterly 75*(2):249–269.

National Research Council. (2002). *Making the Nation Safer: The Role of Science and Technology in Countering Terrorism.* Committee on Science and Technology for Countering Terrorism. Washington, DC: The National Academies Press.

National Research Council. (2009). *Principles and Practices for a Federal Statistical Agency, Fourth Edition.* Committee on National Statistics. C.F. Citro, M.E. Martin, and M.L. Straf, eds. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Pfund, N., and B. Healey. (2011). *What Would Jefferson Do? The Historical Role of Federal Subsidies in Shaping America's Energy Future.* Available: http://www.dblinvestors.com/documents/DBL_energy_subsidies_paper.pdf [October 2011].

Roth, K., K. McKenney, and J. Brodrick. (2008). Emerging technologies: Small devices, big loads. *ASHRAE Journal 50*(6):64–65. Available: http://bookstore.ashrae.biz/journal/journal_s_article.php?articleID=829 [September 2011].

Schwarz, N. (1996). *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation.* Mahwah, NJ: Lawrence Erlbaum Associates.

Smyth, J.D., D.A. Dillman, L.M. Christian, and A.C. O'Neill. (2010). Using the Internet to survey small towns and communities: Limitations and possibilities in the early 21st century. *American Behavioral Scientist 53*(9):1423–1448.

U.S. Department of Energy. (2011a). *Department of Energy Strategic Plan.* Available: http://www.energy.gov/media/DOE_StrategicPlan.pdf [October 2011].

U.S. Department of Energy. (2011b). *Budget Justifications and Supporting Documents.* Available: http://www.cfo.doe.gov/budget/12budget/index12.html [September 2011].

U.S. Energy Information Administration. (2009). *State Energy Data Needs Assessment.* Available: http://www.eia.gov/oiaf/servicerpt/energydata/pdf/sremeu(2009)01.pdf [December 2011].

U.S. Energy Information Administration. (2011a). *International Energy Statistics.* Available: http://www.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm?tid=44&pid=44&aid=2 [September 2011].

U.S. Energy Information Administration. (2011b). *Total Energy Consumption, Price and Expenditure Estimates, 2009.* Available: http://www.eia.gov/states/sep_fuel/html/pdf/fuel_te.pdf [September 2011].

U.S. Energy Information Administration. (2011c). *Annual Energy Outlook 2011 with Projections to 2035.* Available: http://www.eia.gov/forecasts/aeo/pdf/0383(2011).pdf [September 2011].

U.S. Energy Information Administration. (2011d). *How Dependent Are We on Foreign Oil?* Available: http://www.eia.gov/energy_in_brief/foreign_oil_dependence.cfm [September 2011].

U.S. Energy Information Administration. (2011e). *Commercial Buildings Energy Consumption Survey: CBECS Status*. Available: http://www.eia.gov/emeu/cbecs/ [September 2011].

Weisberg, H. (2005). *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago, IL: University of Chicago Press.

Appendix A

Energy Independence and Security Act of 2007 (P.L. 110-140)

Sec. 805. Assessment of Resources.

(a) 5-Year Plan

(1) ESTABLISHMENT—The Administrator of the Energy Information Administration (referred to in this section as the 'Administrator') shall establish a 5-year plan to enhance the quality and scope of the data collection necessary to ensure the scope, accuracy, and timeliness of the information needed for efficient functioning of energy markets and related financial operations.

(2) REQUIREMENT—In establishing the plan under paragraph (1), the Administrator shall pay particular attention to—
(A) data series terminated because of budget constraints;
(B) data on demand response;
(C) timely data series of State-level information;
(D) improvements in the area of oil and gas data;
(E) improvements in data on solid byproducts from coal-based energy-producing facilities; and
(F) the ability to meet applicable deadlines under Federal law (including regulations) to provide data required by Congress.

(b) Submission to Congress—The Administrator shall submit to Congress the plan established under subsection (a), including a description of any improvements needed to enhance the ability of the Administrator to collect and process energy information in a manner consistent with the needs of energy markets.

(c) Guidelines

(1) IN GENERAL—The Administrator shall--
(A) establish guidelines to ensure the quality, comparability, and scope of State energy data, including data on energy production and consumption by product and sector and renewable and alternative sources, required to provide a comprehensive, accurate energy profile at the State level;
(B) share company-level data collected at the State level with each State involved, in a manner consistent with the legal authorities, confidentiality protections, and stated uses in effect at the time the data were collected, subject to the condition that the State shall agree to reasonable requirements for use of the data, as the Administrator may require;
(C) assess any existing gaps in data obtained and compiled by the Energy Information Administration; and
(D) evaluate the most cost-effective ways to address any data quality and quantity issues in conjunction with State officials.
(2) CONSULTATION—The Administrator shall consult with State officials and the Federal Energy Regulatory Commission on a regular basis in--
(A) establishing guidelines and determining the scope of State-level data under paragraph (1); and
(B) exploring ways to address data needs and serve data uses.

(d) Assessment of State Data Needs—Not later than 1 year after the date of enactment of this Act, the Administrator shall submit to Congress an assessment of State-level data needs, including a plan to address the needs.

(e) Authorization of Appropriations—In addition to any other amounts made available to the Administrator, there are authorized to be appropriated to the Administrator to carry out this section--
(1) $10,000,000 for fiscal year 2008;
(2) $10,000,000 for fiscal year 2009;
(3) $10,000,000 for fiscal year 2010;
(4) $15,000,000 for fiscal year 2011;
(5) $20,000,000 for fiscal year 2012; and
(6) such sums as are necessary for subsequent fiscal years.

Appendix B

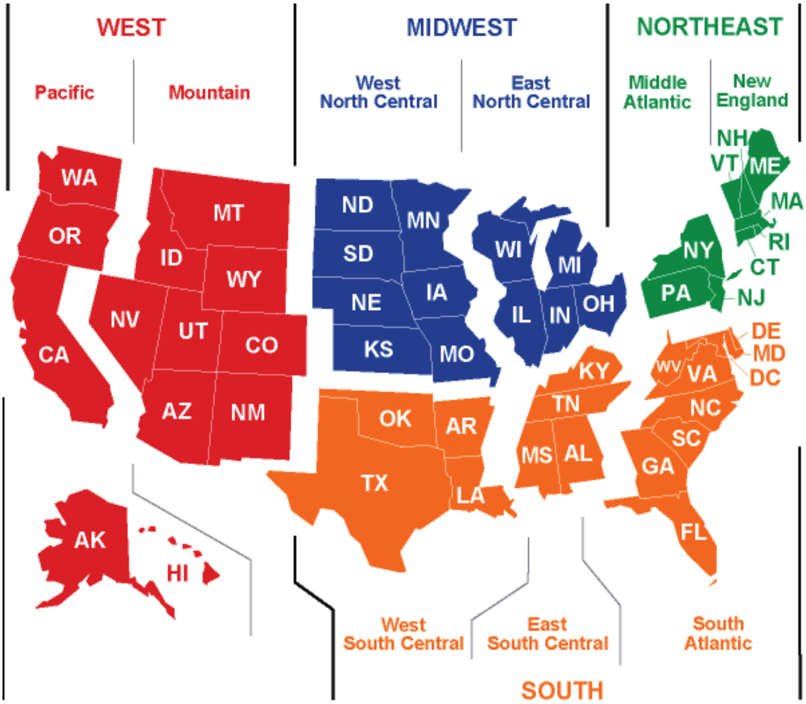
Data Users Who Provided Input to the Panel

