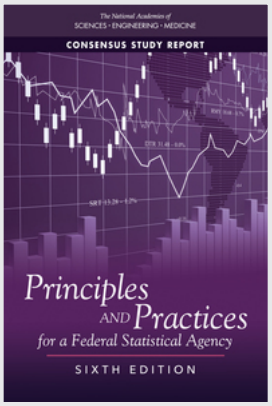


Principles and Practices for a Federal Statistical Agency: Sixth Edition

DETAILS

156 pages | 6 x 9 | PAPERBACK
ISBN 978-0-309-46167-2 | DOI 10.17226/24810

CONTRIBUTORS

Constance F. Citro, Editor; Committee on National Statistics; Division of Behavioral and Social Sciences and Education; National Academies of Sciences, Engineering, and Medicine

Principles and Practices for a Federal Statistical Agency

Sixth Edition

Committee on National Statistics

Constance F. Citro, *Editor*

Division of Behavioral and Social Sciences and Education

A Consensus Study Report of
The National Academies of
SCIENCES • ENGINEERING • MEDICINE

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu

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

This activity was supported by Grant No. SES-1024012 between the National Academy of Sciences and the National Science Foundation, which provides support for the work of the Committee on National Statistics from a consortium of federal agencies. Also supporting the Committee's work are a cooperative agreement from the U.S. Department of Agriculture, to which several agencies contribute, and individual agreements with agencies in the U.S. Departments of Commerce, Health and Human Services, Housing and Urban Development, and Treasury. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of any organization or agency that provided support for the project.

International Standard Book Number-13: 978-0-309-46167-2
International Standard Book Number-10: 0-309-46167-7
Digital Object Identifier: https://doi.org/10.17226/24810

Additional copies of this publication 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 Academies of Sciences, Engineering, and Medicine. (2017). *Principles and Practices for a Federal Statistical Agency, Sixth Edition.* Washington, DC: The National Academies Press. https://doi.org/10.17226/24810.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Marcia McNutt is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The National Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at www.nationalacademies.org.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Consensus Study Reports published by the National Academies of Sciences, Engineering, and Medicine document the evidence-based consensus on the study's statement of task by an authoring committee of experts. Reports typically include findings, conclusions, and recommendations based on information gathered by the committee and the committee's deliberations. Each report has been subjected to a rigorous and independent peer-review process and it represents the position of the National Academies on the statement of task.

Proceedings published by the National Academies of Sciences, Engineering, and Medicine chronicle the presentations and discussions at a workshop, symposium, or other event convened by the National Academies. The statements and opinions contained in proceedings are those of the participants and are not endorsed by other participants, the planning committee, or the National Academies.

For information about other products and activities of the National Academies, please visit nationalacademies.org/whatwedo.

COMMITTEE ON NATIONAL STATISTICS

Dedication

THIS SIXTH EDITION OF *Principles and Practices for a Federal Statistical Agency* (known as *P&P*) is dedicated to Miron L. Straf, third director (1987–1999) of the Committee on National Statistics (CNSTAT), who originated the concept for this report. He coedited the first (1992) edition with the late Margaret E. Martin, coedited the second, third, and fourth editions with Martin and Constance F. Citro, and coedited the fifth edition with Citro.

Straf joined CNSTAT in 1977 and served as its research director before becoming its director. He contributed to CNSTAT studies on environmental monitoring, sharing research data, cognitive aspects of survey methodology (for which he received the Innovators Award of the American Association for Public Opinion Research), and statistical assessments as evidence in the courts. During his time with CNSTAT, he developed over 50 major studies and over 40 conferences on the application of statistics to public policy. He has served the statistics and social science professions in many other ways, including as president of the American Statistical Association (ASA). He is a fellow of the ASA and the Royal Statistical Society and an elected member of the International Statistical Institute.

Following his directorship of CNSTAT, Straf worked at the National Science Foundation to develop the research priorities for the social, behavioral, and economic sciences. He then served as deputy director of the Division of Behavioral and Social Sciences and Education at the National Research Council, where his projects included studies that produced the reports *Using Science as Evidence in Public Policy* and *Furthering America's Research Enterprise*. He is a visiting scholar in the Social and Decision Analytics Laboratory of the Virginia Bioinformatics Institute of Virginia Tech.

Straf's role in moving the concept for *P&P* to fruition cannot be overstated. It is with deep gratitude that the Committee on National Statistics dedicates this sixth edition to Miron L. Straf.

Preface

The Committee on National Statistics (CNSTAT) is a standing unit of the National Academies of Sciences, Engineering, and Medicine, established in 1972 to provide an independent, objective resource for evaluating and improving the work of the decentralized federal statistical system. Under the terms of the 1863 congressional charter to the National Academy of Sciences to provide advice to the government on scientific and technical matters, CNSTAT has assessed a wide range of statistical methods and data sources for information that the public and policy makers need. It has evaluated numerous federal statistical programs, such as censuses and surveys, and addressed a range of statistical policy issues, such as the independence required for statistical agencies to be credible with policy makers, data users and the public, regardless of persuasion or party.

From time to time in its early years, CNSTAT was asked for advice on what constitutes an effective federal statistical agency. When consulted by congressional staff on the matter in the late 1980s and early 1990s during legislative debates over the (unsuccessful) establishment of a Bureau of Environmental Statistics and the (successful) establishment of a Bureau of Transportation Statistics, CNSTAT decided to prepare a document of high-level guidance. The result, *Principles and Practices for a Federal Statistical Agency*, was published in 1992. It defined and discussed reasons for the establishment of a statistical agency, identified three fundamental principles for an effective statistical agency (relevance to policy, credibility with data users, and trust of data providers), and identified 11 practices to enable a statistical agency to operationalize and adhere to the principles. A separate "commentary" section discussed each principle and practice in greater detail, drawing on relevant CNSTAT reports for illustration.

When it became clear not only that the 1992 document served a useful purpose for the federal statistical community, but also that it needed to

be updated to respond to changes in the political, economic, social, and technological environment for statistical agency work, CNSTAT decided to prepare a new edition with additional examples from its reports. The second edition was released in 2001, and subsequent editions have been released every 4 years to be available to new appointees and others at the beginning of a presidential term of office. Officials in the various cabinet departments and independent agencies that house federal statistical agencies are not always cognizant about what is proper for these agencies to be credible sources of objective, relevant, accurate, and timely statistics. The document is designed to assist them, as well as the statistical agencies' leadership and staff, to be fully aware of the standards and ideals that are fundamental to the agencies' work.

Each of the second–sixth editions has included the three original principles. The fourth edition, in light of some threats to independence, elevated statistical agency independence from a practice to a fourth principle. In addition to the four principles, this sixth edition has 13 practices, most of which were in the original 1992 edition; a few others have been added or reworded in subsequent editions based on the conclusions and recommendations in CNSTAT study reports.

This sixth edition has a new format: Part I defines and identifies reasons to establish a statistical agency, briefly relates the U.S. statistical system's evolution over time, and discusses the value of federal statistics; Part II explicates and comments on each principle; and Part III does the same for each practice. Two appendixes provide updated information on legislation and regulations that govern federal statistics and the organization of the federal statistical system.

CNSTAT thanks the many people who contributed their time and expertise to the preparation of this report. We are most appreciative of their cooperation and assistance.

We are particularly grateful to the CNSTAT staff, including director Constance Citro, along with Daniel Cork, Brian Harris-Kojetin, and Eileen LeFurgy. This edition, like its predecessors, benefited from the editing of Eugenia Grohman of the Division of Behavioral and Social Sciences and Education. In addition, we owe to her the concept for reorganizing the sixth edition to be more accessible to readers. We are also indebted to many others who offered valuable comments and suggestions, too numerous to mention.

This Consensus Study Report was reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise. The purpose of this independent review is to provide candid and critical comments that will assist the National Academies of Sciences, Engineering, and Medicine in making each published report as sound as possible and to ensure that it meets the institutional standards for quality, objectivity, evidence, and

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

We thank the following individuals for their review of this report: Mary Jo Hoeksema, Director, Government Affairs, Population Association of America/Association of Population Centers; Felice J. Levine, Executive Director, American Educational Research Association; James P. Lynch, Professor and Chair, Criminology and Criminal Justice Department, University of Maryland; Miron L. Straf, Social and Decision Analytics Laboratory, Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, National Capital Region, Arlington, VA; Roger Tourangeau, Vice President and Associate Director, Westat, Inc.; and Katherine K. Wallman, Chief Statistician (retired), Statistical and Science Policy Office, U.S. Office of Management and Budget.

Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations of this report nor did they see the final draft before its release. The review of this report was overseen by Daniel Kasprzyk, Senior Fellow, NORC at the University of Chicago. He was responsible for making certain that an independent examination of this report was carried out in accordance with the standards of the National Academies and that all review comments were carefully considered. Responsibility for the final content rests entirely with the authoring committee and the National Academies.

Finally, we thank the many federal agencies that support the Committee on National Statistics directly and through a grant from the National Science Foundation, a cooperative agreement from the National Agricultural Statistics Service, and several individual contracts. 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.

Lawrence D. Brown, *Chair*
Committee on National Statistics

Contents

List of Figures and Boxes

FIGURES

BOXES

Executive Summary

Principles and Practices for a Federal Statistical Agency, from its first edition in 1992 to this sixth edition, is intended to support the invaluable role of widely available, trustworthy, relevant, accurate, and timely government statistics. Such statistics are essential not only for policy makers and program administrators at all government levels, but also for individuals, households, businesses, and other organizations to make informed decisions and for scientists to add to knowledge. Even more broadly, the effective operation of a democratic system of government depends on the unhindered flow of impartial, scientifically based statistical information to its citizens on a wide range of issues, including employment, growth in the economy, the cost of living, crime victimization, family structure, physical and mental health, educational attainment, energy use, and the environment.

In the United States, federal statistical agencies, such as the National Center for Health Statistics and the U.S. Census Bureau, are the entities whose principal function is to compile, analyze, and disseminate information for such statistical uses as monitoring key economic and societal indicators, allocating legislative seats and government funds, deciding on the location of services, evaluating programs, and conducting scientific research. Statistical uses encompass only descriptions of groups; they exclude interest in or identification of any individual person, institution, or economic unit. To facilitate understanding, statistical agencies may analyze their data to describe trends, make comparisons, and evaluate data quality, but never to advocate policies or to take partisan positions. The work of these agencies is coordinated by the U.S. Office of Management and Budget.

In today's data-rich environment, in which no single source can provide all of the needed information with acceptable quality, costs, and timeliness, statistical agencies need to draw on multiple data sources, carefully assessing

each source's quality and relevance for statistical use. Such sources include not only surveys or censuses of people and organizations, but also government administrative records, private-sector datasets, sensor data, geospatial data, and a wide variety of web-based data.

Statistical agency decisions—managerial, programmatic, and technical—are guided by four well-established and fundamental principles:

- **Principle 1: Relevance to Policy Issues** A statistical agency must provide objective, accurate, and timely information that is relevant to issues of public policy and useful to a broad range of public- and private-sector users as well as to the public. To establish priorities for its programs, a statistical agency must not only work closely with the executive branch and Congress, but also engage users from state and local governments, businesses, academia, and nongovernmental groups. Interaction with stakeholders is essential to enable a statistical agency to continually reassess the information needs of its users.
- **Principle 2: Credibility among Data Users** Because few data users are in a position to verify the completeness and accuracy of statistical information, they must rely on an agency's reputation as a source of quality work that is free from political and other undue external influence. An agency builds credibility by widely disseminating its data in accessible and easy-to-use ways on an equal basis to all users. Also essential to building credibility are for an agency to be open and transparent about its data sources and their limitations, demonstrate understanding of users' needs and priorities, fully document the processes used to produce and disseminate statistical products, and take proactive steps to preserve data for future use.
- **Principle 3: Trust among Data Providers** Data providers, such as survey respondents or custodians of administrative records, must be able to trust that the information they provide to a statistical agency will be used only for the purposes that the agency has described. In particular, data providers must trust that information collected from them for statistical purposes will not be used for any other purpose, such as law enforcement or regulation, that could affect a specific individual or organization. Data providers must also be able to trust that a statistical agency will honor its pledges to protect the confidentiality of their responses and safeguard individual records against unauthorized attempts to gain access. A statistical agency further earns the trust of data providers and respects their privacy and autonomy through minimizing the intrusiveness of questions and the time and effort needed to respond to information requests, and through informing providers of the importance of the information,

the expected time required to participate in the data collection, and whether the collection is mandatory or voluntary.

- **Principle 4: Independence from Political and Other Undue External Influence** To be credible, trustworthy, and unhindered in its mission, a statistical agency must maintain a position of independence from undue external influences (even as it proactively seeks input on its program and priorities). It must avoid even the appearance that its collection, analysis, or dissemination processes might be manipulated for political or partisan purposes or that individually identifiable data collected under a pledge of confidentiality might be turned over for administrative, regulatory, or law enforcement uses. Protection from undue outside influences requires that a statistical agency have authority for professional decisions on its programs, including authority over the selection and promotion of staff; the processing, secure storage, and maintenance of data; and the timing and content of data releases, accompanying press releases, and documentation. The credibility that comes from independence is essential for users to maintain confidence in the accuracy and objectivity of a statistical agency's data and for data providers to be willing to cooperate with agency requests.

Thirteen practices are critical for the effective, principled operation of a statistical agency:

1. a clearly defined and well-accepted mission,
2. necessary authority to protect independence,
3. use of multiple data sources for statistics that meet user needs,
4. openness about sources and limitations of the data provided,
5. wide dissemination of accessible and easy-to-use data,
6. cooperation with data users,
7. respect for the privacy and autonomy of data providers,
8. protection of the confidentiality of data providers' information,
9. commitment to quality and professional standards of practice,
10. an active research program,
11. professional advancement of staff,
12. a strong internal and external evaluation program, and
13. coordination and collaboration with other statistical agencies.

By adhering to the principles and following the practices, a federal statistical agency will be well positioned to provide the relevant, accurate, timely, credible, and trustworthy statistical information that policy makers and the public require.

Part I

Introduction

Background

Federal statistics have been integral to the growth and development of the United States since the adoption of the U.S. Constitution: Article 1, Section 2, mandates a decennial census to provide population counts for periodic reallocation of seats in the House of Representatives. Today, federal statistics undergird the allocation of billions of dollars of federal funding to states, localities, and other entities, drive financial markets and decision making, contribute to policy debates in many areas of public interest, and support valuable scientific research.

Historically, in the United States, as particular policy concerns have become salient at the federal level, Congress has established statistical agencies to provide the information needed to address them. As a consequence, the federal government has a large number of statistical agencies, as well as agencies with statistical programs, and a statistical coordination function in the U.S. Office of Management and Budget (OMB). The principles and practices for a federal statistical agency that are the subject of this report pertain to individual statistical agencies as separate entities in a decentralized system for providing federal statistics.[1]

This report is the sixth edition of *Principles and Practices for a Federal Statistical Agency*, first published in 1992 by the Committee on National Statistics (CNSTAT) at the National Academies of Sciences, Engineering, and Medicine. The publication draws on CNSTAT's many studies of specific agencies, programs, and topics. It has proved helpful to Congress, OMB, federal statistical agencies and others about what constitutes an effective

[1]This report neither comments on the advantages or disadvantages of the decentralized U.S. system nor compares it with other models for organizing government statistics (see Norwood, 1995, for a comparison). "Statistical agencies" refers to the nation's 13 principal statistical agencies: see "Definition of a Federal Statistical Agency" below.

and credible statistics entity.[2] Beginning with the second edition in 2001, CNSTAT has updated the document every 4 years to provide a current edition to newly appointed cabinet secretaries and other personnel at the beginning of each presidential administration or second term.

The Committee on National Statistics had the following statement of work for this sixth edition:

> In response to recurring requests for advice on what constitutes an effective federal statistical agency, CNSTAT issued the first edition of *Principles and Practices for a Federal Statistical Agency* (*P&P*) in 1992. In early 2001, 2005, 2009, and 2013, CNSTAT issued the second, third, fourth, and fifth editions, respectively, which reiterated the basic principles for federal statistical agencies, revised and expanded the discussion of some of the practices for an effective statistical agency, and updated the discussion with references to recent reports by CNSTAT and others. Changes in laws, regulations, and other aspects of the environment of federal statistical agencies over the past 4 years [see Appendix A] warrant preparation of a sixth edition, which a CNSTAT committee will prepare for release in early 2017.

The committee distinguishes between "principles," which are fundamental and intrinsic to the concept of a federal statistical agency, and "practices," which are ways and means of making the basic principles operational and facilitating an agency's adherence to them. Other documents that present principles, practices, or both for statistical agencies include the *Fundamental Principles of Official Statistics* (United Nations Statistical Commission, 2014), first promulgated by the United Nations Statistical Commission in 1994 and endorsed by the United Nations General Assembly with a revised preamble in 2014; the *European Statistics Code of Practice for the National and Community Statistical Authorities* (European Statistical System Committee, 2011), first issued in 2005 and revised in 2011; and *Statistical Policy Directive No. 1, Fundamental Responsibilities of Federal Statistical Agencies and Recognized Statistical Units* (U.S. Office of Management and Budget, 2014).[3]

This sixth edition presents and comments on four basic principles that statistical agencies must embody to carry out their mission fully:

(1) They must produce data that are relevant to policy issues.
(2) They must achieve and maintain credibility among data users.
(3) They must achieve and maintain trust among data providers.
(4) They must achieve and maintain independence from the appearance and reality of political or other undue external influence in developing, producing, and disseminating statistics.

[2] See, e.g., U.S. Office of Management and Budget (2007, 2014); U.S. General Accounting Office (1995); U.S. Government Accountability Office (2007, 2012); see also Citro (2014b).

[3] The OMB directive cites the committee's fifth edition (National Research Council, 2013b).

This edition also discusses 13 important practices that are the means for statistical agencies to implement the four principles. These practices include authority for an agency to protect its independence, use of multiple data sources for statistics that meet user needs, openness about sources and limitations of the data provided, respect for privacy and autonomy of data providers, protection of the confidentiality of providers' information, a commitment to quality and professional standards of practice, a strong internal and external evaluation program, and coordination and collaboration with other statistical agencies.