American Council for an Energy-Efficient Economy
Carnegie Mellon University Green Design Institute
Consortium for Energy Efficiency
Energy Star Commercial and Industrial Branch Office of Air and Radiation
Energy Star Residential Branch
Institute for Market Transformation
ITRON
KEMA
Koomey, Jonathan, Stanford University and Lawrence Berkeley National Laboratory
Lawrence Berkeley National Laboratory, Energy Analysis Department
Lawrence Berkeley National Laboratory, Environmental Energy Technologies Division
Mississippi Power
National Multi-Housing Council
National Renewable Energy Laboratory, Electricity, Resources and Building Systems Integration Center
National Renewable Energy Laboratory, Residential Research Group
Northeast Energy Efficiency Partnership
Oak Ridge National Laboratory, Building Technologies Research and Integration Center
Resources for the Future

Schipper, Lee, Precourt Energy Efficiency Center, Stanford University
Sudarshan, Anant, Precourt Energy Efficiency Center, Stanford University
TIAX, LLC
U.S. Department of Energy, Building Technologies Program
U.S. Department of Health and Human Services, Agency for Children and Families, Division of Energy Assistance
U.S. Energy Information Administration, Office of Integrated Analysis and Forecasting

Appendix C

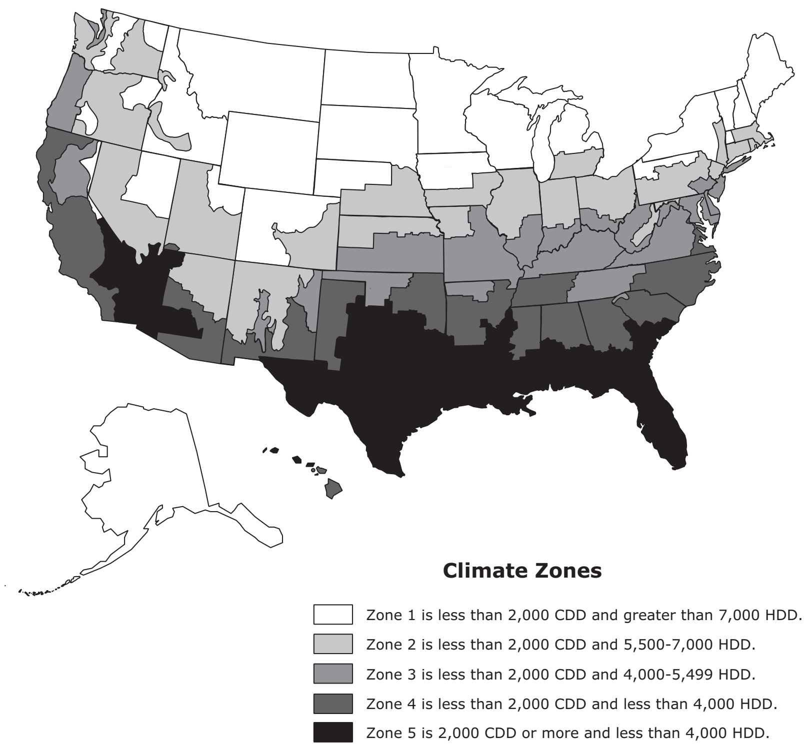
U.S. Census Regions and Divisions



SOURCE: http://www.eia.gov/emeu/reps/maps/us_census.html [December 2011].

Appendix D

U.S. Climate Zones for the 2003 CBECS



SOURCE: http://www.eia.gov/emeu/cbecs/climate_zones.html [December 2011].

Appendix E

Composite Estimation

This appendix briefly presents the simplest composite estimator for a rotating panel design. There is an extensive literature on composite estimation to which the reader can turn for descriptions of more elaborate approaches.

This discussion assumes an annual data collection. Let θ_t denote a parameter to be estimated in year t, and let $\hat{\theta}_t$ denote the "usual unbiased estimate" of θ_t based only on data from time t. Note that it is also possible to compute an estimate $\hat{\theta}_{t,m}$ based only on the overlapping (or "matched") sample, i.e., using only the sample that is being interviewed for the second time (or more). Since values exist from the previous year for the matched sample, it is possible to compute $\hat{\theta}_{t-1,m}$, and so $\hat{\theta}_{t,m} - \hat{\theta}_{t-1,m}$ is an estimate of the change between year t and year $t-1$ based on the same sample units. By adding this estimate of change to last year's estimate of θ_{t-1}, a second estimate of θ_t, namely, $\hat{\theta}_{t-1} + \hat{\theta}_{t,m} - \hat{\theta}_{t-1,m}$ is obtained. (In words, the estimate for this year is equal to the estimate for last year plus the estimated year-to-year change.) Since there are now two estimates of the same quantity θ_t, it is natural to combine them to obtain an improved estimate. Thus the so-called composite estimate of θ_t is $\hat{\theta}_t^c = \alpha\hat{\theta}_t + (1-\alpha)(\hat{\theta}_{t-1} + \hat{\theta}_{t,m} - \hat{\theta}_{t-1,m})$, a weighted average of two estimates (assuming α is between 0 and 1). Because $\hat{\theta}_{t-1}^c$ is a better estimate than $\hat{\theta}_{t-1}$, then in practice, the following composite estimate is used: $\hat{\theta}_t^c = \alpha\hat{\theta}_t + (1-\alpha)(\hat{\theta}_{t-1}^c + \hat{\theta}_{t,m} - \hat{\theta}_{t-1,m})$.

The value of α is chosen to minimize the variance of $\hat{\theta}_t^c$.

Appendix F

Letter Report

This appendix reprints the panel's letter report (National Research Council, May 2010). The only changes to the text as it appeared in the letter report were to drop the Appendixes, which are reproduced elsewhere in the final report.

Dr. Richard Newell
Administrator
U.S. Energy Information Administration
1000 Independence Avenue, SW
Washington, DC 20585

Dear Dr. Newell:

At the request of the U.S. Energy Information Administration (EIA), the Committee on National Statistics (CNSTAT) of the National Research Council convened a panel to conduct a comprehensive 30-month study of the Commercial Buildings Energy Consumption Survey (CBECS) and Residential Energy Consumption Survey (RECS). Many of the design and operational procedures for the CBECS and RECS were developed in the 1970s and 1980s, and resource limitations during much of the time since then have prevented EIA from making significant changes to the survey methodology or operations. With the possibility of additional funding available in the next few years, EIA asked the National Research Council to conduct a comprehensive review to assess how the CBECS and RECS can take advantage of recent developments in survey methods and to ensure the relevance of the data for meeting increased user needs in the next decade and beyond. The panel's charge is to consider possible improvements to data quality, geographic coverage, relevance, and the timeliness of data releases.

Because plans for the upcoming 2011 round of CBECS must be finalized in the near future, the panel was charged to comment as soon as possible on design and data collection options that would enable the upcoming round of this survey to better support U.S. Department of Energy program information needs, reduce respondent burden, and increase the quality and timeliness of the data. This letter responds to that request, and is limited in scope to discussing issues that the panel believes are realistic to consider in the timeframe leading up to the 2011 data collection. At the conclusion of the study, the panel will deliver its comprehensive report on the overall design and conduct of both CBECS and RECS.

At the first meeting of the panel on February 1-2, 2010, EIA staff discussed preparations for the 2011 CBECS and indicated that work will begin on the 2011 CBECS sample design in the summer of 2010. Thus, any changes to this round of the data collection would have to be evaluated before then. EIA staff also informed the panel that the 2011 CBECS is anticipated to have more funding than it has had in the past. The panel also learned in those discussions that EIA has relatively little empirical data on how well the current design and procedures are working in comparison with approaches that have been tried in the past and that EIA has not conducted an analysis of options considered but not pursued. Based on the factors described above, the panel's overarching recommendation is to invest some of the currently available additional funding in research that will enable EIA to make future decisions based on empirical evidence about what is most likely to improve geographic coverage, data quality and relevance, while controlling costs. The panel's specific recommendations for research as part of the 2011 CBECS are described below.

BACKGROUND ON THE CBECS

The CBECS is a survey of commercial buildings in the United States, mandated by Congress to provide comprehensive information about energy use in commercial buildings. In addition to energy consumption and expenditure data, the survey collects information about building characteristics, such as energy source, physical structure, equipment used, and activities performed, which provides researchers with detailed information about commercial sector energy use and how it relates to building characteristics. The CBECS is the only national source of these data, and is used for energy forecasting, program development, and policy development.

The survey collects information from a sample of commercial buildings in the United States, and it is administered in two data-collection stages: a Building Characteristics Survey and an Energy Suppliers Survey. During the first stage of the data collection, interviewers visit the buildings selected into the sample and ask a representative of the building, such as the building's owner, manager, or other knowledgeable person to complete the survey. During the second stage of the data collection, the energy suppliers of buildings that were not able to provide adequate information in the first stage are contacted to obtain actual usage and expenditure data from the supplier's records.

SAMPLING FRAME

There is little comprehensive information about the stock of commercial buildings in the United States, and EIA indicated that the lack of a comprehensive national list of commercial buildings or another suitable source from which to select a sample of buildings to interview is one of the major challenges for the CBECS data collection. Because no complete list of buildings is available to use as a sampling frame, EIA builds a new area probability sampling frame for the CBECS on a decennial basis. The frame is based on field listings of commercial buildings within specified geographic areas. This sampling frame is updated between each data collection. However, field listings are resource intensive and relying on sources that are not comprehensive for updating the sampling frame leads to coverage problems.

The CBECS sample design has undergone numerous revisions over the years, as EIA has attempted to address the cost and coverage issues, but most rounds of the CBECS have relied on a combination of an area frame and a list frame, based on existing lists of commercial buildings from a variety of sources and added at the second stage of the area frame sample. The primary sampling units have been counties or groups of counties, within which smaller geographic areas were randomly selected. All commercial buildings were listed and stratified within these smaller areas, and then a sample of buildings was randomly selected from each stratum. This approach was supplemented with information from existing building lists from other sources to ensure adequate representation of buildings that were of special interest because of their size or type of primary activity.

For the 2007 administration of the CBECS, the 2003 sampling frame had to be updated. At the recommendation of the data collection contractor, the National Opinion Research Center (NORC), the update was based on a U.S. Postal Service (USPS) Delivery Sequence File (DSF) purchased from a vendor licensed by USPS. The DSF is USPS's list of all delivery points in the United States. Using the DSF for updating meant that this list had to be matched to the addresses in the second-stage area frame and the duplicates removed. NORC reported that the unduplication turned out to be a major challenge, in part because of imprecise address records.

As EIA is aware, another major redesign of the CBECS sampling frame could be very productive, but due to the limited time and resources available, this is neither feasible nor recommended for the 2011 data collection. However, leading up to and during the 2011 CBECS, alternative approaches to building a good second-stage sampling frame should be the focus of EIA research, particularly the availability of administrative records and lists. As the EIA staff indicated, and the panel concurs, a sampling frame based on administrative records may have to completely or partially replace the second-stage area frame in the future because of the high costs associated with field listing. Although EIA has considered the use of more lists throughout the years, research on this should continue because the availability of sources of data is constantly evolving, particularly with more information becoming available on the Internet.

For the 2011 CBECS, the most practical approach is to perform another round of updating of the sampling frame using the DSF. Even though unduplication proved to be a challenge when the DSF was first used in 2007, presumably the bulk of the work has now been done, and the 2011 frame can be updated by simply matching the new addresses to the address files used in 2007. We assume that a 2003-2007 cohort of listings is available for use in the 2011 sample based on the matching and updating performed in preparation for the 2007 data collection. As was done in previous years, this approach would have to be supplemented with lists from other sources to assure adequate representation of buildings of special interest.

As an example of such a supplemental source, we recommend exploring the usefulness of local government databases that are available online, such as county property tax databases, some of which include information on square footage and heat source. Two available online databases of which we are aware are those of the Allegheny County Office of Property Assessments in Pennsylvania and of the King County Government in Washington. Although such databases are not universally available online, and their use would undoubtedly present some standardization challenges, their usefulness should be evaluated for two purposes: as a source for a sampling frame and for the possible use of some of the data that are now collected through interviews. Other possible data sources are discussed in the next section, although all require further research to evaluate them.

In rural areas, the DSF often includes only rural route or post office box numbers and so tends not to be very useful. EIA should evaluate information available from vendors who specialize in providing address data to fill these types of gaps. If these sources are found to be inadequate, field listing may still be necessary. Alternatively, half-open interval updating could be considered, if relisting is deemed too inefficient because of the scattered nature of rural areas. This technique involves updating only new or missed units within a small geographic area (an "interval" associated with an address in the sample). In areas where buildings are scattered in unusual ways, half-open interval updating may be difficult to carry out accurately, but the accuracy of the approach in this particular context could be evaluated as part of the 2011 CBECS. For example, in addition to performing half-open interval updating in the rural sample segments, relisting could also be carried out in a subset of these segments to compare the outcome of the two techniques in terms of the number of listings identified and the number that would end up being added to the 2011 CBECS. The relisting could be performed by experienced listers or supervisors to minimize the costs associated with training for these types of assignments.

Further research is needed to understand the quality and future potential of the DSF. In addition to evaluating the performance of the DSF in comparison with other sources for a sampling frame, the panel recommends adding a question to the CBECS questionnaire to better understand the extent to which there is overlap between street addresses and the addresses where the building occupants receive their mail. For example, one challenge is that the DSF contains business-level entries, rather than building-level entries. Furthermore, some businesses have their mail delivered somewhere other than the street address (for example, to a post office box).

SUPPLEMENTARY DATA SOURCES

In addition to considering existing administrative records as an alternative source for a sampling frame, the panel recommends evaluating these records as potential sources for substantive data that could possibly replace an on-site interview at the building's location or could provide additional data for modeling or to conduct new analyses. Relying on data from other sources may become more of a necessity as it becomes increasingly expensive to maintain high response rates, even if an ideal sampling frame of commercial buildings were available. Although gathering and combining data from a variety of administrative records can also be resource intensive, the costs may go down as such data become more widely available, especially online.

Such other sources may also provide higher quality data. For example, EIA staff have expressed concerns regarding the difficulties associated with collecting data about the technical topics covered in the CBECS survey. Neither the typical interviewer nor the typical respondent is particularly knowledgeable about many of the items in the questionnaire, and this raises the question whether there are other sources that could provide better quality data.