Having high-quality information from statistical agencies that can address important issues as they arise is essential to inform public and private decisions, to design, monitor, and evaluate effective public policies,[4] and to support research that contributes to scientific knowledge and evidence-based policy making. To best serve these needs, this sixth edition emphasizes the importance of statistical agencies' actively seeking out new ways of using multiple data sources (such as administrative records, private-sector datasets, and other sources in addition to surveys) as the basis for their statistical products. The goal is to enable agencies to maintain and improve the credibility, relevance, accuracy, timeliness, and cost-effectiveness of their data in the face of the significant challenges they now face. Chief among these challenges are the increasing perceptions of burden and unwillingness to respond to traditional survey inquiries by people and organizations, the correspondingly higher costs of obtaining adequate responses to survey inquiries, and constrained budgets for statistical activities (see Citro, 2014a; Groves, 2011; National Academies of Sciences, Engineering, and Medicine, 2017a; Prewitt, 2010:7–16; Singer, 2016). In facing these challenges, the growing availability of multiple data sources and advances in computing technology and methods for combining data offer many opportunities for more complete and timely statistics.

Unlike previous editions, this sixth edition presents the principles and practices in separate parts (II and III), with commentary on them included. The rest of this Part I covers four topics: the definition of a statistical agency; reasons for establishing a statistical agency or adding responsibilities to an existing agency; a brief history of the U.S. statistical system; and a discussion of the value of federal statistics. Throughout the online version of the report, hyperlinks to other sections and to references are provided wherever possible to assist the reader. The two appendixes update material included in previous editions and add new information to help orient readers; their contents are listed in this document, and they

[4]See also UNECE Task Force on the Value of Official Statistics (2017:8).

are available in full in the on-line edition.[5] Appendix A summarizes the history and current status of key legislation and regulations that affect federal statistical agencies, such as the Confidential Information Protection and Statistical Efficiency Act of 2002; OMB statistical policy directives; and Office of Science and Technology Policy memoranda on scientific integrity and public access to federally sponsored research and data. Appendix B reviews the organization of the U.S. federal statistical system, providing information about each recognized statistical agency and statistical unit and the coordinating function in OMB.

Although focused on federal statistical agencies, many of the principles and practices articulated here likely also apply to statistical activities elsewhere, such as in federal policy, evaluation, research, and program agencies, in state and local government agencies, and in other countries. The principles and practices in this report remain guidelines, not prescriptions. CNSTAT intends them to assist statistical agencies and to inform legislative and executive branch decision makers, data users, and others about the characteristics of statistical agencies that enable them to serve the public good.

[5] Available at doi:https://doi.org/10.17226/24810. The National Academies Press website (www.nap.edu) provides free access to all published reports of the National Academies of Sciences, Engineering, and Medicine in electronic formats.

Definition of a Federal Statistical Agency

A federal statistical agency is a component of the federal government whose principal function is the compilation and analysis of data and the dissemination of information for statistical purposes.[6]

THE *COMPONENT* IS RECOGNIZED as a distinct entity. It may be located within a cabinet-level department or an independent agency, or it may be an independent agency. It may be labeled a bureau, center, division, office, or other entity so long as it is recognized as a distinct entity.

Compilation may include direct collection of data from individuals, organizations, or establishments through surveys. It may also include the acquisition of data from other sources, such as administrative records maintained by government agencies to operate a program, datasets available from the private sector, or data gleaned from sensors or selected Internet websites.[7]

[6]The Confidential Information Protection and Statistical Efficiency of 2002 (P.L. 107-347, Section 502(8)) provides a similar definition of a statistical agency: "An agency or organizational unit of the executive branch whose activities are predominantly the collection, compilation, processing, or analysis of information for statistical purposes." There are 13 principal statistical agencies (the focus of this document), several "recognized statistical units" (components of nonstatistical agencies), and statistical programs, such as a survey or time series, which can be in any agency. The designation of a principal statistical agency and recognized statistical unit is not entirely consistent in legislation and guidance (see "Brief History of the U.S. Federal Statistical System," below, and Appendix B).

[7]"Data," "information," and "statistics" do not have clear definitional boundaries: this document generally uses "data" to refer to individual responses or items in a dataset and "statistics" or "information" to refer to data that have been organized, modified as necessary

Analysis may take various forms. It includes methodological research to improve the quality, usefulness, and usability of data; the use of data gathered in real time about a source (paradata) to make compilation as efficient and error-free as possible; and modeling to combine data from more than one source into useful statistics. It also includes substantive analysis, such as developing indicators from one or more data series, making projections, interpreting data, and explaining differences among statistics obtained by different methods, such as surveys and administrative records. An analysis by a statistical agency does not advocate policies or take partisan positions.

Dissemination means making information available to the public, to the executive branch, and to Congress in easily accessible and usable forms with appropriate documentation to facilitate informed use. Dissemination also means taking care to curate information so that its value and usability are maintained over time.[8]

Statistical purposes include description, evaluation, analysis, and inference for groups of individuals or other units; they do not include interest in or identification of an individual person or economic unit. A statistical agency collects data directly from providers or from other sources for statistical purposes. It does not use these data for nonstatistical purposes, such as regulation or law enforcement. It also protects the confidentiality of responses collected under a confidentiality pledge and safeguards individual records against unauthorized access.[9]

This definition of a federal statistical agency does not include many statistical activities of the federal government, such as statistics compiled by an operating agency for administrative purposes (e.g., U.S. Office of Personnel Management statistics on new hires and retirements). Nor does it include agencies whose primary functions are the conduct or support of problem-oriented research, although their research may be based on information gathered by statistical methods, and they may sponsor important surveys: examples include the National Institutes of Health, the

(e.g., to weight survey responses to national population controls), and otherwise processed for use.

[8]Data curation includes and is more than data archiving; it encompasses an explicit data preservation policy, validation of the data and documentation to be archived, procedures for migration of the data to new media, and other policies and procedures to ensure that authorized researchers and others can access, understand, and use the data over time; see Practice 5.

[9]The Federal Cybersecurity Enhancement Act of 2015 (part of P.L. 114-113) upgraded the protection of individual records against cyberattacks. Concomitantly, to respect the autonomy of respondents, statistical agencies revised their confidentiality pledges to note that respondents' records are screened to protect against cybersecurity risks; see Principle 3 and Practice 8 for a fuller discussion; see also Appendix A.

Agency for Healthcare Research and Quality, and other agencies in the U.S. Department of Health and Human Services.

This definition of a statistical agency also does not usually include agencies whose primary function is policy analysis and planning (e.g., the Office of Tax Analysis in the U.S. Department of the Treasury, the Office of the Assistant Secretary for Planning and Evaluation in the U.S. Department of Health and Human Services). Such agencies may collect and analyze statistical information, and statistical agencies, in turn, may perform some policy-related analysis (e.g., produce reports on trends in after-tax income or child care arrangements of families). However, to maintain credibility as an objective source of accurate, useful information, statistical agencies must be separate from units that are involved in developing policy and assessing policy alternatives.

Statistical agencies, as noted above, typically collect information under a pledge (to a person or organization) of confidentiality. Statistical agencies may collect information that identifies individual government agencies when the data are already public information—as, for example, in the Census Bureau's program of statistics for state and local governments (see National Research Council, 2007a) and the National Center for Science and Engineering Statistics' program to collect information on research and development spending from federal agencies (see National Research Council, 2010a).

Occasionally, statistical agencies are charged to collect information for both statistical and nonstatistical purposes. Three examples are:

- The Bureau of Transportation Statistics (BTS) maintains the Airline On-Time Statistics Program (originated by the former Civil Aeronautics Board), which identifies individual airlines.[10] However, BTS does not itself use the data for administrative or regulatory purposes—those functions are carried out by the Federal Aviation Administration—and the data are not collected under a pledge of confidentiality (see National Research Council, 1997b).
- Higher education institutions that participate in federal student aid programs are required by law (20 USC 1094(a)(17)) to respond to the Integrated Postsecondary Education Data System (IPEDS) from the National Center for Education Statistics (NCES). The data provided to IPEDS on enrollments, graduation rates, faculty and staff, finances, institutional prices, and student financial aid are not collected under a pledge of confidentiality, and NCES makes information on individual institutions available to parents and students to help them in choosing a college, as well as to researchers and others.[11]

[10]See https://www.transtats.bts.gov/ONTIME/Index.aspx [April 2017].
[11]See http://nces.ed.gov/ipeds [April 2017].

- The U.S. Census Bureau participates with a consortium of agencies, led by U.S. Customs and Border Protection, in the Automated Export System (AES), by which exporters electronically file information about shipments. The Census Bureau and other statistical agencies use information in the AES for statistical purposes and do not make the data available to others—that is, they maintain confidentiality following their own data acquisition. Independently, export enforcement agencies use data they acquire from the system to administer U.S. export laws.[12]

Statistical agencies should carefully consider the advantages and disadvantages of undertaking a program with both statistical and nonstatistical purposes. A potential advantage is that there may be improved consistency, quality, and cost-effectiveness when a statistical agency collects information for its own use and that of other agencies. A potential disadvantage is that the program may compromise the public perception of the agency as objective and separate from government administrative, regulatory, and enforcement functions.

When an agency decides to carry out a program that has both statistical and nonstatistical uses, it must clearly describe the program on such dimensions as the extent of confidentiality protection, if any (e.g., some but not all of the data may be collected under a pledge of confidentiality); the statutory basis for the program and the public purposes it serves, including benefits to respondents from having comparative information available of uniform quality; and the role of the agency (e.g., providing information to the public, working with respondents to improve reporting).[13] Should an agency decide that the nature of a program is such that no amount of description or explanation is likely to make it possible for the agency to maintain its credibility as a *statistical* agency, it should decline to carry out the activity.

[12]See https://www.cbp.gov/trade/aes [April 2017].

[13]U.S. Office of Management and Budget (2007:Sec. V) spells out the requirements for statistical agencies to inform the public and respondents regarding any programs with nonstatistical uses.

Establishment of a Federal Statistical Agency

The first principle in the *Fundamental Principles of Official Statistics* (United Nations Statistical Commission, 2014) accords world-wide recognition to the indispensable role of official statistics:

> Official statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information.

In the United States, federal statistical agencies are established to be a credible source of relevant, accurate, and timely statistics in one or more subject areas that are available to the public and policy makers.

Relevant statistics are statistics that measure things that matter to policy and decision making, program implementation, program monitoring, program evaluation, scientific research, and public understanding. Relevance requires agencies to be concerned about providing data that help users meet their current needs for decision making and analysis, as well as anticipating future needs.

Accurate statistics are statistics that capture the phenomena being measured and do so in repeated measurements. Accuracy requires proper concern for consistency across geographic areas and across time, as well as for statistical measures of errors in the data.

Timely statistics are those that are made available close in time to the phenomena they measure. Timeliness also requires concern for issuing data

as frequently—but not more frequently—as is needed to reflect important changes in what is being studied.

Credibility requires concern for both the reality and appearance of impartiality and independence from political and other undue external influence. Credibility also requires that agencies follow such practices as making their data and the information that users need to work with the data readily available in easily accessible formats to all.

It is the primary mission of agencies in the federal statistical system to strive to ensure the relevance, accuracy, timeliness, and credibility of statistical information. *Statistical Policy Directive No. 1* (U.S. Office of Management and Budget, 2014) underlines the importance of this mission; it categorically states that "Federal departments *must* [emphasis added] enable, support, and facilitate Federal statistical agencies and recognized statistical units as they implement these responsibilities."

There is no set rule or guideline for when it is appropriate to establish a separate federal statistical agency or to add responsibility for statistics on a new topic to an existing statistical agency.[14] It may be reasonable to carry on statistical activities within an operating unit of a department or independent agency instead. Establishing a federal statistical agency, or adding responsibilities to an existing agency, should be considered when one or more of the following conditions prevail.[15]

- There is a need for high-quality information on an ongoing basis beyond the capacity of existing operating units. Such needs may require coordinating data from various sources, initiating new data collection programs, or developing regularly updated time series of estimates.
- There is a need, as a matter of credibility, to ensure that major data series are independent of policy makers' control and readily available on an equal basis to all users.
- There is a need to establish the functional separation of data on individuals and organizations that are collected for statistical purposes from data on individuals and organizations that may be used for administrative, regulatory, or law enforcement uses. Such separation, recommended by the Privacy Protection Study Commission (1977), bolsters a culture and practice of respect for privacy and protection of confidentiality. Functional separation is easier to maintain when the data to be used for statistical purposes

[14] Legislation is usually required to establish a federal agency.

[15] National Research Council (2001:Ch. 6) cited such reasons in recommending that the U.S. Department of Health and Human Services establish or identify a statistical agency with responsibility for statistical functions and data collection for social welfare programs and the populations they serve (the recommendation was not adopted).

are compiled and controlled by a unit that is separate from operating units or department-wide data centers.
- There is a need to emphasize the principles and practices of an effective statistical agency, for example, professional practice, openness about the data provided, and wide dissemination of data.
- There is a need to encourage research and development of a broad range of statistics in a particular area of public interest or of government activity or responsibility.
- There is a need to consolidate compilation, analysis, and dissemination of statistics in one unit to encourage high-quality performance, eliminate duplication, and streamline operations.

Commercial, nonprofit, and academic organizations in the private sector also provide useful statistical information, including data they collect themselves and data they acquire from statistical agencies and other data collectors to which they add information or analysis. However, because the benefits of statistical information are shared widely throughout society and because it is often difficult to garner payments for these benefits, private entities are not likely to collect all of the data that are needed for public and private decision making or to make data as widely available as needed for important public purposes.[16]

Federal statistical agencies are established to ensure that a broad array of credible, scientifically based information is available that will inform the electorate and policy makers in a democratic system of government. Federal statistics serve and "should be viewed and treated as part of the nation's scientific infrastructure" (Prewitt, 2010:7).

[16]See "Valuing Federal Statistics," below.

Brief History of the U.S. Federal Statistical System

The U.S. government collected and published statistics long before any distinct federal statistical agency was formed (see Appendix B; see also Anderson, 2015; Citro, 2016; Duncan and Shelton, 1978; Norwood, 1995). The U.S. Constitution mandated a decennial census of population; the first such censuses (beginning in 1790) were conducted by U.S. marshals as one of their many duties. The Constitution also mandated reporting of federal government receipts and expenditures, which led to early collection by the U.S. Department of the Treasury of foreign trade statistics because of the reliance of the federal government on tariffs for revenues in the 19th century. A census of manufactures was first taken in conjunction with the 1810 population census, and the 1820 population census laid the groundwork for additional economic statistics by asking for the number of household members principally employed in agriculture, manufacturing, or commerce.

In the 1860s, Congress enacted laws providing for the compilation of statistics on agriculture, education, and income. It established the Bureau of Labor (forerunner of the Bureau of Labor Statistics) as a separate agency with a mandate to respond to widespread public demand for information on the conditions of industrial workers in 1884. It established the Census Bureau as a permanent agency in 1902.

Many federal statistical agencies that can trace their roots back to the 19th or early 20th century, such as the Bureau of Economic Analysis and the National Center for Health Statistics, were organized in their current form following World War II. Agencies organized since then include the Bureau of Justice Statistics, the Bureau of Transportation Statistics, the

Energy Information Administration, and the National Center for Science and Engineering Statistics.

Today, the U.S. Office of Management and Budget (OMB), through its Statistical and Science Policy Office (which has roots going back to the 1930s), coordinates the work of federal statistical agencies. The chief statistician, who heads the office, chairs the Interagency Council on Statistical Policy (ICSP), which was created by OMB in the late 1980s and authorized in statute in 1995. The current ICSP membership includes the heads of the 13 principal statistical agencies[17] and one other agency:[18]

- Bureau of Economic Analysis (U.S. Department of Commerce)
- Bureau of Justice Statistics (U.S. Department of Justice)
- Bureau of Labor Statistics (U.S. Department of Labor)
- Bureau of Transportation Statistics (U.S. Department of Transportation)
- Census Bureau (U.S. Department of Commerce)
- Economic Research Service (U.S. Department of Agriculture)
- Energy Information Administration (U.S. Department of Energy)
- National Agricultural Statistics Service (U.S. Department of Agriculture)
- National Center for Education Statistics (U.S. Department of Education)
- National Center for Health Statistics (U.S. Department of Health and Human Services)
- National Center for Science and Engineering Statistics (U.S. National Science Foundation)
- National Center for Veterans Analysis and Statistics (NCVAS) (U.S. Department of Veterans Affairs)
- Office of Research, Evaluation, and Statistics (U.S. Social Security Administration)
- Statistics of Income Division (U.S. Department of the Treasury)

Under the guidance issued in 2007 for the Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA) (see Appendixes A and B), OMB may also recognize statistical "units," which are usually

[17]See https://obamawhitehouse.archives.gov/omb/inforeg_statpolicy/bb-principal-statistical-agencies-recognized-units [April 2017].

[18]The Paperwork Reduction Act, as reauthorized in 1995 (44 USC 3504(e)(8)), authorizes the OMB director to "establish an Interagency Council on Statistical Policy to advise and assist the Director in carrying out the functions under this subsection that shall— (A) be headed by the chief statistician; and (B) consist of— (i) the heads of the major statistical programs; and (ii) representatives of other statistical agencies under rotating membership." The current ICSP rotating member is the National Center for Veterans Analysis and Statistics. For many years, the rotating member was the Environmental Protection Agency's Office of Environmental Information.

small components of nonstatistical agencies. The statistical agencies named in CIPSEA and recognized statistical units are authorized to assign agent status to researchers and contractors, which permits sharing individually identifiable information with them for statistical purposes and holding them legally liable for protecting the confidentiality of the information. The fundamental responsibilities enumerated in *Statistical Policy Directive No. 1* apply to both statistical agencies and recognized statistical units.