Because the CBECS is currently the most comprehensive data available on the energy consumption of commercial buildings in the United States, there is no "gold standard" against which the quality of the survey or other potential data sources can be evaluated. However, a variety of other sources exist and can provide at least partial data. Comparing the data from several of these sources will help EIA begin to understand the relative advantages associated with each and the optimal ways of combining information from different sources. Ultimately, conducting in-person interviews for at least a subset of the sample may be necessary for validation purposes, if the research indicates that these interviews produce the highest quality data.

In the rest of this section we discuss some additional sources of data that should be explored and validated at this stage, even if none of them, by itself, represents a realistic replacement for the building interviews. The panel is aware that EIA has considered the use of a variety of administrative records over the years. This option should be revisited periodically as the costs and data quality benefits associated with integrating these data sources evolve, and the 2011 CBECS presents an opportunity to carry out this research. Although different sources may be available for different types of buildings, for the purposes of evaluating them the overlap should be maximized to the extent possible. In other words, a subset of the buildings should be selected for this research and, for these buildings, data should be gathered from all of the available data sources being evaluated. The overlap will be particularly important with the building audits, as discussed below.

Building Audits

We understand that EIA has considered involving professional energy auditors to collect building data instead of relying on interviewers, but there has never been sufficient funding to implement this approach (except in the form of a small study in the past that was not conducted in conjunction with any of the CBECS data collections). We recommend testing the use of auditors on a small scale in the 2011 CBECS to better understand the costs and to collect data that can be used to assess the quality of other data sources. The data collected by the auditors would also be useful for evaluating some of the current back-end procedures, such as data editing, or the regression model used to identify outliers and to initiate a supplier follow-up survey.

For a subset of the buildings we think it would be useful if the same data were collected by both interviewers and auditors to allow the evaluation of the differences between these two in-person data collection approaches, in addition to comparing them to information collected from other data sources. Even if geographic representation cannot be achieved due to cost considerations, to the extent possible the test should include buildings of different sizes and with different characteristics. The data collection should be performed by professional energy auditors, who would carry out their work around the same time as the other data collection efforts relevant to a particular building, and without knowledge of any data already collected or available about the buildings from other sources.

Online Research

The panel recommends selecting a small subset of the buildings in the 2011 CBECS sample and investigating the information that can be obtained about them solely from the Internet. This could be set up in the form of a pilot test involving a small number of buildings (for example, 10 large, 10 medium, and 10 small buildings). The results will provide EIA with a better understanding of what types of data are available online in terms of both quality and extent. If this research is scheduled before the beginning of the data collection, the insights gained could be useful in fine-tuning the data collection instruments and sample design for the 2011 CBECS, but this type of research could be carried out at any time. Once the 2011 CBECS data are available, examining the consistency between the information available from a variety of Internet sources and the data collected through the current methods, as well as from building audits, will provide valuable information about data quality.

LEED and Energy Star Certified Buildings

Data quality can also be assessed by taking advantage of the information that is publicly available online about buildings that have received leadership in energy and environmental design (LEED) or energy star certification. Comparing the data collected through the building interviews about a subsample or all of the buildings that have such certification to the data submitted as part of the certification process for the same buildings can also contribute to a better understanding of possible data quality issues. Naturally, certified buildings are a specific subset of the CBECS sample, and their characteristics are not representative of the entire population of interest for the survey. However, examining any discrepancies in the data available about these buildings can improve EIA's overall sense of the quality of the data and also identify potential areas of concern. The comparison could even be performed on data that have already been collected through a previous round of CBECS to inform the 2011 design.

Data from Energy Suppliers

The CBECS includes an energy supplier survey for about half of the CBECS buildings in the sample. The survey is initiated in cases where the energy usage and cost information cannot be obtained through a building interview or if the data obtained through the building interview are flagged as out of the expected range based on a regression model developed by EIA. To evaluate the quality of the data obtained from the building interviews, as well as the regression model used to identify out of range responses, the next round of the CBECS should include an energy supplier follow-up for more than just the problem cases currently included. For example, the 2011 CBECS could collect supplier data for a random sample of cases that provided usage and consumption data that were deemed valid during the building interview.

Given the increasing interest in time-of-use, hourly, and real-time energy use data, the panel recommends collecting this type of information as well, where available. EIA could aim to collect hourly data or time-of-use data (along with rates) from a random sample of the suppliers contacted for a follow-up interview, all suppliers who are contacted for a follow-up interview, or a random sample of the suppliers for buildings for which interviews were also conducted.

In addition, it is possible to specifically identify a few buildings with real-time metering and explore the data available from this source. The goal, again, would be to start establishing a framework for integrating this type of data into future surveys, understanding what type of information can be collected, and fine-tuning the procedures for working with a variety of (often reluctant) energy suppliers.

Digital Photographs

EIA has considered the use of digital cameras in the past, and this idea should be revisited, at least as a one-time research effort. A test could be accomplished by either providing a small number of the 2011 CBECS interviewers with cameras or in a separate operation from the 2011 data collection, if the latter approach is deemed more cost-effective because of training and operational considerations.

Photographs may provide valuable basic information about buildings. EIA's definition of a building does not always correspond to a respondent's definition of a building, and it is often left up to the interviewer to clarify the definition and come to an understanding with the respondent about what is meant by a building for the purposes of the interview. In addition, EIA staff indicated to the panel that the buildings of interest tend to use a wide range of specialized equipment related to building activity, and respondents' abilities to describe the equipment vary. EIA should evaluate the extent to which the pictures of the buildings and critical equipment, including nameplates, are useful in the data cleaning and editing process (for example, for reconciling ambiguous or questionable entries) and whether their use could translate into cost savings in comparison to the current procedures. EIA should also investigate the privacy and confidentiality concerns and regulations that may be applicable to the potential use of cameras in this context, even if the pictures are only used for data cleaning and editing.

Geocoding

EIA should examine the costs and benefits of supplementing the data available about each case in the sample with the geographic coordinates of the building's address. Geocodes could be added to the sample in house or during the data collection process. The former approach would probably be less precise, so capturing this information during the field work would be preferable if the interviewers can be equipped with the necessary devices at a reasonable cost. Adding geospatial information to each of the cases in the sample will enable researchers to conduct additional analyses of the CBECS data. For example, EIA currently integrates weather data from the National Oceanic and Atmospheric Administration (NOAA) into some of its analyses, and recording the building's proximity to the closest weather station would expand the analytic possibilities. Again, EIA should conduct research on any potential confidentiality concerns related to the use of this type of data and whether there are ways of appending additional geographic information to the data while maintaining confidentiality.

Other Data Sources

The panel recommends evaluating other existing data sources that EIA has considered in the past, as well as the breadth and consistency of information that could be obtained from local governments. The availability of more funding for the 2011 CBECS than has been available in the past provides a unique opportunity to carry out research that can inform future decisions about the design of the survey. Even if the information available from the various data sources is limited in scope, the recommended research can provide valuable feedback about the quality of the self-reported building data and identify options for integrating a variety of data sources in the future.

DATA COLLECTION

Process

The panel understands that EIA staff participate in all interviewer training, but even more active involvement may be necessary to share the study's goals and communicate how the quality of the data determines their usefulness. In addition, EIA staff members are also best qualified to conduct training on topics and concepts that are complicated, as a result of a long institutional history, such as the definition of a building and of a qualified respondent.

Additional resources should be invested in analyzing the characteristics of the field operations and in identifying opportunities for increased efficiency. EIA should review any information available from the data collection contractor regarding the amount of time spent on cases of various types (such as buildings with different characteristics, respondents with different backgrounds, etc.). If the case level contact history is not recorded in sufficient detail, efforts should be made to capture this information in the future. In addition, EIA should ask the data collection contractor to schedule debriefings with the interviewers soon after the beginning of the field period, and EIA staff should attend these debriefings to better understand how interviewers spend their time in the field, what types of cases are presenting the biggest challenges, and why. A detailed analysis of the time allocation should reveal whether there are subsets of cases that require a disproportionately large amount of time to complete and whether the effort is justified in the context of data needs and statistical techniques available to compensate for missing information.

EIA should also work closely with the data collection contractor to review the procedures used to select the best respondent for the building interviews and identify opportunities to streamline this process. Again, debriefings with interviewers can provide invaluable feedback that can help fine-tune the process and contribute to the development of new interviewer protocols. More efficient procedures for identifying a qualified respondent can not only reduce costs, but also address some of the concerns related to the technical nature of the questions. The qualitative feedback from the interviewers can then be further examined with an analysis of the quantitative responses by respondent type to identify possible differences in data quality. In other words, it is possible that most of the questions are not "too technical" if posed to the right respondent.

Additional activities for which analyzing existing data could identify opportunities for increased efficiency include the handling of partial interviews, both in terms of the field operations and from the perspective of data editing. Given that the CBECS interview is relatively long (with an estimate of 30 to 45 minutes provided to respondents), interviewer debriefings could reveal new strategies for approaching buildings and asking for appointment times. It would also be useful to understand whether there is a pattern to at what point the partial interviews end and whether the order of the items in the questionnaire could be rearranged to make the partial interviews more useful to EIA for either weighting or imputation.

Interviewers can be a good source of background and contextual information on questions that are difficult to administer, especially on whether particular questions are leading to partial interviews or possible data quality concerns. Discussions with the interviewers could represent the beginning of a close examination of the questionnaire that has evolved with a face-to-face administration in mind and may need revising or simplifying to accommodate different future modes of data collection, as described in the next section.

Modes

Although EIA has considered the use of other modes of data collection, CBECS data are still collected primarily by in-person interviewing. In part because applying the CBECS definitions to determine the boundaries of a building is not always a straightforward task, as discussed above, EIA has continued to rely on face-to-face interviews. Identifying the most appropriate respondent is another task that is thought to benefit from the presence of an interviewer. Interviewers also carry hard-copy "show cards" that list the answer options for specific items and can be handed to the respondent to assist with answering questions that may otherwise be too difficult to remember if only read by an interviewer. In addition, one of the roles of the interviewers is to scan utility bills if they are available.

To prevent declines in the response rates and to limit costs, EIA will have to revisit the use of other modes of data collection, particularly the possibility of a multimode approach, with at least a portion of the interviews being conducted online. Transitioning at least a subset of the buildings to the web will free up some resources in the long run, which then can be allocated to the more complex cases and possibly invested into increasing the sample size. Although collecting this type of data on the web will present some methodological challenges, the panel believes that these challenges can be addressed and that web data collection may also represent some methodological advantages, in addition to the likely cost savings.

One possible approach that should be explored is to divide the sample into buildings that can be relatively easily transitioned to a web administration and buildings with more complicated characteristics that may benefit from interviewer administration. It may also be necessary to treat large buildings differently from smaller ones. The review of the case histories and the interviewer debriefings described above will be helpful in beginning to identify the building types for which data collection is fairly straightforward.

Until reliable auxiliary data sources can be integrated into the data collection process, a first in-person visit to each building will still be useful. During this visit, interviewers should follow a protocol developed by EIA to determine whether a second in-person visit is necessary (as is currently done) or whether the building is a good candidate for a web interview. Given the concerns related to the definition of a building, the decision of whether a case can be transitioned to the web will likely have to depend in part on whether the definition seems straightforward, as it would be, for example, for a small, standalone building occupied by one business. Resources should be invested in testing ways of communicating the definition of a building through a self-administered format, in anticipation of possibly being able to transition more and more complex buildings to web administration in the future.

The logistics of the best way to collect contact information for a web survey would have to be explored. Possible options include obtaining the information during the first visit or by telephone. Sometimes information on how to access a web survey is included in a hard-copy advance letter mailed to respondents, even though this is less ideal than an e-mail invitation because it requires respondents to manually enter the web address of the survey and the login information. Since contacting respondents by mail may be the only option if an e mail address cannot be obtained, investigating the extent to which building addresses and the mail delivery addresses overlap (as recommended above) will also be useful for this purpose.

There is no question that identifying the best respondent for completing the interview is crucial in the case of the CBECS, but exploring ways of accomplishing this without involving an interviewer should be examined. It is possible that a web option could in fact contribute to more interviews being conducted with qualified respondents. In some cases, it may be easier to forward a questionnaire to the right person than to locate him or her in a building and arrange an interview. In-person interviewers may also have an incentive to complete an interview as quickly as possible by settling for a willing respondent rather than pursuing the most appropriate one.

If a questionnaire is available on the web, it is also easier for several respondents to collaborate, each completing the sections he or she is most knowledgeable about. In addition, a web option could result in more complete data because it gives respondents the option to obtain information for questions they are not sure about and resume the survey later. Naturally, if respondents stop or forward a survey, there is a risk that they will not return to complete it, so an extensive follow-up effort is likely to be necessary. However, when a topic is too technical for many respondents, such as is the case of the CBECS, this kind of follow-up could make a significant difference in data quality.

The CBECS questionnaire relies very heavily on show cards, which is another reason why the survey is administered face to face. However, the use of show cards raises the concern of order effects, especially because many of the show cards have a large number of answer options listed, making it difficult for respondents to focus equally on all of them. For example, the show card listing the answer options for the primary activity in the building contains 16 items. A respondent for a building with multiple activities may be tempted to select the first one that is applicable as the "primary" activity instead of carefully reviewing the entire list. A web questionnaire would make it easier to restructure these questions into layered sets of items, with fewer answer options, or to reduce the possibility of primacy effects with the use of innovative methods, such as the animated presentation of response choices or an eye-catching emphasis on the end of the list.

The additional funding available for the 2011 CBECS represents an opportunity to test various ways of asking questions that EIA has identified as problematic because of their technical nature. For example, different approaches to obtaining the square footage information from respondents can be tested in the form of a split-sample experiment. Deconstructing this kind of an item into a series of questions would introduce complex skip patterns, but it would be easy to implement on the web without increasing the cognitive burden on respondents. Web administration can integrate various aids and tools for respondents, such as definitions or diagrams that can pop up if a respondent seems to be having trouble with a question or requests help. The interviewer debriefings described above will be useful in pinpointing specific questions that could benefit from a different approach and whether web administration is a promising option to pursue.

As is always the case with self-administered surveys, providing respondents with an e-mail address and toll-free telephone number they can use if they have questions may be valuable. The staff accessible through these means should be able to provide assistance related to the technical topics in the questionnaire, as well as answer to questions specific to the web administration.