In its annual compilation *Statistical Programs of the United States Government* (see, e.g., U.S. Office of Management and Budget, 2017:Table 1), OMB lists over 100 agencies that are neither a principal statistical agency nor a recognized unit but that carry out statistical activities with annual budgets of $500,000 or more. The purpose of the listing is to indicate the breadth of statistical work in the federal government. Figure I.1 depicts the 13 principal statistical agencies and other agencies (including recognized statistical units) with significant statistical programs by cabinet department. Figure I.2 depicts reporting relationships for the 13 principal statistical agencies, rotating members of the ICSP, other statistical units, and the Statistical and Science Policy Office from the relevant congressional appropriations subcommittee to the cabinet secretary and any other intermediate levels of authority. Both figures make clear the decentralization of the U.S. statistical system.

The principles for federal statistical agencies presented here are relevant to recognized statistical units and other federal agencies that carry out statistical activities, and many of the detailed practices are also pertinent. Similarly, the principles and practices may be relevant to statistical units in state and local government agencies, as well as for international statistical agencies.[19]

[19] As cited in National Academies of Sciences, Engineering, and Medicine (2017b), several agencies have issued federal program evaluation guidelines that include such principles as rigor, relevance, transparency, independence, and ethics. See also the call in the President's fiscal 2017 budget for a government-wide "principles and practices for evaluation offices," modeled on *Statistical Policy Directive No. 1* (U.S. Office of Management and Budget, 2016:71).

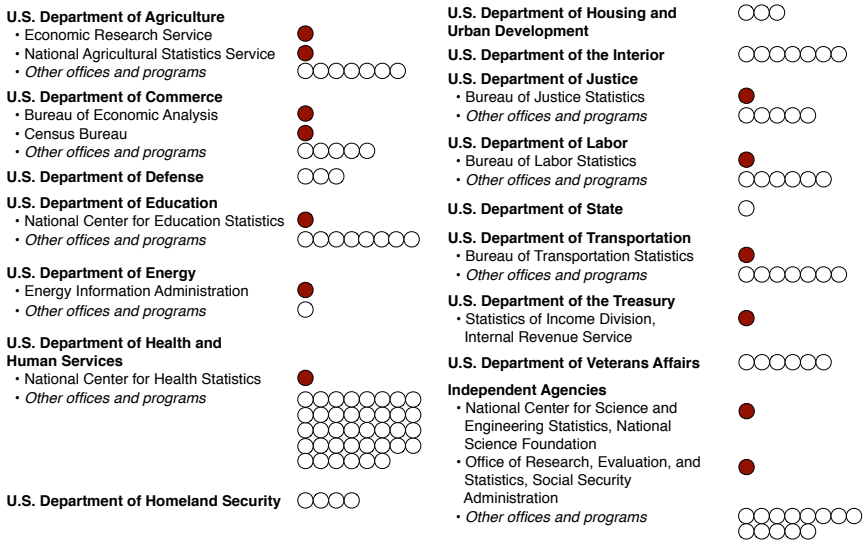
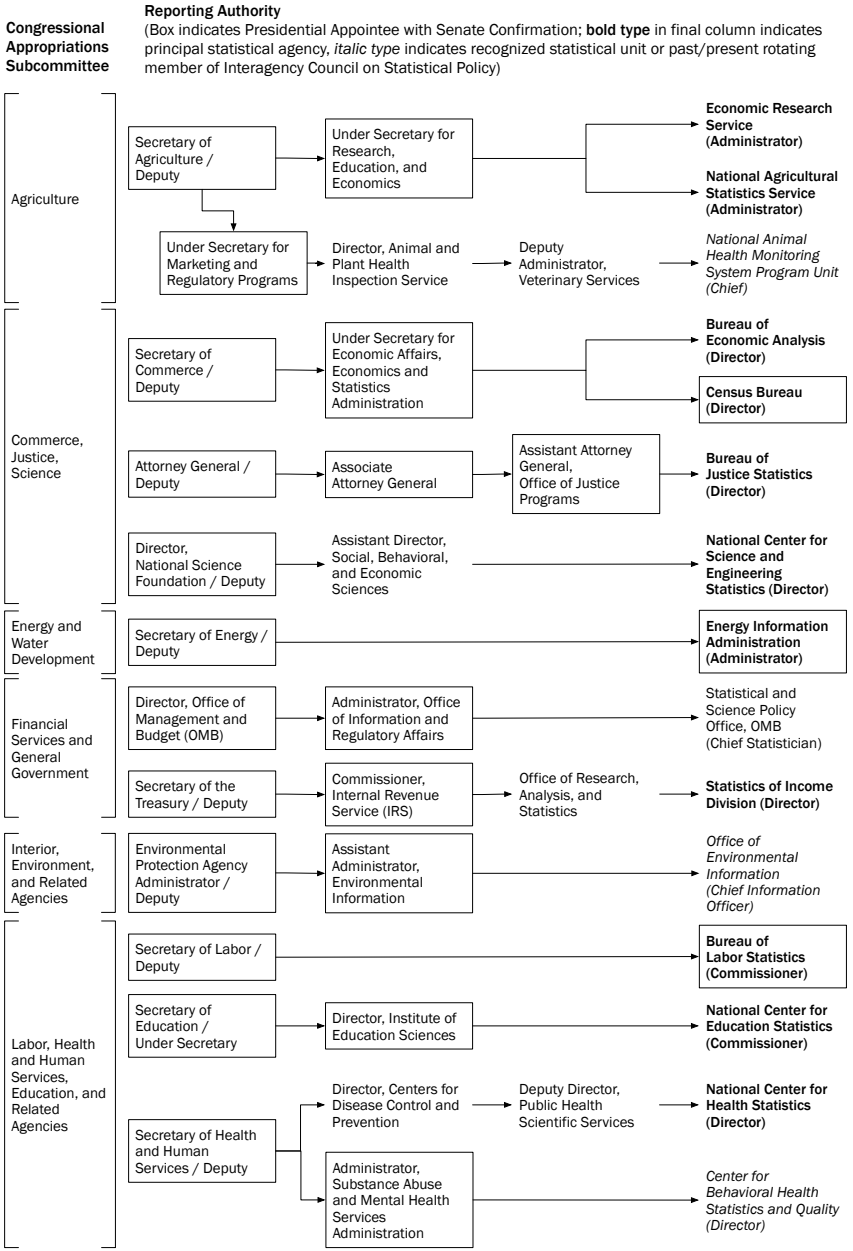


Figure I.1 Organization of principal federal statistical agencies and other statistical units and programs, by department, 2017.

NOTES: Filled-in circles are the 13 principal statistical agencies; empty circles are other statistical units and programs. See text and Appendix B for discussion.

SOURCE: U.S. Office of Management and Budget (2017:Table 1).



(Continued on next page)

Figure I.2 Principal statistical agencies, rotating ICSP members, other recognized statistical units, and the Statistical and Science Policy Office, by congressional appropriations committee and parent department/agency.

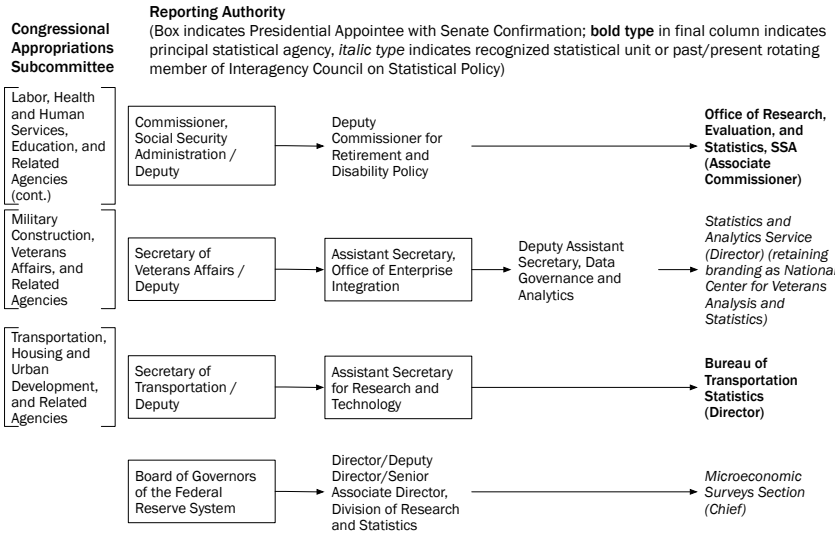


Figure I.2 (cont.)

NOTE: The National Center for Veterans Analysis and Statistics and the Office of Environmental Information are the current and previous ICSP rotating member, respectively. The organizational location and relevant congressional appropriations subcommittees are based on jurisdictions in the 115th Congress, which began on January 1, 2017.

Valuing Federal Statistics

The value of federal statistics rests on their unique character.[20] Federal statistics meet the definition of a "public good," similar to the judiciary, clean air, and national defense (see Box I.1), and are valuable for many important public purposes.

That federal statistics are a valuable public good justifies expenditure of public funds to ensure that relevant, accurate, timely, and credible statistics are collected and made available to all current and potential users. In fact, spending on statistical programs is a tiny fraction of overall federal spending: in fiscal 2016, the $6.5 billion budget request for all statistical programs identified by the U.S. Office of Management and Budget (2017:Table 1), excluding the decennial census, amounted to less than 0.2 percent of the budget authority of about $3.6 trillion for the federal government. The $6.5 billion is equal to about $20 annually for every U.S. resident (325 million as of December 1, 2016).[21]

Whether current spending on federal statistics, in total or as divided among individual statistical programs, is optimal in terms of the value of the statistics produced is not easily determined. Yet there are compelling justifications in terms of the uses of the data, which include political representation, economic decision making in the public and private sectors, administration of programs, scientific research, and contributions to public opinion and debate. After briefly reviewing these uses, the section concludes

[20] This section draws heavily on Citro (2016) and National Research Council (2013b:101–102); see also U.S. Office of Management and Budget (2017:3–4); see Prewitt (2010) for an overview of the federal statistical system and its data users and uses.

[21] The decennial census adds about $40 per resident over a 10-year period, or an additional $4 annually. Of the total of $6.5 billion, the 13 principal statistical agencies accounted for 38 percent.

with a brief consideration of the limits of a formal cost-benefit approach to assessing value.

POLITICAL REPRESENTATION

The bedrock justification for federal statistics on the population rests in the U.S. Constitution, which mandated a decennial census of population every 10 years (the first census was taken in 1790) for determining the allocation of seats in the U.S. House of Representatives among the states. At first, Congress added seats to the House each decade so that the decennial apportionment would not reduce the number of seats for any state even when its percentage share decreased. In 1929, Congress fixed the size of the House at 435 seats, so that reapportionment means that slower growing states lose numbers of seats in addition to percentage shares.

Reapportionment in turn triggers the redistricting process by which states, using census data, redraw the boundaries of congressional districts to accommodate changes in the number of seats and in the distribution of population geographically (e.g., movement from urban to suburban areas) and by race and ethnicity. Typically, data on voting patterns by party are also used in redistricting as political parties attempt to maximize the number of seats they can expect to win in elections. States and many local governments also use census and other data to reapportion and redistrict their legislative bodies. The redistricting process for congressional and local seats can become highly contentious, generating legal cases that, in some instances, reach the U.S. Supreme Court (see McMillen, 2012; National Research Council, 1995:App. C).

Legislation that affects voting participation and hence political representation and power in at least one instance hinges explicitly on federal statistics. Thus, section 203 of the Voting Rights Act mandates provision of language assistance to voters in areas that meet certain population criteria as determined by the U.S. Census Bureau. The current method uses the American Community Survey (ACS) to estimate single language minority groups with limited English speaking proficiency and educational attainment in counties and tribal areas.[22]

ECONOMIC DECISION MAKING

In the early days of the republic, James Madison expressed the view that expanding census statistics beyond the headcount and obtaining

[22]See, e.g., "Voting Rights Act Amendments of 2006, Determinations Under Section 203," 81 *Federal Register* 87532 (December 5, 2016); available: https://www.federalregister.gov/d/2016-28969 [April 2017]. The ACS asks questions formerly included in the decennial census "long-form sample."

Box I.1 Federal Statistics as a Public Good

The definition of a public good has two components (see Gravelle and Rees, 2004). First, a public good must be *nonrivalrous*—that is, when one individual consumes it, there is no actual or potential reduction in the amount available for another individual to consume. A public good must also be *nonexcludable*—that is, it must be difficult for a business or other private entity to try to establish a market for such a good that is open only to those willing to pay the price. Federal statistics, similar to other public goods, such as the judiciary and national defense, satisfies the first component and practically speaking satisfies the second (see discussion in National Research Council, 1999:Ch. 2).

In regard to the second component, there are many examples of privately marketed statistical information, based on surveys or other data sources and including information that adds value to an underlying federal data series. Yet the private sector (whether a business, academic institution, or other organization) does not typically attempt to provide statistics on the scale of critically important federal surveys, censuses, and time series. For example, there has never been a private-sector initiative equal to the Current Population Survey (CPS), which is the source of official monthly unemployment estimates and has a sample size of about 60,000 households (see https://www.census.gov/programs-surveys/cps.html [April 2017]). The costs of federal statistical programs for data collection, processing, quality review, documentation, and dissemination would likely be too high (e.g., in the case of a major nationwide federal household or business survey like the CPS), or the potentially interested market of data users too small (e.g., in the case of longitudinal surveys for scientific research), or both, to make for a viable business proposition. Business reasons may also lead a private-sector data provider to modify or abandon useful statistical series.

Private-sector series often depend on federal statistics in various ways. For example, private-sector price indexes based on web-scraping (e.g., the MIT Billion Prices Index, see http://bpp.mit.edu [April 2017]) use federal Consumer Price Index (CPI) information for market-basket weights and benchmarking, and public opinion polls use federal statistics on gender, age, ethnicity, and other characteristics to adjust the raw data to represent population groups.

Thus, federal statistics not only are useful and often indispensable themselves, but also are a necessary ingredient to many other data products and services that fulfill information needs. Paraphrasing a sentence in Abraham Lincoln's 1858 "House Divided" speech, without government-funded federal statistics, the nation would not know where it has been or where it is tending and would be ill-positioned to judge what to do and how to do it.

detailed information on the economy as well as the population would be of great benefit for legislative deliberations (Eberstadt et al., 2017:1–2). From these beginnings, the scope of federal economic statistics, broadly defined, expanded enormously as described in the "Brief History of the U.S. Federal Statistical System" section above.

For some economic statistics, it is possible to assess their value by comparing the dollars spent on them to the dollars they drive in the economy and society. The federal government currently labels 38 statistics—e.g., gross domestic product (GDP), unemployment, monthly retail trade, weekly natural gas storage, grain production, money stock, and others—as "principal federal economic indicators." *Statistical Policy Directive No. 3* (U.S. Office of Management and Budget, 1985), issued in the 1970s in response to attempted manipulation of Bureau of Labor Statistics monthly unemployment statistics (see Norwood, 2016) and strengthened in 1985, requires these indicators to be published by the cognizant statistical agency on specified release dates under procedures designed to protect the integrity and credibility of the estimates (see Practice 2).

An example of the consequential effects of these key statistics is provided by the CPI. It determines annual cost-of-living adjustments to Social Security monthly benefits, which in April 2017 amounted to $76.8 billion provided to 61.4 million people.[23] Consequently, a 1 percentage point difference in the CPI would amount to $768 million in additional or reduced Social Security benefits on a monthly basis, or about $9 billion annually. Annual changes in the CPI also affect commercial and residential rents, public- and private-sector wages, and components of the federal income tax code. Monthly changes in the CPI are a major input to Federal Reserve Board decisions on short-term interest rates and financial decision making throughout the public and private sectors. By comparison, the BLS program to produce the CPI and other price series had an estimated budget authority of $216 million in fiscal 2016.[24]

Federal statistics also feed directly into allocations of billions of dollars of federal funds to states and local governments by use of legislated formulas. For example, formulas for allocating federal funds for education of disadvantaged children to the nation's more than 13,000 school districts use statistical estimates of school-age children in poverty produced by the U.S. Census Bureau from a model that incorporates data from the ACS and administrative records (see National Research Council, 2000). Reamer (2010) traced the allocations of $416 billion in federal funds in fiscal 2008 that depended on estimates from the ACS, postcensal population estimates

[23]See http://www.ssa.gov/policy/docs/quickfacts/stat_snapshot [April 2017].

[24]See https://www.dol.gov/sites/default/files/documents/general/budget/2016/CBJ-2016-V3-01.pdf [April 2017].

developed by updating census data with administrative records, per capita income estimates from the Bureau of Economic Analysis, and other federal statistics (see also National Research Council, 2003).

Less obvious, perhaps, but just as important as the use of federal statistics in fund allocation formulas is their use to inform legislative proposals to modify existing programs or add new ones in a wide range of areas. Federal statistics are used extensively, for example, in microsimulation and other kinds of policy models to estimate the likely benefits and costs of proposed changes to tax policy, social welfare programs such as food assistance, and many other policy proposals. Such models are used by executive branch agencies and nongovernmental organizations to make a case for or against particular proposals. The Congressional Budget Office and the Joint Committee on Taxation have formal roles in costing out policy proposals before they can be acted on by Congress, and they make extensive use of survey and other data from federal statistical agencies and other sources (see, e.g., National Research Council, 1991, 1997a, 2010b).

The Economics and Statistics Administration (2014) discusses additional uses of government economic data for public- and private-sector decision making. This study also estimated the size of the "government-data-intensive-sector," including news media, market researchers, investment advisors, pollsters, and firms that repackage government data to add value for their customers.

OTHER USES

There may not be and often is not a dollars-and-cents calculus that establishes value by relating the data from a statistical program to specific consequential decisions as can be done, say, for the CPI. Nonetheless, qualitative assessments that trace through the uses of statistical programs can often provide evidence of substantial value. An example is an effort to comprehensively identify nonfederal uses of data from the American Community Survey by state and local governments, nongovernmental organizations, the media, and others (National Research Council, 2013a; see also National Research Council, 2007b:Ch. 3). This effort identified consequential nonfederal uses in such areas as disaster preparedness, economic development and workforce planning, public health surveillance, and regional transportation planning.