Finally, when evaluating the implications of transitioning to a mixed mode administration, options for collecting the utility bills that are currently collected during the interview should also be considered. Some respondents may be able to easily upload an electronic copy of their bills through the questionnaire website, and this possibility should be investigated. Asking respondents to mail a copy of their utility bills would probably not be cost-effective because extensive follow-up would likely be necessary. The options should be assessed in the context of the research conducted to evaluate the possibility of increased reliance on supplier data.

The ideal time for beginning to explore the feasibility of transitioning some of the sample to web administration and conducting experiments on question wording is in parallel with the other data validation efforts, especially the involvement of the energy auditors. This timing will allow EIA to allocate some of the funding currently available to cover the cost of the transition, and it will provide an opportunity to take maximum advantage of the data collected from different sources. An analysis of the data collected from different sources can also guide decisions on whether the in-person interviews can be used in the future to calibrate the data collected through the web. All of the research should keep long-term plans in mind, such as the characteristics of the web as a data collection mode, even if the integration of web interviews is not realistic for the 2011 CBECS.

DATA RELEASES

The panel learned that the schedule of the data releases is a major concern to users who would like to see the lag between the data collection and release date reduced. EIA has been working on taking greater advantage of the Internet to facilitate data distribution.[1] We note that introducing a web option during the data collection stage can, in the long run, reduce the time necessary for preparing the files for release by reducing data editing and cleaning time.

The panel also recommends evaluating the possibility of eliminating some of the editing steps by reducing the number of editing rules or the number of variables edited or by focusing on cases that have the most impact on the estimation. Many data users will not only appreciate a shorter lag between data collection and release, but may also prefer access to data with fewer edits.

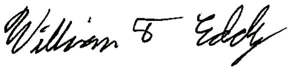
SUMMARY

The 2011 CBECS presents an opportunity to conduct research that can guide the redesign of the survey on the basis of empirical data about the most cost effective approach for collecting valid and reliable information about the energy consumption of commercial buildings. This letter report of the panel outlines a variety of research topics that seem most promising to pursue before or as part of the 2011 CBECS data collection. EIA should

[1] Energy Information Administration. (2009). State Energy Data Needs Assessment. Report SR-EMEU(2009)01. Washington, DC: U.S. Department of Energy.

focus its efforts on (1) evaluating the availability and quality of alternative data sources that could assist with sampling frame development and potentially provide substantive data, and (2) developing a strategy for transitioning some of the interviews to a web-based data collection mode. This research will inform a possible future redesign of the sampling methodology and revisions to the data collection procedures that could be considered for subsequent rounds of the CBECS.

We hope this letter and our recommendations are helpful to you in planning the 2011 CBECS.



William F. Eddy
Chair

Appendix G

Glossary

area probability sample a sample generated by dividing a geographic area into a number of smaller areas and then sampling from a subset of these areas

coverage error bias resulting from the omission of units from the sampling frame

end use a specific activity that requires energy (for example, space heating or refrigeration)

Energy Star certification a joint program of the U.S. Environmental Protection Agency and the U.S. Department of Energy that certifies energy-efficient products

geocoding the process of appending geographic identifiers (codes or coordinates) to an address

hot deck imputation a method for handling missing data by filling in a missing value with a response from another case within the same data set

LEED certification a building certification system providing verification that a building was designed and built according to a set of "green" standards

list sample a sample generated from a sampling frame that exists in a list form (such as a list of housing unit addresses)

multistage sampling a sampling process involving several stages, in which units at each subsequent stage (for example, households) are subsampled from previously selected larger units (for example, city blocks or neighborhoods)

sampling frame the set of units from which the sample is selected

sampling units the individual units selected from the sampling frame

show card an interviewing aid consisting of a paper version of answer options or definitions associated with questionnaire items and used during an in-person interview when the questions are read to the respondent and may be too difficult to understand or remember without a visual aid (also referred to as hand cards or flash cards)

smart meter electric meter that transmits electricity consumption data to the utility and typically enables a consumer to see a detailed breakdown of the electricity usage information collected by the meter

stratified sample a sampling technique that involves dividing the sampling frame into distinct subgroups of similar units and then selecting a separate sample from each of the subgroups

Appendix H

Biographical Sketches of Panel Members and Staff

WILLIAM F. EDDY (*Chair*) is John C. Warner professor of statistics at Carnegie Mellon University, and he also holds appointments in the School of Computer Science and the Department of Biological Sciences. He is an elected fellow of the American Statistical Association, the Institute of Mathematical Statistics, and the American Association for the Advancement of Science and is an elected member of the International Statistical Institute. He has served on two National Research Council (NRC) boards, seven NRC panels, and eight NRC committees. He is former chair of the Committee on National Statistics (CNSTAT) and was previously the chair of the Committee on Applied and Theoretical Statistics. Some of the panels he has served on are the Panel to Update the Fourth Edition of Principles and Practices for a Federal Statistical Agency; the Committee on Assessing the Feasibility, Accuracy, and Technical Capability of a National Ballistics Database; the Panel on Enhancing the Data Infrastructure in Support of Food and Nutrition Programs, Research, and Decision Making; the Panel to Review the Statistical Procedures for the Decennial Census; and the Committee to Review the Bureau of Transportation Statistics Survey Programs. He completed his term as chair of CNSTAT in June 2010. He received a Ph.D. in statistics from Yale University.

MARILYN A. BROWN is a professor of public policy at the Georgia Institute of Technology and a member of the board of directors of the Tennessee Valley Authority. Previously, she was the interim director of the

Engineering Science and Technology Division at the Oak Ridge National Laboratory (ORNL). During her 22 years at ORNL, she researched the impacts of policies and programs aimed at advancing the market entry of sustainable energy technologies and led several energy technology and policy scenario studies. Prior to serving at ORNL, she was a tenured associate professor in the Department of Geography at the University of Illinois, Urbana–Champaign, where she conducted research on the diffusion of energy innovations. She has authored over 200 publications and has been an expert witness in hearings before committees of both the U.S. Senate and the House of Representatives. She was a member of Working Group III of the International Panel on Climate Change (IPCC) and coauthor of the 2007 Assessment Report on Mitigation of Climate Change. The IPCC was honored with the 2007 Nobel Peace Prize in part for this work. A recent study that she co-led, Scenarios for a Clean Energy Future, was the subject of two Senate hearings, has been cited in proposed federal legislation, and has had a significant role in international climate change debates. She serves on the boards of directors of several energy, engineering, and environmental organizations, including the Alliance to Save Energy and the American Council for an Energy Efficient Economy, and she serves on the editorial board of the *Journal of Technology Transfer*. She is also a member of the National Research Council's (NRC's) Board on Energy and Environmental Systems and the National Commission on Energy Policy. She served on the NRC Committee on America's Climate Choices and the NRC Panel on Limiting the Magnitude of Future Climate Change. She has a Ph.D. in geography from Ohio State University.