More generally, statistical programs have important uses in taking the pulse of social and economic trends, designing and evaluating government programs, and supporting foundational social science research:

- ***Monitoring the social and economic health of the nation, states, and localities***—Regularly published social and economic indicators

from statistical agencies are widely cited in the media and consulted by the public to identify trends and, when estimates are available for states and local areas, to compare across areas. Some of these series, published annually or biannually, include *America's Children: Key National Indicators of Well-Being* from the Interagency Forum on Child and Family Statistics; *The Condition of Education* from the National Center for Education Statistics; *Income and Poverty in the United States* from the U.S. Census Bureau; and *Science and Engineering Indicators* from the National Science Board and National Center for Science and Engineering Statistics.[25]

- ***Providing empirical evidence for developing and evaluating federal, state, local, and private-sector programs***—For example, the American Housing Survey, sponsored by the Office of Policy Development and Research in the U.S. Department of Housing and Urban Development and conducted by the Census Bureau, provides valuable data on housing condition and housing finance to inform housing policy (see National Research Council, 2008). The Commercial Buildings Survey and the Residential Energy Consumption Survey, sponsored by the Energy Information Administration in the U.S. Department of Energy, provide valuable data for public- and private-sector policy making on end uses of various types of energy for heating, cooling, information technology, and other uses (see National Research Council, 2012a:Ch. 1).
- ***Providing input to important social science research that, in turn, informs the public and policy makers***—Many policy-relevant insights have resulted from analysis of long-running federally funded surveys, including longitudinal surveys that follow individuals over time (see, e.g., National Research Council, 2005). A few of these surveys are the Education Longitudinal Study of 2002 sponsored by the National Center for Education Statistics; the Health and Retirement Study sponsored by the National Institute on Aging and the Social Security Administration; and the National Longitudinal Surveys of Labor Market Behavior sponsored by BLS.[26]

[25]Websites [as of April 2017] for the cited series are: https://www.childstats.gov/americaschildren/; https://nces.ed.gov/programs/coe/; https://www.census.gov/library/publications/2016/demo/p60-256.html; and https://www.nsf.gov/statistics/2016/nsb20161/.

[26]Websites [as of April 2017] for the referenced surveys are: https://nces.ed.gov/surveys/els2002/; http://hrsonline.isr.umich.edu/; and https://www.bls.gov/nls/.

COST-BENEFIT ANALYSIS

Some attempts have been made to estimate costs and benefits of statistical programs or improvements in programs as input for making decisions about program priorities. However, formal cost-benefit analysis applied to federal statistical programs arguably is too dependent on questionable assumptions to be useful for this purpose.

A line of research begun in the 1970s has focused on the costs and benefits of marginal improvements in the accuracy of statistics used in congressional reapportionment, legislative redistricting, and allocation of federal funds to states and localities. As reviewed by the National Research Council (2015:51–70) (see also National Research Council, 1995:40–43), studies have found only modest effects on any of these from correcting for known net census population undercount. Studies have found more marked effects on fund allocation of using alternative estimates of per capita income.

Spencer (1997) has attempted to measure the impact of spending for improved data quality on specific policy decisions. As he acknowledges, a daunting problem is to determine an appropriate counterfactual—what the decision would have been with better (or worse) data—given that decision making rarely moves linearly from reviewing the evidence to framing appropriate policy alternatives to deciding among them.[27]

Moreover, a focus on specific policy decisions is too narrow for valuing investment in federal statistics. As one example, there is no immediately pending policy decision about the nation's rising income and wealth inequality and declining social mobility (see, e.g., Chetty et al., 2014, which used federal income tax data), but estimates of those changes have brought them to the forefront of public and policy attention. This, in turn, argues for improving federal statistics on income and wealth—increasingly problematic in surveys (see, e.g., Czajka, 2009)—to ensure as accurate a picture as possible of trends in inequality and mobility in the United States over time on which to base future policy making.

In reviewing data needs on natural gas for the Energy Information Administration, the National Research Council (1985:Ch. 3 and App. 3A) raised valid concerns about formal cost-benefit analysis of federal statistical programs. The report concluded that such analysis is largely infeasible because of the difficulty of: (1) uncovering who has used an agency's statistics—many uses are indirect without the user being aware of the data source, a problem magnified by the Internet; and (2) putting a quantitative value on each known use. The report recommended that statistical agencies emphasize regular in-depth contact with users through advisory committees

[27]For a discussion of how social science evidence is filtered in various ways in policy formation, see National Research Council (2012b).

and other mechanisms to set priorities among their programs, in preference to formal cost-benefit analyses.

CONCLUSION

The fundamental characteristic of federal statistics as a public good and the demonstrated policy, planning, research, and informational value of today's portfolio of statistical programs justify adequate budgets for federal statistics. Such funding needs to provide for research and development for continuous improvement in relevance, accuracy, timeliness, and accessibility (see Practice 10). In turn, it is incumbent on federal statistical agencies to communicate the value of their programs to policy makers and others and to analyze the costs of their programs on a continuing basis so that they can ensure the best return possible on the tax dollars invested in them.

References for Part I

All URL addresses valid as of April 2017. *Please note that, as of February 2017, U.S. Office of Management and Budget documents previously on the main whitehouse.gov site were relocated to an archived site. Wherever possible, a citation is provided to a stable site such as federalregister.gov.*

Anderson, M. J. (2015). *The American Census: A Social History* (3rd ed.). New Haven, CT: Yale University Press.

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014). Is the United States still a land of opportunity? Recent trends in intergenerational mobility. *American Economic Review Papers and Proceedings 104*(5), 141–147. Available: https://doi.org/10.1257/aer.104.5.141; National Bureau of Economic Research Working Paper version available: https://doi.org/10.3386/w19844.

Citro, C. F. (2014a). From multiple modes for surveys to multiple data sources for estimates. *Survey Methodology 40*(2), 137–161. Available: http://www.statcan.gc.ca/pub/12-001-x/2014002/article/14128-eng.htm.

Citro, C. F. (2014b). Principles and practices for a federal statistical agency: Why, what, and to what effect. *Statistics and Public Policy 1*(1), 51–59. Available: http://dx.doi.org/10.1080/2330443X.2014.912953.

Citro, C. F. (2016). The US federal statistical system's past, present, and future. *Annual Review of Statistics and Its Application 3*(1), 347–373. Available: http://dx.doi.org/10.1146/annurev-statistics-041715-033405.

Czajka, J. (2009). SIPP data quality. Pp. 157–178 in National Research Council, *Reengineering the Survey of Income and Program Participation*. Panel on the Census Bureau's Reengineered Survey of Income and Program Participation, C.F. Citro and J.K. Scholz, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12715.

Duncan, J. W. and W. C. Shelton (1978). *Revolution in United States Statistics, 1926–1976*. Office of Federal Statistical Policy and Standards, U.S. Department of Commerce. Washington, DC: U.S. Government Printing Office. Available: https://catalog.hathitrust.org/Record/007886111/Home.

Eberstadt, N., R. Nunn, D. W. Schanzenbach, and M. R. Strain (2017). *"In Order That They Might Rest Their Arguments on Facts": The Vital Role of Government-Collected Data*. Washington, DC: The Hamilton Project (Brookings Institution) and American Enterprise Institute. Available: https://www.brookings.edu/wp-content/uploads/2017/02/thp_20170227_govt_collected_data_report.pdf.

Economics and Statistics Administration (2014). *Fostering Innovation, Creating Jobs, Driving Better Decisions: The Value of Government Data*. Washington, DC: U.S. Department of Commerce. Available: https://esa.gov/sites/default/files/revisedfosteringinnovationcreatingjobsdrivingbetterdecisions-thevalueofgovernmentdata.pdf.

European Statistical System Committee (2011). *European Statistics Code of Practice for the National and Community Statistical Authorities*. Luxembourg: Eurostat. Available: http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-32-11-955.

Gravelle, H. and R. Rees (2004). *Microeconomics* (3rd ed.). Essex, UK: Pearson Education Limited. Available: https://ignorelist.files.wordpress.com/2012/01/microeconomics-gravelle-and-rees.pdf.

Groves, R. M. (2011). Three eras of survey research. *Public Opinion Quarterly 75*(5), 861–871. Available: http://dx.doi.org/10.1093/poq/nfr057.

McMillen, D. (2012). Apportionment and districting. Pp. 49–58 in *Encyclopedia of the U.S. Census: From the Constitution to the American Community Survey (ACS)* (2nd edition). M.J. Anderson, C.F. Citro, and J.J. Salvo, eds. Washington, DC: CQ Press.

National Academies of Sciences, Engineering, and Medicine (2017a). *Innovations in Federal Statistics Using New Data Sources While Respecting Privacy*. Panel on Improving Federal Statistics for Policy and Social Science Research Using Multiple Data Sources and State-of-the-Art Estimation Methods, R.M. Groves and B.A. Harris-Kojetin, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/24652.

National Academies of Sciences, Engineering, and Medicine (2017b). *Principles and Practices for Federal Program Evaluation: Proceedings of a Workshop—In Brief*. J. White, rapporteur. Committee on National Statistics. Washington, DC: National Academies Press. Available: https://doi.org/10.17226/24716.

National Research Council (1985). *Natural Gas Data Needs in a Changing Regulatory Environment*. Panel on Natural Gas Statistics, Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/19272.

National Research Council (1991). *Improving Information for Social Policy Decisions: The Uses of Microsimulation. Volume I: Review and Recommendations*. Panel to Evaluate Microsimulation Models for Social Welfare Programs, C.F. Citro and E.A. Hanushek, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/1835.

National Research Council (1995). *Modernizing the U.S. Census*. Panel on Census Requirements in the Year 2000 Census and Beyond, B. Edmonston and C. Schultze, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/4805.

National Research Council (1997a). *Assessing Policies for Retirement Income: Needs for Data, Research, and Models*. Panel on Retirement Income Modeling, C.F. Citro and E.A. Hanushek, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/5420.

National Research Council (1997b). *The Bureau of Transportation Statistics: Priorities for the Future*. Panel on Statistical Programs and Practices of the Bureau of Transportation Statistics, C.F. Citro and J.L. Norwood, eds. Committee on National Statistics and Transportation Research Board. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/5809.

National Research Council (1999). *Sowing Seeds of Change: Informing Public Policy in the Economic Research Service of USDA*. Panel to Study the Research Program of the Economic Research Service, J.F. Geweke, J.T. Bonnen, A.A. White, and J.J. Koshel, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/6320.

National Research Council (2000). *Small-Area Income and Poverty Estimates: Priorities for 2000 and Beyond*. Panel on Estimates of Poverty for Small Geographic Areas, C.F. Citro and G. Kalton, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/9957.

National Research Council (2001). *Evaluating Welfare Reform in an Era of Transition*. Panel on Data and Methods for Measuring the Effects of Changes in Social Welfare Programs, R.A. Moffitt and M. Ver Ploeg, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/10020.

National Research Council (2003). *Statistical Issues in Allocating Funds by Formula*. Panel on Formula Allocations, T.A. Louis, T.B. Jabine, and M.A. Gerstein, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/10580.

National Research Council (2005). *Expanding Access to Research Data: Reconciling Risks and Opportunities*. Panel on Data Access for Research Purposes, Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11434.

National Research Council (2007a). *State and Local Government Statistics at a Crossroads*. Panel on Research and Development Priorities for the U.S. Census Bureau's State and Local Government Statistics Program. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12000.

National Research Council (2007b). *Using the American Community Survey: Benefits and Challenges*. Panel on the Functionality and Usability of Data from the American Community Survey, C.F. Citro and G. Kalton, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11901.

National Research Council (2008). *Rebuilding the Research Capacity at HUD*. Committee to Evaluate the Research Plan of the Department of Housing and Urban Development, Center for Governance, Economic, and International Studies. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12468.

National Research Council (2010a). *Data on Federal Research and Development Investments: A Pathway to Modernization*. Panel on Modernizing the Infrastructure of the National Science Foundation Federal Funds Survey, Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12772.

National Research Council (2010b). *Improving Health Care Cost Projections for the Medical Population: Summary of a Workshop*. G. Wunderlich, rapporteur. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12985.

National Research Council (2012a). *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 and Board on Energy and Environmental Systems. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/13360.

National Research Council (2012b). *Using Science as Evidence in Public Policy*. Committee on the Use of Social Science Knowledge in Public Policy, K. Prewitt, T.A. Schwandt, and M.L. Straf, eds. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/13460.

National Research Council (2013a). *Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop*. D.L. Cork, rapporteur. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/18259.

National Research Council (2013b). *Principles and Practices for a Federal Statistical Agency, Fifth Edition*. Committee on National Statistics, C.F. Citro and M.L. Straf, eds.. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/18318.

National Research Council (2015). *The Bicentennial Census: New Directions for Methodology in 1990: 30th Anniversary Edition*. Panel on Decennial Census Methodology, C.F. Citro and M.L. Cohen, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://www.nap.edu/catalog/21728/the-bicentennial-census-new-directions-for-methodology-in-1990-30th.

Norwood, J. L. (1995). *Organizing to Count: Change in the Federal Statistical System*. Washington, DC: The Urban Institute Press.

Norwood, J. L. (2016). Politics and federal statistics. *Statistics and Public Policy 3*(1), 1–8. American Statistical Association Distinguished Statistician Lecture, 2000. Available: http://dx.doi.org/10.1080/2330443X.2016.1241061.

Prewitt, K. (Ed.) (2010, September). *The Federal Statistical System: Its Vulnerability Matters More Than You Think*, Volume 631 of *The ANNALS of the American Academy of Political and Social Science*. Thousand Oaks, CA: SAGE Publications. Available: http://journals.sagepub.com/toc/anna/631/1.

Privacy Protection Study Commission (1977). *Personal Privacy in an Information Society: The Report of The Privacy Protection Study Commission*. Washington, DC: U.S. Government Printing Office. Available: https://catalog.hathitrust.org/Record/000293901.

Reamer, A. D. (2010). *Surveying for Dollars: The Role of the American Community Survey in the Geographic Distribution of Federal Funds*. Metropolitan Policy Program. Washington, DC: The Brookings Institution. Available: https://www.brookings.edu/wp-content/uploads/2016/06/0726_acs_reamer.pdf.

Singer, E. (2016). Reflections on surveys' past and future. *Journal of Survey Statistics and Methodology 4*(4), 463–475.

Spencer, B. D. (Ed.) (1997). *Statistics and Public Policy*. Oxford, UK: Clarendon Press.

UNECE Task Force on the Value of Official Statistics (2017, March). *Value of Official Statistics: Recommendations on Promoting, Measuring and Communicating the Value of Official Statistics*. Report of the task force to the United Nations Economic Commission for Europe, Conference of European Statisticians. Geneva: United Nations. Available: https://www.unece.org/statistics/statstos/task-force-on-the-value-of-official-statistics.html.

United Nations Statistical Commission (2014, March 3). *Fundamental Principles of Official Statistics*. New York: United Nations. Endorsed by the General Assembly as Resolution A/RES/68/261. Available: http://unstats.un.org/unsd/dnss/gp/FP-New-E.pdf.

U.S. General Accounting Office (1995). *Statistical Agencies: Adherence to Guidelines and Coordination of Budgets*. Report No. GGD-95-65. Washington, DC: U.S. Government Printing Office. Available: http://www.gao.gov/products/GGD-95-65.

U.S. Government Accountability Office (2007). *Bureau of Justice Statistics: Quality Guidelines Generally Followed for Police-Public Contact Surveys, but Opportunities Exist to Help Assure Agency Independence*. Report No. GAO-07-340. Washington, DC: U.S. Government Printing Office. Available: http://www.gao.gov/products/GAO-07-340.

U.S. Government Accountability Office (2012). *Federal Statistical System: Agencies Can Make Greater Use of Existing Data, but Continued Progress Is Needed on Access and Quality Issues*. Report No. GAO-12-54. Washington, DC: U.S. Government Printing Office. Available: http://www.gao.gov/products/GAO-12-54.

U.S. Office of Management and Budget (1985). Statistical Policy Directive No. 3: Compilation, release, and evaluation of principal federal economic indicators. 50 *Federal Register* 38932 (September 25, 1985). Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/omb/inforeg/statpolicy/dir_3_fr_09251985.pdf.

U.S. Office of Management and Budget (2007). Implementation guidance for Title V of the E-Government Act, Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA). 72 *Federal Register* 33362 (June 15, 2007). Available: https://www.federalregister.gov/d/E7-11542.

U.S. Office of Management and Budget (2014). Statistical Policy Directive No. 1: Fundamental responsibilities of federal statistical agencies and recognized statistical units. 79 *Federal Register* 71609 (December 2, 2014). Available: https://www.federalregister.gov/d/2014-28326.

U.S. Office of Management and Budget (2016). *Analytical Perspectives, Budget of the U.S. Government, Fiscal Year 2017*. Washington, DC: U.S. Government Printing Office. Available: https://www.govinfo.gov/content/pkg/BUDGET-2017-PER/pdf/BUDGET-2017-PER.pdf.

U.S. Office of Management and Budget (2017). *Statistical Programs of the United States Government: Fiscal Year 2017*. Statistical and Science Policy Office, Office of Information and Regulatory Affairs. Washington, DC: U.S. Government Printing Office. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/information_and_regulatory_affairs/statistical-programs-2017.pdf.

Part II

Principles for a Federal Statistical Agency

Principle 1: Relevance to Policy Issues

A federal statistical agency must be in a position to provide objective, accurate, and timely information that is relevant to issues of public policy.

Federal statistical agencies exist to provide objective, accurate, and timely information that is relevant for policy and public use. Federal statistics are used both inside and outside the government not only to delineate problems and to guide policy and program action, but also to evaluate the results of government activity or lack of activity. This role places a heavy responsibility on federal statistical agencies to ensure the impartiality, objectivity, and quality of federal statistics and their continued relevance as social circumstances and public and policy concerns change.

A statistical agency must be knowledgeable about the issues and requirements of public policy and federal programs within its scope and able to provide correspondingly relevant information for use by managers, overseers, and policy makers in the executive branch and in Congress. A statistical agency's programs must also be relevant for everyone who requires objective statistical information that pertains to its mission areas. Such information is needed for many purposes, including economic decision making, state and local government program planning, research to improve knowledge of social systems and behaviors, and people's participation in civic affairs. A free enterprise economic system depends on the availability of economic information to all participants; a vibrant research enterprise that can contribute to societal well-being depends on the availability of relevant information to all researchers; and a democratic political system depends on—and has a fundamental duty to provide—access to information on

education, health, transportation, the economy, the environment, criminal justice, and other areas of public concern.

In order to provide relevant information, statistical agencies need to reach out to the users of their data. Although agencies are usually in touch with the primary users in their own departments, they may need to expend considerable energy and initiative to open avenues of communication more broadly with other current and potential users. The range of such users includes analysts and policy makers in other federal departments, members of Congress and their staffs, state and local government agencies, academic researchers, private-sector businesses and other organizations, organized constituent groups, associations that represent data users, and the media.

One way to obtain the views of users outside a statistical agency, as well as people with relevant technical expertise, is through advisory committees (see National Research Council, 1993:Ch. 8, 2007:Ch. 7). Many agencies obtain advice from committees that are chartered under the Federal Advisory Committee Act: examples include the Advisory Committee on Agriculture Statistics for the National Agricultural Statistics Service; the Board of Scientific Counselors for the National Center for Health Statistics; the Data Users Advisory Committee and the Technical Advisory Committee for the Bureau of Labor Statistics; and the Scientific Advisory Committee and the National Advisory Committee on Race, Ethnic, and Other Populations for the U.S. Census Bureau. The Federal Economic Statistics Advisory Committee (FESAC), chartered in November 1999, provides substantive and technical advice to three agencies—the Bureau of Economic Analysis, the Bureau of Labor Statistics, and the Census Bureau—thereby providing an important cross-cutting perspective on major economic statistics programs.[28] Some agencies obtain advice from committees and working groups that are organized by an independent association, such as the American Statistical Association's Committee on Energy Statistics for the Energy Information Administration.

Other means to get input from current and potential users include workshops and conferences, which are valuable for facilitating interchange among users and agency staff (see National Research Council, 2013a). Online mechanisms, such as blogs and web surveys, may also assist a statistical agency to obtain input from users.

It is important for an agency's own staff to engage in analysis of its data to improve them and make them more relevant to users (Martin, 1981; Norwood, 1975; Triplett, 1991). Such analyses may use the agency's data to examine correlates of key social or economic phenomena or to study the statistical error properties of the data. Similarly, agencies can use web analytics to better understand their user base, as well as to assess

[28]See http://www.census.gov/fesac/ [April 2017].

the accessibility and usability of their website and data products. Such in-house analysis can lead to improvements in the quality of the statistics, to the identification of new needs for information and data products, to a reordering of priorities, and to closer cooperation and mutual understanding with data users.

Statistical agencies are sometimes asked to work with policy analysis units. Statistical agencies may properly advise on the availability and strengths and limitations of relevant information. The substantive analyses that statistical agencies produce as a regular part of their dissemination and research activities will also likely be helpful to policy analysis units, as well as other data users. These analyses typically describe relevant conditions and trends and may estimate the likely effects of future population growth or other factors on trends (e.g., trends in high school completion rates, poverty rates, or fuel usage) or the differences in trends from the use of alternative statistical measures (e.g., of unemployment or income). A statistical agency may properly extend such analyses at the request of a policy analysis unit or other data user (e.g., by examining trends for particular population groups). However, a statistical agency neither makes policy recommendations nor carries out substantive analyses of policies that go beyond description. The distinction between substantive analysis that is part of the mission of a statistical agency and policy analysis is not always clear, and a statistical agency will need to consider carefully the extent of policy-related activities that are appropriate for it to undertake.

Practices that are particularly relevant for Principle 1

1 mission
2 authority for independence
3 use of multiple data sources
5 wide dissemination
6 cooperation with users
10 research program
12 internal and external evaluation
13 collaboration with other agencies

Principle 2: Credibility among Data Users

A federal statistical agency must have credibility with those who use its data and information.

It is essential that a statistical agency strive to maintain credibility for itself and for its data. Few data users are in a position to verify the completeness and accuracy of statistical information; they must rely on an agency's reputation as a trustworthy source of accurate and useful statistics derived in an objective and impartial manner. With the appearance and reality of credibility, an agency is perceived to be working in the national interest, not in the interest of a particular administration or one set of users over another (Ryten, 1990).

Credibility derives from the respect and trust of users in the statistical agency and its data. Such respect results, first and foremost, from an agency's production of data that merit acceptance as relevant, accurate, timely, and free from political and other undue external influence. Respect also results from many aspects of an agency's policies and practices. Thus, credibility is enhanced when an agency actively engages users on priorities for data collection and analysis; fully informs users of the strengths and weaknesses of the data; makes data available widely on an equal basis to all users; and exhibits a posture of respect toward and trust in users. That respect includes a willingness to understand and strive to meet users' needs, even though users may not clearly articulate them. Credibility is further enhanced when an agency shows commitment to quality and professional practice and maintains a strong internal and external evaluation program to assess and improve its statistical programs, methods, and processes.

Essential to credibility is that a statistical agency's website be as accessible as possible for locating, working with, and understanding the strengths and limitations of the agency's data. Such accessibility—fully tested with a broad array of users—recognizes the central role of users as the ultimate consumers of a statistical agency's products.

An agency's website, as its primary means of communication with users, should have readily accessible information about its policies on such topics as confidentiality and privacy protection; scientific integrity; standards for data quality and for documenting sources of error in data collections and estimation models; procedures and schedules for the release of new and continuing data series; procedures for timely notice of errors and corrections to previously released data; and procedures and schedules for archiving historical data. Links to policies of an agency's parent department or independent agency that clearly specify the authority that is delegated to the statistical agency also enhance credibility and build trust with users.

Today, with many alternate sources of data available and with public trust in many government institutions at a low ebb,[29] it is undeniably challenging for a statistical agency to be perceived as credible and trustworthy. Statistical agencies, individually and collectively, need to respond proactively to this challenge, seeking input from knowledgeable data users on how best to do so.

Practices that are particularly relevant for Principle 2

2 authority for independence
3 use of multiple data sources
4 openness about sources and limitations
5 wide dissemination
6 cooperation with users
9 commitment to quality
11 professional advancement of staff
12 internal and external evaluation

[29] A 2012–2013 Gallup survey, commissioned by several federal statistical agencies, asked respondents whether they had "a great deal/quite a lot" of confidence for several institutions. The results were 70 percent for the military, 50 percent for universities, 25–30 percent for federal statistical agencies and newspapers, and 10 percent for Congress (Childs et al., 2015).

Principle 3: Trust among Data Providers

A federal statistical agency must have the trust of those whose information it obtains.

The statistical programs of the federal government rely on information supplied by many data providers, including individual members of the public, other agencies of the federal government, and organizations outside the federal government, such as state and local governments, businesses, and other organizations. Some of this information is a by-product of data collections that are required by law or regulation for use in the administration of government tax and transfer programs, such as employers' wage reports to state employment security agencies or records of payments to program beneficiaries. Much of it is obtained through the voluntary cooperation of respondents in statistical surveys. Even when response is mandatory, as in the case of such essential statistical programs as the population and economic censuses, the willing cooperation of respondents reduces costs and likely promotes accuracy (see National Academies of Sciences, Engineering, and Medicine, 2016; National Research Council, 1995, 2004, 2013b).

A high standard of ethical conduct on the part of a statistical agency is critical for obtaining the cooperation of data providers, whether they are individuals or organizational entities or custodians of administrative records. Thus, trust is engendered when data providers can rely on the word of a statistical agency that the information they are asked to provide is important and legitimate for the government to collect, is being collected in an impartial and competent manner, and will be used only for the purposes that the agency has described.

Data providers must be able to trust the word of a statistical agency that it will scrupulously honor its promises to protect individual responses collected under a pledge of confidentiality. Such protection should preclude that any individually identifiable information collected for statistical purposes is used for any administrative, regulatory, or law enforcement purpose. It should require, moreover, that agencies use state-of-the-art methods to protect—to the extent possible—against reidentification of individual records in statistical data products that are available for public use (see National Academies of Sciences, Engineering, and Medicine, 2017:Ch. 5). Agency and contractor staff and researchers who have been given controlled access to individual records should follow proper procedures to guard against disclosure and suffer penalties if a disclosure occurs.

In today's world, when intrusions into computer networks are distressingly frequent, protection of confidentiality and respect for privacy further requires the use of state-of-the-art technology to ensure cybersecurity. The Federal Cybersecurity Enhancement Act of 2015 (see Appendix A) requires all agencies to use a system known as EINSTEIN (currently in version E3A) to guard against and detect cybersecurity breaches. The added protection from this advanced technology is a benefit to statistical agencies.[30]

To engender trust, a statistical agency should also respect the privacy of data providers in other ways and ensure that individuals' decisions to respond to a survey are made with full information (i.e., are autonomous).[31] Such respect requires an agency to minimize the intrusiveness of questions and the time and effort to respond to them to the maximum extent possible that is consistent with the agency's requirements for information. It also requires that an agency's data collection staff take care to treat respondents with courtesy and show appreciation for their time (see National Academies of Sciences, Engineering, and Medicine, 2016). Respect further requires that an agency provide sufficient information for the provider to make an informed decision about whether to supply the requested data, including the intended uses of the data being collected, their relevance for important public purposes, and the extent of confidentiality protection that will be provided.

[30]However, because the Act allows U.S. Department of Homeland Security staff to monitor traffic conducted over any federal agency network and take follow-up actions as necessary, agencies have modified their confidentiality pledges to inform respondents of screening of records for cybersecurity purposes (see Practice 8).

[31]The requirements for informed consent to participate in surveys and other research under the Federal Policy for the Protection of Human Subjects (Common Rule) (https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/index.html [April 2017]) rest on the concept of autonomy in the "Belmont Report": "To respect autonomy is to give weight to autonomous persons' considered opinions and choices" (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1978:5).

When data are obtained from the administrative records of other federal, state, or local government agencies or other third-party providers, the same principle of trust applies in order to secure the fullest possible cooperation of these providers with a statistical agency's needs for the records and associated documentation.[32] Provider agencies and organizations need to be able to trust that their records are important and legitimate for a statistical agency to obtain, that their own restrictions on data access will be honored, and that the statistical agency will make every effort to minimize the burden for the provider agency of responding to the statistical agency's requests.

Practices that are particularly relevant for Principle 3

2 authority for independence
7 respect for privacy and autonomy
8 protection of confidentiality

[32] U.S. Office of Management and Budget (2014a) asserts the legitimacy and benefits of use of administrative data from other federal agencies for statistical agency purposes. It also provides guidance for best practices and procedures to engender mutual respect and trust and facilitate such data sharing.

Principle 4: Independence from Political and Other Undue External Influence

A federal statistical agency must be independent from political and other undue external influence in developing, producing, and disseminating statistics.

The reason for a statistical agency to exist is to serve as a trustworthy source of objective, relevant, accurate, and timely information for decision makers, analysts, and others—both inside and outside the government—to help them understand present conditions, draw comparisons with the past, and guide plans for the future.[33] For these purposes, it is essential that a statistical agency be distinct from those parts of a department that carry out administrative, regulatory, law enforcement, or policy-making activities. It is also essential that a statistical agency have a widely acknowledged position of independence from political and other undue external influence in developing, producing, and disseminating statistics, together with the necessary authority to protect independence (see Practice 2). A statistical agency actively works to obtain a broad range of external input to develop its programs and for its programs to be relevant to policy needs: "undue external influences" are those from outside the agency that seek to undermine its impartiality, nonpartisanship, and professional judgment.

A statistical agency must be able to execute its mission without being subject to pressures to advance a political agenda. It must be impartial

[33] See United Nations Statistical Commission (2014), European Statistical System Committee (2011), and U.S. Office of Management and Budget (2014b).

and avoid even the appearance that its collection, analysis, and reporting processes might be manipulated for political purposes or that individually identifiable data collected under a pledge of confidentiality might be turned over for nonstatistical purposes. Independence from any undue outside influence is an essential element of credibility with data users and the public and of trust among data providers.

The independence of a statistical agency is exercised in a broad framework. Legislative authority usually gives ultimate responsibility to the secretary of the department rather than to the head of the statistical agency. In addition, a statistical agency is subject to the normal budgetary processes and to various coordinating and review functions of the U.S. Office of Management and Budget, as well as the legislative mandates and oversight of Congress. Within this broad framework, a statistical agency has to maintain its credibility as an impartial purveyor of information. In the long run, the effectiveness of an agency depends on its maintaining a reputation for impartiality: thus, an agency must be continually alert to possible infringements on its credibility and be prepared to argue strenuously against such infringements.

For an agency head, independence and protection from undue political influence can be strengthened by the method of the person's appointment. Two methods that can bolster the professional independence of an agency head are appointment by the President with confirmation by the Senate for a fixed term, as is the case for the Bureau of Labor Statistics (BLS) and the U.S. Census Bureau, and departmental appointment of a career civil servant.[34]

For a fixed term, it is desirable that it not coincide with the presidential term so that professional considerations, rather than political ones, are more likely to be paramount in the appointment process. Appointment by the President with Senate confirmation for a term that is at the pleasure of the President, as is the case for the head of the Energy Information Administration (EIA), is not ideal for agency independence. However, EIA does have strong legislative protection for the authority of its administrator.

In 2012, the heads of two agencies, the Bureau of Justice Statistics (BJS) and the National Center for Education Statistics (NCES), were changed from presidential appointees with Senate confirmation (the NCES commissioner

[34] Agencies headed by career civil servants, many of whom hold their positions for long periods of time, include the Bureau of Economic Analysis; the Bureau of Transportation Statistics; the Economic Research Service in the U.S. Department of Agriculture; the National Agricultural Statistics Service; the National Center for Health Statistics; the National Center for Science and Engineering Statistics; the Office of Research, Evaluation, and Statistics in the Social Security Administration; and the Statistics of Income Division in the Internal Revenue Service.

having a fixed term) to presidential appointees.[35] The consequences for the independence of these two agencies are not clear. However, the change can be viewed as part of the "layering of statistical agencies," that is, positioning them lower in their department's administrative structure, which National Research Council (2009:226) identified as "a subtle, but increasingly common" threat to independence.

It is valuable for the head of a statistical agency to have direct access to the secretary of the department or the head of the independent agency in which the statistical agency is located. Such access allows the head to inform new secretaries about the role of a statistical agency and be able to directly present the case for new or changed statistical initiatives. Such direct access currently is provided by legislation for BLS and EIA but not for other agencies, which are often several administrative rungs below the secretary (see Figure I.2).

A statistical agency should have its own funding appropriation from Congress and not be dependent on allocations from the budget of its parent department or agency, which may be subject to reallocation. In addition, it can be helpful when a specified percentage of a department's budget is set aside for its statistical agency, which signals the importance of the statistical agency function.

While such organizational aspects as direct access to the secretary of the agency's department and separate budgetary authority facilitate a strong position of independence, they are neither necessary nor sufficient. In that regard, agencies can do much to maintain their independence through other practices, procedures, and policies.

Practices that are particularly relevant for Principle 4

1 mission
2 authority for independence
10 research program
11 professional advancement of staff
12 internal and external evaluation

[35]This change for BJS and NCES was part of the Presidential Appointment Efficiency and Streamlining Act of 2011 (P.L. 112-166), which made a large number of subcabinet-level posts presidential appointments that no longer required Senate confirmation. The act also provided that the director of the Census Bureau remain a presidential appointee with Senate confirmation but have a fixed 5-year term (with one renewal permitted) for terms beginning on January 1 of years ending in 2 and 7. Previously, the Census Bureau director served at the pleasure of the President.

References for Part II

All URL addresses valid as of April 2017. *Please note that, as of February 2017, U.S. Office of Management and Budget documents previously on the main whitehouse.gov site were relocated to an archived site. Wherever possible, a citation is provided to a stable site such as federalregister.gov.*

Childs, J. H., R. King, and A. C. Fobia (2015). Confidence in US federal statistical agencies. *Survey Practice 8*(4). Available: http://www.surveypractice.org/index.php/SurveyPractice/article/view/314/html_38.

European Statistical System Committee (2011). *European Statistics Code of Practice for the National and Community Statistical Authorities*. Luxembourg: Eurostat. Available: http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-32-11-955.

Martin, M. E. (1981). Statistical practice in bureaucracies. *Journal of the American Statistical Association 76*(373), 1–8. Available: http://www.tandfonline.com/doi/abs/10.1080/01621459.1981.10477593.

National Academies of Sciences, Engineering, and Medicine (2016). *Reducing Respondent Burden in the American Community Survey: Proceedings of a Workshop*. T.J. Plewes, rapporteur. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/23639.

National Academies of Sciences, Engineering, and Medicine (2017). *Innovations in Federal Statistics Using New Data Sources While Respecting Privacy*. Panel on Improving Federal Statistics for Policy and Social Science Research Using Multiple Data Sources and State-of-the-Art Estimation Methods, R.M. Groves and B.A. Harris-Kojetin, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/24652.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1978). *The Belmont Report: Ethical Principles and Guidelines*

for the Protection of Human Subjects of Research. Bethesda, MD: U.S. Government Printing Office. Available: http://archive.org/details/belmontreporteth00unit.

National Research Council (1993). *The Future of the Survey of Income and Program Participation*. Panel to Evaluate the Survey of Income and Program Participation, C.F. Citro and G. Kalton, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/2072.

National Research Council (1995). *Modernizing the U.S. Census*. Panel on Census Requirements in the Year 2000 Census and Beyond, B. Edmonston and C. Schultze, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/4805.

National Research Council (2004). *The 2000 Census: Counting Under Adversity*. Panel to Review the 2000 Census, C.F. Citro, D.L. Cork, and J.L. Norwood, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/10907.

National Research Council (2007). *Using the American Community Survey: Benefits and Challenges*. Panel on the Functionality and Usability of Data from the American Community Survey, C.F. Citro and G. Kalton, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11901.

National Research Council (2009). *Ensuring the Quality, Credibility, and Relevance of U.S. Justice Statistics*. Panel to Review the Programs of the Bureau of Justice Statistics, R.M. Groves and D.L. Cork, eds. Committee on National Statistics and Committee on Law and Justice. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12671.