MICHAEL L. COHEN is a senior program officer for the Committee on National Statistics. He is currently serving as study director for the Panel on Industrial Methods for the Effective Test and Development of Defense Systems, the Panel on the Theory and Application of Reliability Growth Modeling to Defense Systems, and the Workshop on Future Directions for the National Science Foundation National Patterns of Research and Development Program. Previously, he has directed studies involving census and survey methodology and the testing and evaluation of defense systems in development, as well as a study on handling missing data in clinical trials. He was a mathematical statistician at the U.S. Energy Information Administration, an assistant professor at the School of Public Affairs at the University of Maryland, and a visiting lecturer in statistics at Princeton University. His general area of interest is the use of statistics in public policy, with particular focus on census undercount, model validation, and robust estimation. He is a fellow of the American Statistical Association and an elected member of the International Statistical Institute. He received a Ph.D. in statistics from Stanford University.

FREDERICK CONRAD is research professor in the Institute for Social Research at the University of Michigan, adjunct associate professor of psychology at the University of Michigan, and research associate professor in the Joint Program in Survey Methodology at the University of Maryland. He spent more than 10 years at the Bureau of Labor Statistics as a senior research psychologist, where he twice received the U.S. Department of Labor Secretary's Exceptional Achievement Award. His research involves new data collection methods such as interactive web surveys and virtual interviewers; interviewer–respondent interaction, including the effects of spoken language on outcomes of survey invitation interactions; reliability of pretesting techniques; and the role of public events in personal memory. He has received National Science Foundation grants for research on responding to surveys on mobile, multimodal devices; animated agents in self-administered surveys; adaptive interfaces for collecting survey data from users; costs and benefits of conversational interviews; and the usability of electronic voting systems. He served on the National Research Council's panel to review the U.S. Department of Agriculture's Agricultural Resource Management Survey. He served as associate editor of the *Journal of Official Statistics* from 2002 to 2011 and coedited a special issue of *Applied Cognitive Psychology* on cognitive aspects of survey methodology. He is a past member of the editorial board of *Public Opinion Quarterly.* He has a Ph.D. in cognitive psychology from the University of Chicago.

DONALD A. DILLMAN is regents professor in the Department of Sociology at Washington State University. He also serves as deputy director for research and development in the Washington State University Social and Economic Sciences Research Center (SESRC), and he founded the SESRC's Public Opinion Laboratory, one of the first university-based telephone survey laboratories in the United States. He served as a senior survey methodologist in the Office of the Director of the U.S. Bureau of the Census, leading to the development of new questionnaire designs and procedures for the 2000 decennial census and other government surveys. He has written or co-written 13 books and has 245 other publications. He has had a career-long emphasis on research to improve survey methodologies and apply them to the conduct of high-quality surveys. His research has also emphasized the development and use of new technologies, including the impact of information technologies on rural people and organizations. He has received many awards, including the Distinguished Rural Sociologist Award for Career Achievement from the Rural Sociological Society, the Exceptionally Distinguished Achievement Award from the American Association for Public Opinion Research, the Lester F. Ward Distinguished Contributions to Applied Sociology Award from the Society for Applied Sociology, and the Roger Herriot Award for Innovation in Federal Statistics from the American Statistical Association and the Washington Statistical Society. He is a past president of the American Association of Public Opinion Research. He served on the National Research Council panel on census residence rules and on the Survey of Earned Doctorates advisory panel. He served on the Federal Economic Statistics Advisory Committee providing advice to the Bureau of Labor Statistics, the Bureau of Economic Analysis, and the Census Bureau. He has a Ph.D. in sociology from Iowa State University.

DWIGHT K. FRENCH retired as director of the Energy Consumption Division, Office of Energy Markets and End Use, U.S. Energy Information Administration (EIA), U.S. Department of Energy, in March 2007. The Energy Consumption Division is responsible for planning and conducting the Residential Energy Consumption Survey, the Commercial Buildings Energy Consumption Survey, and the Manufacturing Energy Consumption Survey. He was responsible for the overall management of the division's energy consumption information programs and research studies as well as for providing technical input into these programs and studies. He was recognized as EIA's foremost technical expert on energy use. He served on many EIA cross-organizational teams tasked with improving the operations and data quality of information systems across EIA as well as the technical and management quality of the organization. He authored research reports, papers presented at professional meetings, and a summary history of the Energy Consumption Program at EIA. Previously, he had also served as the chief mathematical statistician of the division, responsible for review of and technical input into statistical methods for information and research programs and developing, adapting, and applying statistical techniques related to overall survey conceptual planning, complex sample design, parameter and variance estimation, statistical analysis and reporting, and quality control. Prior to joining EIA, he was a mathematical statistician at the National

Center for Health Statistics. He has an M.A. with an emphasis in statistics from the Ohio State University.

JACK G. GAMBINO is director of the Household Survey Methods Division at Statistics Canada. He manages a methodology division responsible for the design, development, implementation, and maintenance of almost all of Statistics Canada's household surveys. He has provided guidance and advice to methodologists and client program managers on methodological issues, including research and development of innovative methods; and he chaired the Methodology Research and Development Committee, which was responsible for overseeing research in the methodology branch. He participated in two redesigns of the Canadian labor force survey, proposing innovations to improve the survey design and planning studies to evaluate these proposals. He is a member of the American Statistical Association and the Statistical Society of Canada and is an elected member of the International Statistical Institute. He holds a Ph.D. in statistics from the University of Toronto.

CLARK W. GELLINGS is a fellow at the Electric Power Research Institute (EPRI). He has been at EPRI since 1982; prior to that, he was with the Public Service Enterprise Group in New Jersey. He is both an electrical and mechanical engineer with a strong background in the development of new products and services for the energy industry, especially those applied to the power industry. He has many accomplishments in developing systems for demand-side management and optimal and cost-effective utility management and in applying digital technology in the power sector in order to gain efficiencies in generation, dispatching, and end use. He is a member of numerous professional associations and has received many prizes for his work over the years. He has authored or coauthored more than 400 articles or papers and 10 books. He has an M.S. in mechanical engineering from the New Jersey Institute of Technology, an M.S. in management science from Stevens Institute of Technology, and a B.S. in electrical engineering from the Newark College of Engineering.

JANE F. GENTLEMAN is director of the Division of Health Interview Statistics at the National Center for Health Statistics (NCHS) of the Centers for Disease Control and Prevention. She was honored as the first recipient of the University of Alabama at Birmingham's Janet L. Norwood Award for Outstanding Achievement by a Woman in the Statistical Sciences in 2002.

She was honored by the University of Waterloo's Faculty of Mathematics with an achievement medal as one of the top graduates in 2005. Prior to coming to the NCHS in 1999, she was a senior research statistician, then chief of the health status and vital statistics section, and finally the assistant director of analytic methods at Statistics Canada. She taught statistics at the University of Waterloo prior to moving to Statistics Canada. She has a Ph.D. in mathematics from the University of Waterloo.

DAVID G. HUNGERFORD is the special advisor to Commissioner Anthony Eggert and former special advisor to Commissioner Arthur Rosenfeld at the California Energy Commission. His areas of responsibility include energy efficiency, alternative transportation fuels, smart grid, and demand forecasting. He most recently served as the energy commission's lead staff member on demand-response policy development. His professional career has focused on conducting and overseeing evaluation research of energy efficiency and demand-response programs and using those results to analyze the impacts of policy change for the purpose of developing and guiding policy initiatives. He has also served on numerous technical advisory committees for investor-owned utility programs and public interest energy research projects. His professional focus is in energy policy analysis, and his research interests are in technology and society issues, technology adoption, consumer behavior, and social change applied to the problem of energy consumption. He received his Ph.D. in 2003 from the University of California, Davis, in human ecology.