National Research Council (2013a). *Benefits, Burdens, and Prospects of the American Community Survey: Summary of a Workshop*. D.L. Cork, rapporteur. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/18259.

National Research Council (2013b). *Nonresponse in Social Science Surveys: A Research Agenda*. Panel on a Research Agenda for the Future of Social Science Data Collection, R. Tourangeau and T.J. Plewes, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/18293.

Norwood, J. L. (1975). Should those who produce statistics analyze them? How far should analysis go? An American view. *Bulletin of the International Statistical Institute [Proceedings of the 40th Session]* 46, 420–432.

Ryten, J. (1990). Statistical organization criteria for inter-country comparisons and their application to Canada. *Journal of Official Statistics* 6(3), 319–332. Available: http://www.jos.nu/Articles/abstract.asp?article=63319.

Triplett, J. (1991). The federal statistical system's response to emerging data needs. *Journal of Economic and Social Measurement 17*(3–4), 155–201. Available: http://content.iosPress.com/articles/journal-of-economic-and-social-measurement/jem17-3-4-03.

United Nations Statistical Commission (2014, March 3). *Fundamental Principles of Official Statistics*. New York: United Nations. Endorsed by the General Assembly as Resolution A/RES/68/261. Available: http://unstats.un.org/unsd/dnss/gp/FP-New-E.pdf.

U.S. Office of Management and Budget (2014a). Guidance for providing and using administrative data for statistical purposes. OMB Memorandum M-14-06. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2014/m-14-06.pdf.

U.S. Office of Management and Budget (2014b). Statistical Policy Directive No. 1: Fundamental responsibilities of federal statistical agencies and recognized statistical units. 79 *Federal Register* 71609 (December 2, 2014). Available: https://www.federalregister.gov/d/2014-28326.

Part III

Practices for a Federal Statistical Agency

Practice 1: A Clearly Defined and Well-Accepted Mission

A CLEAR UNDERSTANDING of the mission of a statistical agency, the scope of its programs, and its authority and responsibilities is basic to planning and evaluating its programs and maintaining credibility and independence from political control (National Research Council, 1986, 1997b). Some agency missions are clearly spelled out in legislation; other agencies have only general legislative authority. On occasion, specific requirements are set by legislation or regulation (see Appendix B).

An agency's mission should include at least three major elements: The first is responsibility for all elements of its statistical programs: sources of data, measurement methods, efficient methods of data collection and processing, and appropriate methods of analysis. The second is responsibility for evaluating, implementing, and documenting new methods and processes that better serve users' needs for relevant, accurate, and timely statistics within budget constraints. The third is responsibility for curation to ensure the availability for future use of the agency's data together with documentation of the methods used and the quality of the estimates.

A statistical agency's mission should also include primary responsibility for determining the subjects and questions for which it will collect data, acknowledging that Congress and the executive branch may properly exercise authority in this area (see Practice 2). In all instances, a statistical agency's mission must focus on information that is to be used for statistical purposes. When an agency is charged to carry out an activity that could undermine its credibility as an objective source of information (e.g., collecting data for both statistical and administrative purposes), the agency should carefully describe and structure the activity (e.g., perhaps locating

it within a clearly demarcated office). If the agency concludes that it is not possible to develop a satisfactory arrangement that is both responsive to the charge and credible, it should request that the activity be assigned elsewhere.

A statistical agency should clearly communicate its mission to others. The Internet is today the primary means for an agency, not only to disseminate its data and associated documentation, but also to publicize its mission to a broad audience and to provide related information, including enabling legislation, the scope of the agency's statistical programs, confidentiality provisions, and data quality guidelines. Consequently, careful design of an agency's website to maximize the ability to easily navigate the site is imperative.

A statistical agency should periodically review its mission. As part of strategic planning to carry out its mission within its budget, it should review priorities among different programs (see, e.g., National Research Council, 1976, 2000b, 2009a), the infrastructure (e.g., computing capabilities) needed to support them, and the relative importance and urgency of needed improvements, say, in timeliness versus accuracy.

Such a process will take time and may open an agency to charges that its priorities are misguided (e.g., when an agency decides to cut back on an existing program in favor of a new initiative). Nonetheless, the process should put the agency in a stronger position to justify its portfolio to stakeholders and to more efficiently allocate scarce staff time and other resources among programs and initiatives. Regular reviews may also identify authority that would be useful for an agency to have, such as authority to acquire administrative records or to be a clearinghouse for information collected in a topic area from its own programs and those of other statistical agencies.

Advice from outside groups should be sought, formally and informally, on the agency's priorities for data collection and data products and services and on its statistical and operational methods. Such advice should be obtained from both data users and providers and from professional and technical experts in the subject-matter area, in statistical methods, and in information technology (see Practice 6). A strong research program in the agency's subject-matter field can assist in setting priorities and identifying ways to improve an agency's statistical programs because subject-matter researchers at an agency are well positioned to communicate data user needs to agency statistical and operational staff and vice versa (Triplett, 1991; see also Practice 10).

Practice 2: Necessary Authority to Protect Independence

Protection from political and other undue external influence over a statistical agency's data collection, production, and dissemination requires that an agency has the necessary authority for professional decisions in key aspects of its work. Only with such authority can a statistical agency establish and maintain its reputation for objectivity and quality in carrying out its mission to inform policy makers and the public. Within the limits of budgetary resources, departmental requirements, review by the U.S. Office of Management and Budget (OMB), and congressional mandates (see "Limits on Authority," below), a statistical agency requires the following:

- authority for professional decisions over the scope, content, and frequency of data compiled, analyzed, and published within the framework set by an agency's authorizing legislation;
- authority for the statistical agency head and qualified staff to discuss the agency's statistics with members of Congress, congressional staff, and in appropriate public forums;
- authority for selection and promotion of professional, technical, and operational staff;
- authority to control information technology (IT) systems to securely maintain the integrity and confidentiality of individual records and reliably support timely and accurate production of key statistics; and
- authority—and recognition by policy officials outside the statistical agency of that authority—to release statistical information, including

accompanying press releases and documentation, without prior clearance regarding the statistical content of the release.[36]

Necessary authority may come from legislation, OMB directives, and policies and practices that are communicated by agency leadership to political appointees. The two most recent additions to the policies that help protect statistical agency independence are: (1) "Statement of Commitment to Scientific Integrity by Principal Statistical Agencies;"[37] and (2) *Statistical Policy Directive No. 1* (U.S. Office of Management and Budget, 2014b). The statement affirms the agencies' commitment to the principles and practices in National Research Council (2009c), OMB statistical policy directives, and guidelines required by the Information Quality Act, all of which uphold the necessity for statistical agency independence from undue outside influence. The directive affirms the need for statistical agency autonomy in professional decisions and the obligation of departments to support their statistical agencies in this regard (see Appendix A).

ASPECTS OF INDEPENDENCE

Authority to decide the scope and specific content of the data collected or compiled, within the scope of a statistical agency's mission, and to make decisions about technical aspects of data collection programs is an important element of independence. Similarly, authority for a statistical agency head and qualified staff to discuss the agency's statistics with members of Congress, congressional staff, and appropriate public entities helps to maintain an agency's standing.

Another important aspect of independence is control over personnel actions, especially the selection of qualified professional staff, including senior executive career staff. Agency staff who report directly to the agency head should have formal education and deep experience in the substantive, methodological, operational, and management issues facing the agency as appropriate for their positions. For the head of a statistical agency, professional qualifications are of the utmost importance, whether the profession is that of statistician or a relevant subject-matter field (National Research Council, 1997b). Relevant professional associations can be a source of valuable input about suitable candidates.

The authority to ensure that IT systems fulfill the specialized needs of a statistical agency is another important aspect of independence. A statistical agency must be able to vouch for the integrity, confidentiality, and impartiality of the information collected and maintained under its

[36]Such authority is provided in *Statistical Policy Directives Nos. 3* and *4* (U.S. Office of Management and Budget, 1985, 2008).

[37]See, for example, http://www.bls.gov/bls/integrity.htm [April 2017].

authority so that it retains the trust of its data providers and data users (see Practices 6, 7, and 8). Such trust is fostered when a statistical agency has control over its IT resources, and there is no opportunity or perception that policy, program, or regulatory agencies could gain access to records of individual respondents.

A statistical agency also needs control over its IT resources to support timely and accurate release of official statistics, which are often produced under stringent deadlines. To streamline and increase the efficiency of the acquisition, upgrading, and use of IT resources across the federal government, the 2014 Federal Information Technology Acquisition Reform Act (FITARA) gave IT authority to the chief information officer (CIO) of each cabinet department. Recognizing the special needs of statistical agencies, guidance for FITARA implementation includes an attachment that spells out ways in which CIOs are expected to work together with statistical agency heads to ensure that IT decisions and policies of the CIO appropriately support the agencies. The attachment explicitly states that implementation of FITARA must ensure that statistical information acquired under a pledge of confidentiality is used only for statistical purposes (U.S. Office of Management and Budget, 2016b; see also Appendix A).

Authority to release statistical information and accompanying materials (including press releases) without prior clearance for the statistical content by department policy officials is essential so that there is no opportunity for or perception of political manipulation of any of the information.[38] When a statistical agency releases information publicly, a clear distinction should be made between the statistical information and any policy interpretations of it. Not even the appearance of manipulation for political purposes should be allowed. This essential requirement is one reason that statistical agencies are required by *Statistical Policy Directive No. 3* (U.S. Office of Management and Budget, 1985) to adhere to predetermined schedules for the public release of key economic indicators and to take steps to ensure that no person outside the agency has access to such indicators before the official release time except under carefully specified conditions.[39] *Statistical Policy Directive*

[38]The Energy Information Agency had its independence authorized in this regard in Section 205 of the Department of Energy Organization Act of 1977; 42 USC 7135(d)): "The Administrator [of EIA] shall not be required to obtain the approval of any other officer or employee of the Department in connection with the collection or analysis of any information; nor shall the Administrator be required, prior to publication, to obtain the approval of any other officer or employee of the United States with respect to the substance of any statistical or forecasting technical reports which he has prepared in accordance with law." The Bureau of Transportation Statistics had its independence similarly strengthened in Section 6017 of the 2015 Fixing America's Surface Transportation (FAST) Act (P.L. 114-94).

[39]The economic indicators are given to the chair of the Council of Economic Advisers as soon as they are available and, through the chair, to the President. At the determination of the issuing statistical agency, they may also be provided to others under strict conditions to ensure

No. 4 (U.S. Office of Management and Budget, 2008) requires agencies to develop and publish schedules for release of other important social and economic indicators and to announce and explain any changes in schedules as far in advance as possible (see Appendix A).

For press releases, *Statistical Policy Directive No. 4* (U.S. Office of Management and Budget, 2008) encourages statistical agencies to use them to publicize and thereby expand the dissemination of data to the public. The directive states: "statistical press releases must be produced and issued by the statistical agency and must provide a policy-neutral description of the data." Policy pronouncements must be issued separately by executive branch policy officials and not by the statistical agency, and "policy officials of the issuing department may review the draft statistical press release [solely] to ensure that it does not include policy pronouncements."

Statistical agencies should also have dissemination policies that foster regular, frequent release of major findings from the agency's programs to the public through the traditional media, the Internet, and other means. In these ways, an agency can guard against even the perception of political and other undue external influence that might inhibit or bias its operations.

LIMITS ON AUTHORITY

The authority for a statistical agency to decide on the data collected or compiled can never be without limits. Congress may specify particular data to be collected (e.g., data on job openings and labor turnover by the Bureau of Labor Statistics, data on family farms by the Economic Research Service and National Agricultural Statistics Service). For the decennial census, Congress requires an opportunity to review the proposed questions.

The OMB Office of Information and Regulatory Affairs, under the Paperwork Reduction Act, is responsible for designating a single data collection instrument for information wanted by two or more agencies. It is also responsible under the same act for reviewing all questionnaires and other instruments for the collection of data from 10 or more respondents to determine whether they meet applicable standards and are well justified (see Appendix A). The courts sometimes become involved in interpreting laws and regulations that affect statistical agencies, as in issues concerning data confidentiality, Freedom of Information Act requests, and the use of sampling in the population census.

The budgetary constraints on statistical agencies and OMB's review of data collections are ongoing. Other pressures depend, in part at least, on the relations between a statistical agency and the people or entities that have

there is no unauthorized dissemination or use. Executive branch officials outside the issuing agency may not comment on indicators until at least 1 hour after they are publicly released.

supervisory or oversight functions for them. Agencies need to enhance their skills in communicating to oversight groups the need for statistical series and in estimating the costs of statistical work in a credible manner. In turn, although it is standard practice for the secretary of a department or the head of an independent agency to have ultimate responsibility for all matters in the department or agency, the head of a statistical agency, for credibility, should be allowed full authority in professional and technical matters. For example, decisions to revise the methodology for calculating the Consumer Price Index, the gross domestic product, and the supplemental poverty measure[40] have been and are properly made by the relevant statistical agency heads or their designees.[41]

[40] See Interagency Technical Working Group on Developing a Supplemental Poverty Measure (2010:1, 3).

[41] For a somewhat different view, see Bradburn and Fuqua (2010), who discuss the potential dangers to statistical agency credibility in the development of indicators, defined as statistics used for assessment (e.g., the poverty rate). They argue for policy officials to establish normative aspects of indicators, such as the poverty threshold.

Practice 3: Use of Multiple Data Sources for Statistics That Meet User Needs

Federal statistical agencies and programs cannot be static. To provide information of continued relevance for public and policy use, they must regularly engage with a broad range of users to identify emerging needs and look for ways to develop new information that can serve broad purposes. To improve the quality and timeliness of their products, they must keep abreast of methodological and technological advances and be prepared to implement new procedures in a timely manner (see Practice 9). They must also continually seek ways to make their operations more efficient and less burdensome (see Practice 10).

Preparing for the future requires that agencies periodically assess the justification, scope, and frequency of existing data series, plan new or modified data series as required, and be innovative and open to new ways to improve their programs. Because of the decentralized nature of the federal statistical system, innovation often requires cross-agency collaboration (see Practice 13) and a willingness to implement different kinds of data collection efforts to answer different needs.

Two changes in policy and outlook can help statistical agencies foster the needed spirit of innovation. The first is to focus on the desired outputs of their programs by defining their primary business as that of providing relevant, accurate, and timely statistics obtained in a cost-effective manner. An output-oriented focus should help agencies justify and implement difficult decisions to modify or replace data collection and estimation programs that have lost their relevance, timeliness, or accuracy.

The second is to adopt as a matter of stated policy a paradigm of using multiple data sources to generate needed information as an expansion of the long-dominant paradigm of using probability sample surveys. This new paradigm, which federal statistical agencies are already embracing, recognizes the continued importance of surveys, both cross-sectional and longitudinal. At the same time, it explicitly recognizes the roles of administrative records and other third-party sources, along with the use of new methods for combining data from multiple sources, as key elements of a cost-effective strategy to serve users' needs.[42]

In considering new data collection, estimation, and dissemination strategies for the future, statistical agencies must be mindful of tradeoffs among relevance, accuracy, timeliness, comparability over time and with other data sources, transparency, costs, and respondent burden. It will not usually be possible to maximize all seven criteria at the same time, given constrained budgets, but using multiple data sources will enable statistical agencies to better address this challenge.

ROLES FOR SURVEYS

Many current statistical programs rely on well-established probability sampling methods that draw representative samples of a population, such as household members or business establishments, interview the sample units, and produce estimates that account for known errors in population coverage and missing data and have a quantifiable level of uncertainty from sampling variability. Box III.1 provides a brief history of probability sampling for federal statistics and lists examples of long-running federal surveys.

Declining rates of response over the past 30 years in the United States (and in other countries), however, are making it increasingly difficult to contain the costs of data collection with traditional surveys in ways that do not risk compromising the quality of the data (see, e.g., Brick and Williams, 2013; de Leeuw and de Heer, 2002).[43] User demands for timeliness and granularity of estimates also strain the ability of statistical agencies to respond using established survey techniques.

Survey researchers are actively seeking ways to maintain and improve both the quality and the cost-effectiveness of surveys. For example, more surveys are using multiple modes to facilitate response (Internet,

[42]See Lohr and Raghunathan (2017) on methods for combining survey and nonsurvey data with examples of applications.

[43]Lower response rates reduce the effective sample size and increase the sampling error of survey estimates; lower rates may also increase response bias in survey estimates.

Box III.1 Federal Surveys: Brief History and Examples

The probability sampling paradigm represented a quantum leap forward in providing cost-effective information on a variety of subjects when it was first introduced for federal surveys beginning in the late 1930s (Citro, 2014; Duncan and Shelton, 1978). For example, no longer did everyone in the United States have to answer a long battery of questions in the decennial census, since the use of a separate "long-form" questionnaire administered to a sample of the population could produce reliable estimates for the nation and smaller areas (see National Research Council, 2007b; the American Community Survey replaced the long-form sample in 2005). Moreover, probability surveys produced much more reliable results than surveys that did not use probability techniques.

Survey estimates (e.g., unemployment rates produced from the predecessor to the Current Population Survey in the early 1940s) quickly gained widespread public and policy acceptance. In the following decades, the federal government came to rely on probability surveys to collect information on a wide array of topics, as shown in the examples below (with the agency or program responsible for the survey).

Repeated Cross-Sectional Surveys

[Ask the same questions of new samples of respondents each month, year, or other reporting period; provide regularly updated estimates of key statistics.]

Household Surveys

Current Population Survey (CPS): 1940–present; monthly unemployment rate (BLS)*

CPS Annual Social and Economic Supplement: 1947–present; annual median income, poverty rate, health insurance coverage (U.S. Census Bureau)

Consumer Expenditure Survey: 1950, 1960–61, 1970–72; 1980–present; quarterly interviews, 2-week diaries; market-basket weights for the Consumer Price Index (BLS)*

National Health Interview Survey: 1957–present; annual health status, health insurance coverage (updated quarterly) (NCHS)*

National Crime Victimization Survey: 1972–present; annual estimates of crimes, both reported and not reported to the police (BJS)*

American Housing Survey: 1973–present; annual estimates of housing conditions from 1973–1981; biennial estimates since that time (PD&R)*

Business Establishment Surveys

Monthly Wholesale Trade Survey: 1940–present; principal economic indicator (U.S. Census Bureau)

Business R&D [Research and Development] and Innovation Survey: 1953–present; annual business expenditures on R&D (formerly Survey of Industrial R&D) (NCSES)*

National Hospital Care Survey: 1965–present; annual statistics on in-patient, out-patient, and emergency medical care (NCHS)

Commercial Buildings Energy Consumption Survey: periodically 1979–present; energy use for stores, malls, etc. (EIA)

Agricultural Resource Management Survey: 1996–present; annual statistics on farm practices and farm and farm household income (similar statistics collected back to 1975 in two surveys) (ERS, NASS)

Box III.1 (continued)

Panel Surveys
[Follow the same respondents over time to understand antecedents and consequences of changes in their circumstances]

Survey of Doctorate Recipients: 1979–present (NCSES)
National Longitudinal Survey of Youth: 1979–present; 1997–present (BLS)
Early Childhood Longitudinal Study, Kindergarten Class of 1988–89: 1998–2007 (NCES)

NOTES: *: Data collected by U.S. Census Bureau. BLS, Bureau of Labor Statistics; BJS, Bureau of Justice Statistics; EIA, Energy Information Administration; ERS, Economic Research Service; NASS, National Agricultural Statistics Service; NCES, National Center for Education Statistics; NCHS, National Center for Health Statistics; NCSES, National Center for Science and Engineering Statistics; PD&R, Office of Policy Development and Research, U.S. Department of Housing and Urban Development.

smartphone, telephone, mail, in person), as well as using paradata[44] to improve survey operations and facilitate "responsive" or "adaptive" survey designs (see National Research Council, 2013a).

Surveys should remain an important component of federal statistical agencies' portfolios for two major reasons: (1) some information is not readily ascertained except by asking questions; and (2) surveys can collect information on many characteristics at the same time, thereby permitting rich multivariate analysis. Yet the challenges to the survey paradigm make it essential to consider how use of other data sources can bolster the completeness, quality, and utility of statistical estimates while containing costs and reducing respondent burden (see National Academies of Sciences, Engineering, and Medicine, 2016, 2017b).

ROLES FOR ADMINISTRATIVE RECORDS

Administrative records include records of federal, state, and local government agencies that are used to administer a government program. Examples include U.S. Social Security Administration records of payroll taxes collected from workers and benefits paid out to retirees and other beneficiaries; state agency records of information provided by applicants for assistance programs and payments to applicants deemed eligible; and property tax records of local governments.

Administrative records are not generated probabilistically, as are surveys, but they are not unlike household or business censuses and can be evaluated

[44]Paradata are data about the source that are gathered in real time, such as the length of time to complete a survey.

in similar ways. Administrative records are designed to capture information for all instances of a specified population (e.g., program beneficiaries) according to a set of rules typically based in statute or regulation, and, like censuses, they may have omissions or duplications, and the variables in the records may differ in accuracy.[45] The records may also be stored in difficult-to-use formats, not well documented, or not provided on a timely basis. Acquiring the records requires negotiations with the custodial agency, and their contents may change when program rules change. Yet efforts to develop error profiles for administrative records (see Practice 9) and productive relationships with the custodial agency (see Practice 7) can have sizable payoffs, as is evident in several well-established statistical uses of records.

For example, administrative records are used to generate up-to-date population estimates by age, race and ethnicity, and gender. In turn, these estimates are used to adjust population survey weights for coverage errors and for many other purposes.[46] Tax records are used instead of questionnaires for the Census Bureau's economic censuses and surveys for nonemployer businesses. Administrative records are more and more used with survey data to produce model-based estimates with improved accuracy for small geographic areas or population groups.[47]

There are many other statistical uses that agencies should consider for administrative records, for which the investment in data collection has already been made. In some instances, records could improve the cost-effectiveness and data quality of current statistical programs (e.g., by substituting administrative records for survey questions). In other instances, they could add richness to the combined dataset (e.g., by appending administrative records variables to matched survey records).[48] Box III.2 presents examples of innovative uses of administrative records for federal statistics. National Research Council (2009e) provides a comprehensive strategy for using administrative records to improve income information in the Census Bureau's Survey of Income and Program Participation.

[45] For example, payments to beneficiaries may be more accurate than information provided at the time of application regarding a beneficiary's characteristics.

[46] See, e.g., National Research Council (2004a, 2007b).

[47] See, e.g., National Research Council (2000c,d), on the Census Bureau's Small-Area Income and Poverty Estimates (SAIPE) Program and recommended improvements.

[48] Extant matches include: (1) matches of Social Security earnings histories and Medicare benefits with the Health and Retirement Study and other surveys to analyze retirement decisions and the effect of medical care use on income security (see National Research Council, 1997a; National Research Council and Institute of Medicine, 2012); and (2) matches of employer and employee survey data with state employment security agency records in the Census Bureau's Longitudinal Employer-Household Dynamics Program to analyze business and employment dynamics (see National Research Council, 2007a). Access to matched datasets must be restricted to protect confidentiality (see Practice 8).

Box III.2 Uses of Administrative Records for Federal Statistics: Examples

Demographics and Socioeconomic Data The U.S. Census Bureau has built a record linkage infrastructure, based on assigning protected identification keys to census and survey records for matching with administrative records. One project (sponsored by ERS) linked SNAP records for New York State, 2010 census responses, and county-level labor market indicators, for research on how local labor market conditions affect the likelihood that SNAP participants leave the program. See https://www.census.gov/library/working-papers/2016/adrm/carra-wp-2016-10.html. Another project, the National Longitudinal Mortality Study (sponsored by NIH and NCHS), links selected data from the CPS and census records with death certificate information. See https://www.census.gov/did/www/nlms/about/index.html.

Education NCES has for many years combined surveys of secondary and postsecondary students with administrative data, including school transcripts and student loan records maintained by the U.S. Department of Education. See https://nces.ed.gov/statprog/handbook/index.asp.

Energy EIA introduced monthly statistics in March 2015 on crude oil transport by rail, based on waybill data from the U.S. Surface Transportation Board and administrative data from Canada's National Energy Board. See http://www.eia.gov/petroleum/transportation/methodology.pdf.

Health NCHS and CMS have a joint program to link Medicare records to NCHS surveys (e.g., the National Health and Nutrition Examination Survey). See https://www.cdc.gov/nchs/data/datalinkage/cms_medicare_methods_report_final.pdf. Beginning in 2012, NCHS replaced its separate surveys of nursing homes and other long-term care providers with the biennial National Study of Long-Term Care Providers; for nursing facilities, home health agencies, and hospices, NCHS now uses only administrative data from CMS. See https://www.cdc.gov/nchs/nsltcp/nsltcp_questionnaires.htm.

NOTE: All URL addresses were active as of April 2017. CMS, Centers for Medicare and Medicaid Services; CPS, Current Population Survey; EIA, Energy Information Administration; ERS, Economic Research Service; NCES, National Center for Education Statistics; NCHS, National Center for Health Statistics; NIH, National Institutes of Health; SNAP, Supplemental Nutrition Assistance Program.

Some uses may require not only state-of-the-art confidentiality protection techniques (see Practice 8), but also explicit legal authority (see "Toward the Paradigm of Multiple Data Sources," below).[49]

ROLES FOR NONTRADITIONAL DATA SOURCES

Statistical agencies are currently exploring the use of data sources, other than surveys and administrative records, that hold promise to improve the

[49]Statistics Canada, the national Canadian statistical agency, which has full access to administrative datasets, asks respondents to the Canadian Income Survey (and its predecessor) for permission to use income tax records in place of questions: see http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5200 [April 2017].

relevance, accuracy, and timeliness of federal statistics. These nontraditional data sources include, among others, data gleaned from relevant Internet websites (e.g., price quotes or social media postings), extracted from sensors (e.g., traffic cameras), and obtained from the private sector (e.g., scanner data on consumer purchases). Often, these sources generate large volumes of data that require data mining and other computationally intensive techniques for extracting information (see National Research Council, 2008a, esp. App. H).[50]

Some agencies are already using nontraditional data sources. For example, the Economic Research Service (in the U.S. Department of Agriculture) obtains expenditure data scanned by households from store receipts from a private vendor and has evaluated the quality of the data (Muth et al., 2016). The National Center for Health Statistics (NCHS), in its surveys of hospitals and other health care providers, obtains data from questionnaires, abstracts of samples of patient records, and providers' electronic medical care claim records.[51] See Box III.3 for other examples of current uses of nontraditional data sources.

Most nontraditional data sources present significant challenges to statistical agencies to evaluate the accuracy and error properties of the information. For example, harvesting website data to develop up-to-the-minute consumer price indexes[52] may offer significant timeliness and cost savings compared with traditional methods, but it is not clear how to adjust these data for consumer expenditures that occur off-line so that they accurately represent the universe of purchases. More generally, information that is taken from the Internet cannot usually be described or evaluated according to either a probability survey paradigm or a rules-based administrative records paradigm. For example, people who post items to sell on an auction website do not comprise any specified population. Another challenge of nontraditional data is that statistical agencies lack control over the consistency over time or among vendors or sites, so that deciding to rely heavily on such data sources carries high risks of compromising key time series if a vendor or site ceases operation or there are marked changes in data content or population coverage.

Yet in an era when data users expect timeliness and when budgets are constrained, statistical agencies should actively explore means by

[50]Such data are often referred to as "big data," which are characterized by high volume, velocity, and variety and require new tools and methods to capture, curate, manage, and process in an efficient way. See https://unstats.un.org/bigdata/ and https://unstats.un.org/unsd/statcom/doc15/BG-BigData.pdf [April 2017].

[51]See, for example, "What Does Participation in the NHCS [National Hospital Care Survey] Entail," and related frequently asked questions at http://www.cdc.gov/nchs/nhcs/faq.htm [April 2017].

[52]This is currently being done by the Billion Prices Project at the Massachusetts Institute of Technology; see http://bpp.mit.edu [April 2017].

Box III.3 Uses of Nontraditional Data for Federal Statistics: Examples

Health Insurance Claims Records BEA's health care satellite account, first released in 2015 for 2000–2010 and currently available through 2013, estimates spending by disease treatment rather than provider type, using a blend of Medical Expenditure Panel Survey data, Medicare claims records, and Truven Health MarketScan® Commercial Claims and Encounters Database. See https://bea.gov/national/health_care_satellite_account.htm.

Sales and Transactions Data Aggregators BEA and the Census Bureau's Economic Directorate are conducting pilot projects on the use of third-party aggregators of scanned store sales and credit card transactions to improve monthly retail trade estimates. These data are currently provided by the Monthly Retail Trade Survey, which has experienced significant declines in response rates. Initial results of the pilot projects show close correlations between the transactions data and survey data. The two agencies believe it may be possible to use the transactions data to enhance early estimates of consumer spending and produce estimates for metropolitan areas. See https://www.bea.gov/about/pdf/2016%20BEA%20Strategic%20Plan.pdf.

Web Scraping In the BLS consumer prices program, the web is scraped for characteristics of selected products (e.g., televisions, camcorders, cameras, washing machines) to use in models for quality adjustments. See https://www.bls.gov/cpi/cpihqaitem.htm.

NOTE: All URL addresses were active as of April 2017. BEA, Bureau of Economic Analysis; BLS, Bureau of Labor Statistics.

which nontraditional data sources can contribute to their programs. Such means could include: (1) augmenting information obtained from traditional sources; (2) replacing information elements previously obtained from traditional sources; (3) providing earlier estimates that are later benchmarked with traditional sources; and (4) analyzing information streams to identify needed changes (e.g., in types of jobs, education majors) in statistical classifications and survey questions.

Just as more and more surveys use multiple data collection modes, so a growing number of statistical programs will likely benefit from using multiple data sources, including nontraditional sources. To garner acceptance by policy makers and the public, statistical agencies should invest resources in documentation and user training and education. Agencies may need to "wall off" data series that are derived from nontraditional sources by labeling them as experimental or for research use until their statistical characteristics can be fully understood. If it is not possible to evaluate a nontraditional source sufficiently to establish its quality and suitability for inclusion in a statistical program, then a statistical agency should not use the data, although it may assist users by informing them of the problems with the source.

INTEGRATION AND SYNCHRONIZATION OF DATA ACROSS AGENCIES

Statistical agencies that collect similar information should pursue integration of their microdata records for specified statistical uses as another way to improve data quality, develop new kinds of information, and increase cost-effectiveness. One cost-effective approach is for a large survey to provide the sampling frame and additional content for a smaller, more specialized survey. Currently, the National Health Interview Survey of NCHS serves this function for the Medical Expenditure Panel Survey of the Agency for Healthcare Research and Quality. Similarly, the American Community Survey serves this function for the National Survey of College Graduates that the Census Bureau conducts for the National Center for Science and Engineering Statistics (see National Research Council, 2008b).

Another cost-effective approach is to synchronize or harmonize similar data held by different agencies. For example, both the Bureau of Labor Statistics (BLS) and the Census Bureau maintain business establishment lists. The lists derive from different sources (state employment security records for BLS and a variety of sources, including federal income tax records, for the Census Bureau). Research has demonstrated that synchronization of the lists would improve the accuracy of the information and the coverage of business establishments in the United States (National Research Council, 2006b, 2007a).

A major step toward synchronization was taken in the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) of 2002. The act authorized the synchronization of business data among the three principal statistical agencies that produce the nation's key economic statistics—the Bureau of Economic Analysis (BEA), BLS, and the Census Bureau (see Appendix B). However, synchronization of business establishment lists between BLS and the Census Bureau cannot currently be done because the Census Bureau is prohibited by law (Title 26 of the U.S. Code) from sharing with BLS (or BEA) any tax information of businesses or individuals that it may acquire from the Internal Revenue Service (IRS), even for statistical purposes.[53]

TOWARD THE PARADIGM OF MULTIPLE DATA SOURCES

There have been several initiatives in the past decade to further the use of multiple data sources by statistical agencies. The Federal Committee on Statistical Methodology established a Subcommittee on

[53]Efforts have been under way since CIPSEA was enacted to permit business data synchronization involving IRS records, but, to date, it has not occurred.

Administrative Records in 2008, subsequently renamed the Subcommittee on Administrative, Alternative, and Blended Data, which to date has developed examples and protocols for accessing, using, and evaluating administrative records (Federal Committee on Statistical Methodology, 2009, 2013). OMB guidance issued in 2014 states that federal administrative records as a matter of practice should be considered for federal statistics (U.S. Office of Management and Budget, 2014a; see also Appendix A). The United States participates in the United Nations Global Working Group on Big Data for Official Statistics, established in 2014.[54] The bipartisan 15-member commission created by the Evidence-Based Policymaking Commission Act of 2016 will report in September 2017 on such topics as how to integrate administrative and survey data and make them available for research and evaluation while protecting privacy and confidentiality; how data infrastructure, database security, and statistical protocols should be modified toward these ends; and whether a federal clearinghouse should be created for government survey and administrative data.[55]

In moving to expand their use of administrative records or other nontraditional data sources together with surveys, statistical agencies should consider at least six factors in assessing the benefits and costs:

(1) the need for upfront investment to facilitate the most effective approach to acquisition and use of administrative records or another nontraditional data source, accompanied by estimates of the likely longer term cost savings or reductions in respondent burden (or both) and timeliness or data quality improvements (or both);
(2) the protocols and criteria to follow to ensure full understanding by the statistical agency of the properties of a specific nonsurvey data source (e.g., its population coverage, frequency of updating): see the quality frameworks discussed in Practice 9;
(3) changes to established processing and estimation methods required to incorporate a nonsurvey source to maximize quality and timeliness of the resulting estimates: see Practice 10;
(4) the means by which the confidentiality of linked or augmented datasets can be protected while allowing access for research purposes: see National Research Council (2005b) and Practice 8;
(5) the requirements for expanded documentation, user outreach, and user education to assure acceptance of the resulting estimates and understanding of their strengths and limitations: see Practice 4; and

[54]See https://unstats.un.org/bigdata/ [April 2017].
[55]See www.cep.gov [April 2017].

(6) the risks to the availability, consistency, and quality of statistical series if a provider discontinues or changes a nontraditional data source in significant ways and how those risks can be mitigated.

In the United States, a challenge that must first be addressed is the legality and feasibility of acquisition of data sources not already "owned" by a statistical agency. National Academies of Sciences, Engineering, and Medicine (2017b:1) concluded that "legal and administrative barriers limit the statistical use of administrative datasets by federal statistical agencies." The problem is particularly acute for records for which states have custody; incentives (as some federal-state cooperative statistics programs provide) are likely needed for states to share their records for federal statistical purposes. For nontraditional data sources, such as scanner data and the like, the problems are more that such data vary greatly in their fitness for use in official statistics.

National Academies of Sciences, Engineering, and Medicine (2017b) recommended stepped-up efforts led by the Interagency Council on Statistical Policy to coordinate research to evaluate different nontraditional data sources, develop appropriate estimation methods based on combining data, develop quality metrics for combined estimates, implement state-of-the-art data privacy protection techniques, and similar matters (see Practice 13). The report also recommended (p. 2) that "a new entity or an existing entity should be designated to facilitate secure access to data for statistical purposes to enhance the quality of federal statistics." The report argues that the current system—whereby individual statistical agencies must negotiate separately with other federal agencies, state agencies, and the private sector—is burdensome on all parties and precludes the necessary economies of scale for resolving the common challenges to realizing the potential of multiple data sources for statistics that meet user needs.