NANCY J. KIRKENDALL joined the staff of the Committee on National Statistics as a senior program officer in May 2009. Currently, she is study director for the Panel on Using the ACS to Estimate the Percent of Children Eligible for the School Meals Program. She served as director of the Statistics and Methods Group, U.S. Energy Information Administration (EIA), and as a member of EIA's senior staff from 2002 to 2008. She has held a variety of other positions in EIA. From 1996 to 1999, she served as senior mathematical statistician in the Statistical Policy Branch, Office of Information and Regulatory Affairs, Office of Management and Budget. There, she also served as the desk officer for the U.S. Bureau of Census, chaired the Federal Committee on Statistical Methodology, and led a variety of interagency activities. She taught part-time at the George Washington University in the Statistics Department from 1978 to 1996 and in the Engineering Management and Systems Engineering Department from 1996 to 2002. She is a past vice president of the American Statistical Association and a past president of the Washington Statistical Society. She received the Roger Herriot Award for Innovation in Federal Statistics in 2007 and the American Statistical Association's Founder's Award in 2001, and was awarded the status of fellow of the American Statistical Association in 1993. She is a member of the American Statistical Association's Committee on Energy Statistics and the Accreditation Implementation Committee. She holds a Ph.D. in mathematical statistics from George Washington University.

NINA S.-N. LAM is a professor and former chair of the Department of Environmental Sciences at Louisiana State University (LSU). She is also adjunct professor in the Department of Oceanography and Coastal Science and the Department of Geography and Anthropology at LSU. She was a program director of the Geography and Regional Science Program at the National Science Foundation and the president of the University Consortium on Geographic Information Science. She is an expert in geographic information science, remote sensing, spatial analysis, and environmental health. She has published on topics, such as spatial interpolation, fractals in geography, cancer mortality and spatial data mining, HIV/AIDS in America, and, lately, on modeling business return in New Orleans after Hurricane Katrina and on measuring community resilience. She has served on numerous national and international advisory panels and journal editorial boards, including review panels and committees for the National Research Council, the National Science Foundation, the National Institutes of Health, and the National Aeronautics and Space Administration. She has received several awards, including an Outstanding Contributions in Remote Sensing Award from the Association of American Geographers, an LSU Distinguished Faculty Award in 2006, an LSU Rainmaker Award in 2008, an LSU Distinguished Research Master Award in 2009, and an Outstanding Faculty Research Award by the LSU School of the Coast and Environment in 2011. She received her Ph.D. in geography from the University of Western Ontario, Canada.

KRISZTINA MARTON (*Study Director*) is senior program officer with the Committee on National Statistics. She is currently serving as study director for the Panel on Measuring the Group Quarters Population in the American Community Survey. She previously served as the study director of the Workshop on the Future of Federal Household Surveys. Previously, she was a survey researcher at Mathematica Policy Research (MPR), where she conducted methodological research and oversaw data collections for the National Science Foundation, the U.S. Department of Health and Human Services, the Agency for Healthcare Research and Quality, the Robert Wood Johnson Foundation, and other clients. Prior to joining MPR, she was a survey director in the Ohio State University Center for Survey Research. She has a Ph.D. in communications with an interdisciplinary specialization in survey research from the Ohio State University.

ALAN K. MEIER is a senior scientist and principal investigator at Lawrence Berkeley National Laboratory (LBNL), where he leads the Technology, Energy, Markets and Analysis group. He is also associate director of the University of California, Davis, Energy Efficiency Center. He spent four years as senior advisor in energy efficiency at the International Energy Agency in Paris. His research has focused on understanding how people (and machines) use energy and the opportunities to conserve. His research on standby power use in appliances—1 percent of global CO_2 emissions—led him to propose an international plan to reduce standby power in all devices to less than 1 watt, which has now been endorsed by the G8 countries. Other research topics include energy use of consumer electronics, energy test procedures, indicators of performance for commercial buildings, and international policies to promote energy efficiency. He has developed techniques for the collection, analysis, and display of regional electricity supply and consumption in real time. Previously, he supervised all activities of the Building Energy Analysis Group at LBNL. He is the executive editor of the magazine *Home Energy* and past editor of the journal *Energy and Buildings*. He is the author of many articles and two books, *Supplying Energy through Greater Efficiency* and *Saving Electricity in a Hurry*. He earned his Ph.D. in energy and resources from the University of California, Berkeley.

MICHAEL M. MEYER works for Google, Inc., and is based in Seattle, Washington. Previously he was cofounder and chief scientist at Intelligent Results, Inc., a software company specializing in collecting raw data from interviews and developing software to analyze large databases for the purpose of predicting behavior with regard to debt collection and the financing industry. He also held mathematics and engineering analyst positions at the Boeing Company and at Amazon.com. He served as a senior research scientist in the Department of Statistics at Carnegie Mellon University, where he was the director of special projects in computing services and the director of the applications software group. He held an academic appointment in the Department of Statistics at the University of Wisconsin–Madison. He served on the National Research Council Committee to Assess the Feasibility, Accuracy, and Technical Capability of a National Ballistics Database and the Panel on the Research on Future Census Methods. In 1991–1992, on a part-time basis, he served as the study director for the Panel to Review Evaluation Studies of Bilingual Education. He is a fellow of the American Statistical Association, in part for his role in the development and maintenance of the Statlib archive of statistical software and data resources. He received a Ph.D. in statistics from the University of Minnesota.

MAXINE L. SAVITZ is retired general manager of Technology Partnerships at Honeywell, Inc., and has more than 35 years of experience managing research, development, and implementation programs for the public and private sectors, including in the aerospace, transportation, and industrial sectors. From 1979 to 1983, she served as deputy assistant secretary for conservation in the U.S. Department of Energy. She currently serves as vice president of the National Academy of Engineering. She serves on advisory bodies for the Sandia National Laboratory and Pacific Northwest National Laboratory and is a member of the board of directors of the American Council for an Energy Efficient Economy. She served on the National Academies' Committee on America's Energy Future and was vice chair of the Panel on Energy Efficiency. She was recently appointed to the President's Council of Advisors for Science and Technology. She has a Ph.D. in chemistry from the Massachusetts Institute of Technology.

COMMITTEE ON NATIONAL STATISTICS

The Committee on National Statistics (CNSTAT) 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.

BOARD ON ENERGY AND ENVIRONMENTAL SYSTEMS

The National Academies Board on Energy and Environmental Systems conducts a program of studies and other activities to provide independent advice to the executive and legislative branches of government and the private sector on issues in energy and environmental technology, and related public policy. The Board directs expert attention to (1) energy supply and demand technologies and systems, including resource extraction through mining and drilling, energy conversion, distribution and delivery, and efficiency of use; (2) environmental consequences of energy-related activities; (3) environmental systems and controls in areas related to fuels production, energy conversion, transmission and use; and (4) related issues in national security and defense. In its work, the Board mobilizes a wide range of expertise in engineering and the physical and social sciences. In pursuit of its goals, the Board develops strategic plans, meets with sponsors and other interested organizations to exchange ideas and information, and provides stewardship of sponsored activities involving studies, briefings, workshops, symposia, and a variety of information dissemination activities.