Practice 4: Openness about Sources and Limitations of the Data Provided

A STATISTICAL AGENCY SHOULD BE OPEN about the strengths and limitations of its data, taking as much care to understand and explain how its statistics may fall short of accuracy as it does to produce accurate data. Metadata, or "data about data," are a critical element of data dissemination and curation (see Practice 5). All data contain some uncertainty and error, which does not mean the data are wrong, but that they need to be used with understanding of their limitations.[56]

Openness requires that data releases from a statistical program include a full description of the purpose of the program; the methods and assumptions used for data collection, processing, and estimation; what is known and not known about the quality and relevance of the data; sufficient information for estimating variability and other errors in the data, when possible; appropriate analysis methods that take account of variability and other sources of error; and the results of research on the methods and data. Openness also means that a statistical agency should describe how decisions on methods and procedures were made for a data collection program and provide ready access to research results that entered into such decisions. Such transparency is essential for credibility with data users and trust of data providers.

Openness about data limitations requires much more than providing estimates of sampling error for surveys or basic attributes for administrative

[56] Manski (2015) points to a need for fuller measurement and communication of uncertainty in official statistics.

records or other nonsurvey data sources. In addition to a discussion of aspects that statisticians characterize as nonsampling errors—such as coverage errors, nonresponse errors, measurement errors, and processing errors—it is valuable to have a description of the concepts used and how they relate to the major uses of the data. Descriptions of the shortcomings of and problems with the data should be provided in sufficient detail to permit a user to take them into account in analysis and interpretation. Descriptions of how the data relate to similar data collected by other agencies should also be provided, particularly when the estimates from two or more surveys or other data sources exhibit large differences that may have policy implications.

On occasion, the objective of presenting the most accurate data possible may conflict with user needs for timely information. When concerns for timeliness prompt the release of preliminary estimates (as is done for some economic indicators), consideration should be given to the frequency of revisions and the mode of presentation from the point of view of the users as well as the issuers of the data. Agencies that release preliminary estimates need to educate the public about differences among preliminary, revised, and final estimates.

To meet their responsibility to users for openness, some statistical agencies in the 1990s developed detailed "quality profiles" for major surveys, including the American Housing Survey (Chakrabarty and Torres, 1996); the Residential Energy Consumption Survey (Energy Information Administration, 1996); the Schools and Staffing Survey (Kalton et al., 2000); and the Survey of Income and Program Participation (U.S. Census Bureau, 1998). Previously the Federal Committee on Statistical Methodology (1978) developed a quality profile for employment as measured in the Current Population Survey. These profiles documented what was and was not known about errors in estimates as a help to experienced users and agency personnel (see Federal Committee on Statistical Methodology, 2001; National Research Council, 1993a, 2007b). As print publications, however, they quickly became outdated and were rarely updated given the burden on agency staff. Today, the Internet enables easier maintenance and updating of quality profile-type information (e.g., separate web pages for major types of error).[57]

[57] The Census Bureau posts basic quality indicators for the American Community Survey for the nation and states; these include sample size, population coverage, household response rates, and item response rates. See https://www.census.gov/acs/www/methodology/sample_size_and_data_quality/ [April 2017]. The Bureau of Labor Statistics issued a prototype data quality report for the Consumer Expenditure Quarterly and Diary Surveys, which it characterized as the "first in a series of iterations towards a single reference source on a comprehensive set of CE data quality metrics that are timely and routinely updated." The metrics provided in the prototype report refer to data for 2009–2013. Available: https://www.bls.gov/cex/ce_methodology.htm#dqreports0 [April 2017].

Error (conversely, "accuracy") is not the only dimension of quality of concern to statistical agencies and their data users. Building on a seminal paper by Brackstone (1999), many statistical agencies around the world have adopted "quality frameworks," which are typologies of key attributes or dimensions to use in systematically measuring, improving, and documenting data quality. For example, the Eurostat (2000) framework includes relevance, accuracy, timeliness and punctuality, accessibility and clarity, comparability (across time and geography), and consistency (with other series). Biemer et al. (2014) decompose "accuracy" into sampling error (where applicable) and seven types of nonsampling error. Daas et al. (2012) and Federal Committee on Statistical Methodology (2013) address quality attributes for administrative records-based data series.

In the United States, the Information Quality Act of 2000 required all federal agencies to develop written guidelines for how they ensure the quality of the information they disseminate to the public. Using a framework developed by the Interagency Council on Statistical Policy, individual statistical agencies developed quality guidelines (see Practice 9 and Appendix A). However, the guidelines, in response to the legislation, are process oriented (e.g., indicating how to request correction of a datum) and are not quality frameworks as described above. Federal statistical agencies should consider adopting and implementing quality frameworks for their programs.

An important aspect of openness not addressed in quality frameworks concerns the treatment of mistakes that are discovered subsequent to data release. Openness means that an agency has an obligation to issue corrections publicly and in a timely manner. The agency should use not only the same dissemination avenues to announce corrections that it used to release the original statistics, but also additional vehicles, as appropriate, to alert the widest possible audience of current and future users of the corrections in the information. Agencies should be proactive in seeking ways to alert known and likely users of the data about the nature of a problem and the corrective action that it is taking or that users should take.

Overall, agencies should treat the effort to provide information on the quality, limitations, and appropriate use of their data as an essential part of their mission. Such information should be made available in ways that are readily accessible to all known and potential users (see National Research Council, 1993a, 1997b, 2007b).

Practice 5: Wide Dissemination of Data

A STATISTICAL AGENCY SHOULD STRIVE CONTINUALLY for the widest possible dissemination of the data it compiles, consistent with its obligations to protect confidentiality. Data should be disseminated in formats that are accessible and accompanied by documentation that is clear and complete. Dissemination should be timely, and information should be made readily available on an equal basis to all users. Agencies should have data curation policies and procedures in place so that data are preserved, fully documented, and accessible for use in future years.[58]

Planning for dissemination should be undertaken from the viewpoint that the public has contributed the data elements and paid for the data collection and processing. In return, the information should be accessible in ways that make it as useful as possible to the largest number of users—for decision making, program evaluation, scientific research, and public understanding.

An effective dissemination program is comprised of a wide range of elements:

- It should have an established publications policy, which describes, for a data collection program, the types of reports and other data releases to be made available, the formats to be used, the audience to be served, and the frequency of release.[59]

[58] Data curation involves the management of data from collection and initial storage to archiving (or deletion should the data be deemed of no further use—e.g., a data file that represents an initial stage of processing). The purpose of data curation is to ensure that information can be reliably retrieved and understood by future users.

[59] *Statistical Policy Directive No. 3* (U.S. Office of Management and Budget, 1985) prescribes a yearly calendar of firm, fixed release dates for key economic indicators; *Statistical Policy Directive*

- It should have a variety of avenues for disseminating information about data availability and upcoming releases. Those avenues should be chosen to reach as broad a public as reasonably possible—including, but not limited to, an agency's Internet website, conference exhibits and programs, newsletters and journals, email address lists, and social media and blogs. A statistical agency should also regularly communicate major findings to the media, which helps build the expectation of statistical agency releases without political interference or partisan spin.
- The public release of data should occur in a variety of forms (suitably processed to protect confidentiality), so that information can be accessed by users with varying skills and needs for retrieval and analysis. Useful data products include not only understandable maps, graphs, indicators, and tables on statistical agency websites, but also public-use microdata samples (PUMS) and other computer-readable files with richly detailed information.
- For data that are not publicly available, agencies should provide access for research and other statistical purposes through restricted modes that protect confidentiality, such as protected data enclaves and contractual licensing agreements.
- All data releases should be accompanied by careful and complete documentation or metadata, including explanatory material to assist users in appropriate interpretation (see Practice 4). For a complex database (such as a PUMS file), user training should be provided through webinars, online tutorials, and sessions at appropriate conferences.[60]
- The program should include archiving policies that guide which data to retain, where they are to be archived (with the National Archives and Records Administration, or an established archive maintained by an academic or other nonprofit institution, or both), and how they are to be accessible for future secondary analysis while protecting confidentiality.[61]

PUBLIC DATA PRODUCTS

Data release of aggregate statistics may take the form of regularly updated time series, cross-tabulations of aggregated characteristics of

No. 4 (U.S. Office of Management and Budget, 2008) lays out best practices for dissemination of other federal statistics (see Appendix A).

[60]For example, the Bureau of Labor Statistics holds an Annual CE [Consumer Expenditure Surveys] Microdata Users' Workshop: see https://www.bls.gov/cex/csxannualworkshop.htm [April 2017].

[61]See, for example, http://www.data-archive.ac.uk/ [April 2017].

respondents, analytical reports, interactive maps and charts, and brief reports of key findings. Such products should be readily accessible through an agency's website, which should also make available more detailed tabulations in formats that are downloadable from the website. Agencies should take care in designing their websites to make it as easy as possible for users to locate and access information, testing accessibility and usability with a range of users.

A relatively new way for agencies to expand public use of their aggregate statistics is by providing selected data through application programming interfaces (APIs) to developers who, in turn, build custom applications for the Internet, smartphones, and similar media. For example, the Census Bureau's APIs include neighborhood population characteristics and county-level information on business activity.[62]

Yet another form of dissemination involves access to individual-level microdata files, which make it possible to conduct in-depth research in ways that are not possible with aggregate data. PUMS files can be developed for general release. Such files contain data for samples of individual respondents that have been processed to protect confidentiality by deleting, aggregating, or modifying any information that might permit individual identification.[63]

RESTRICTED DATA ACCESS

While honoring their obligation to be proactive in seeking ways to provide data to users, statistical agencies must be vigilant in their efforts to protect against disclosure of data obtained under a pledge of confidentiality (see Practices 7 and 8). The stunning improvements over the past three decades in computing speed, power, and storage capacity, the growing availability of information from a wide range of public and private sources on the Internet, and the increasing richness of statistical agency data collections have increased the risk that individually identifiable information can be obtained through reidentification of data thought to have been suitably protected (see Doyle et al., 2001; National Academies of Sciences, Engineering, and Medicine, 2017b:Ch. 5; National Research Council, 2003b, 2005b:Ch. 5). In response, statistical agencies may have to scale back the detail that is provided in PUMS files or other public data products.

As an alternative to public access, statistical agencies have pioneered several methods of restricted access. One method is to provide or arrange

[62]See http://www.census.gov/data/developers/about.html [April 2017]; see also National Research Council (2012).

[63]For a review of methods for confidentiality protection of PUMS files, see Federal Committee on Statistical Methodology (2005).

for a facility on the Internet to allow researchers to analyze restricted microdata to suit their purposes, with safeguards so that the researcher is not seeing the actual records and cannot obtain any output, such as too-detailed tabulations, that could identify individual respondents.[64] A second method, pioneered by the National Center for Education Statistics (NCES), is to grant licenses to individual researchers to analyze restricted microdata at their own sites: such licenses require that the researchers agree to follow strict procedures for protecting confidentiality and accept liability for penalties if confidentiality is breached.[65] A third method is to allow researchers to analyze restricted microdata at a secure site, such as one of the Federal Statistical Research Data Centers (FSRDCs) currently located at two dozen universities and research organizations around the country. The FSRDC network began as a Census Bureau initiative and now includes data from other agencies.[66] Statistical agencies should continually seek to enlarge their suite of restricted access methods and, for each, to reduce as much as possible the cost, time, and burden of access for users.

[64]The Data Enclave of NORC at the University of Chicago is such a facility: see http://www.norc.org/Research/Capabilities/Pages/data-enclave.aspx [April 2017]. It provides secure access by researchers to selected microdata sets of the Economic Research Service, the National Center for Science and Engineering Statistics, and several other federal agencies and private foundations. NCES provides similar functionality for access to its data sets: see, e.g., https://nces.ed.gov/datalab/ [April 2017].

[65]For NCES's licensing procedures and terms, see https://nces.ed.gov/statprog/instruct.asp [April 2017].

[66]See https://www.census.gov/fsrdc [April 2017].

Practice 6: Cooperation with Data Users

Within the limitations of its confidentiality procedures, a statistical agency should seek to provide maximum access to its data on an equal basis to all. Having data accessible for a wide range of uses, including for secondary analysis by external researchers, increases the return on the investment in data collection (National Research Council, 1985, 2005b). Ensuring equal access requires avoiding release of standard data products to any users in advance of their being released to all users.[67] In addition, agencies that prepare special tabulations of their data on request at cost should list all such tabulations on their website and make them available to other users for the cost of reproduction.[68]

Researchers and other users frequently request data from statistical agencies for specific purposes. Agencies should have guidance materials on their websites to help users find what they need and that specify the procedures for contacting staff who can further assist them in accessing and understanding relevant data. Statistical agencies should view these services as a part of their dissemination activities.

Once statistical data are made public, they will be used in numerous ways, including ways not originally envisaged, and by many different types of users, ranging from government officials to activists and students. An effective statistical agency continually strives to learn about its data users and obtain input from them on its programs, including what they want and how they use data. The goal is to make an agency's data as relevant,

[67] Such advanced release is proscribed in *Statistical Policy Directives Nos. 3* and *4* (U.S. Office of Management and Budget, 1985, 2008).

[68] See https://www.census.gov/population/www/cen2010/spec-tab/stp_faq.html [April 2017] for an example of such a policy.

accurate, timely, and accessible as possible for the broadest possible range of uses and users.

Keeping abreast of the interests of current and potential new users requires continual attention to changes in policy issues, social and economic conditions, and technologies for data access. To do so, statistical agencies should work with professional associations, institutes, universities, and scholars to determine current and emerging needs of research communities. They should also work with relevant associations and other organizations to determine the needs of business and industry, with state and local governments, and with user groups focused on federal statistics generally or on particular statistical programs.

Statistical agencies can proactively explore the needs of users through advisory committees, focus groups, analysis of requests and Internet activity, formal surveys of users, and user conferences.[69] To optimize the data products from a statistical program (e.g., tabulations), statistical agencies should seek advice from knowledgeable users. They should inform these and other users of decisions about products.

A statistical agency shows cooperation with data users not only by learning about their needs for information and inviting their input on data products, but also by facilitating their access to and ability to use data through well-designed websites and other dissemination vehicles, careful and complete documentation, and user training adapted to varying skills and needs. In this regard, it is critical that agencies beta test website design changes and new data products with a range of users outside the agency. Agency staff can provide useful input in this regard (see Practice 10), but there is no substitute for testing with outside users.

In developing and implementing new and nontraditional methods and data sources to produce information, statistical agencies should involve policy makers and other key data users at an early stage so that they understand an agency's criteria and decision process and have an opportunity to provide input. Statistics that are based on models (e.g., for small geographic areas) or that use nontraditional data sources will likely require more detailed explanations of their benefits and limitations than is usually provided for more traditional products and sources (see Practice 3).

[69] See, e.g., https://www.cdc.gov/nchs/events/2015nchs/index.htm [April 2017], which summarizes the 2015 National Conference on Health Statistics organized by the National Center for Health Statistics and held every 2–3 years.

Practice 7: Respect for the Privacy and Autonomy of Data Providers

To maintain a relationship of respect and trust with survey participants and other data providers, a statistical agency should respect their privacy. One aspect of this practice is to minimize the burden imposed on them. A statistical agency should also respect and enhance respondents' autonomy—in the sense of their ability to make informed decisions—when they are asked to participate in a voluntary program to collect data. Even for mandatory data collections, such as the decennial and economic censuses, a statistical agency respects its respondents by giving them as much information as possible about the reasons for the collection and making it as easy as possible for them to respond. The principles and practices of respect apply not only to individuals asked to participate in a survey, but also to representatives of organizations (e.g., businesses, state and local governments) asked to participate in a survey and to custodians of already collected data, such as administrative records, who are asked to share their data for statistical purposes.

Clear policies and effective procedures for respecting privacy and enhancing autonomy are critical to maintaining the quality and comprehensiveness of the data that federal statistical agencies provide to policy makers and the public. Part of the challenge for statistical agencies is to develop effective means of communicating not only the agency's privacy protection procedures and policies, but also the importance of the data being collected for the public good.

RESPECTING PRIVACY IN SURVEYS

To promote trust and encourage accurate response from survey respondents, it is important that statistical agencies respect their privacy by reducing, to the extent possible, the time and effort required to respond and by giving them adequate information with which to decide if a survey is worthy of response—that is, to give their informed consent (see below). Thus, when individuals or organizations are asked to participate in a survey, they should be told whether it is mandatory or voluntary, how the data will be used, and who will have access to the data.[70] They should also be informed of the likely duration of an interview and, if the survey involves more than one interview, how many, and whether they will be asked to consult records.

To reduce the burden of replying to surveys (see National Academies of Sciences, Engineering, and Medicine, 2016; National Research Council, 2013a:Ch. 4), statistical agencies should write clear questions that fit respondents' common understanding; minimize the intrusiveness of questions and explain why intrusive-seeming questions serve important purposes; allow alternative modes of response when appropriate (e.g., Internet, smartphone); and use administrative records or other data sources, if sufficiently complete and accurate, to provide some or all of the needed information. In surveys of businesses or other organizations, agencies should seek to obtain information directly from the organization's records and so minimize the need for duplicate responses to multiple requests. Agencies should also use qualified, well-trained data collection staff who can establish trust with respondents.

Agencies can explore ways in which participation can be made easier for respondents and result in more accurate data through such means as focus group discussions and surveys. Many agencies apply the principles of cognitive psychology to questionnaire design (National Research Council, 1984) and make use of paradata, in order to increase the accuracy of the survey results and make more efficient use of the time and effort of respondents (National Research Council, 2013a:Ch. 3).[71] Some agencies

[70]Statistical agencies typically use general language about the intended uses and users of the survey data, such as that the data will be used for statistical and research purposes. A good example of helpful language for a respondent that is more specific is the information provided as an attachment to the advance letter soliciting participation in the Bureau of Labor Statistics' National Longitudinal Survey of Youth 1979 (see https://www.bls.gov/nls/handbook/2005/nlshc9.pdf [April 2017].

[71]Paradata are data collected automatically and in real time about the data collection process, such as the number of seconds to complete each question. Paradata can be used after the fact to evaluate data quality and alter design features or, increasingly, in adaptive or responsive designs, to make real-time decisions about nonresponse follow-up efforts.

thank respondents for their cooperation by providing them with summaries of the information after a survey is compiled.

It is likely that increasing concerns about privacy are contributing to the recent declines in survey response rates. At a time when people are inundated with requests for information from public and private sources, there are publicized instances of identity theft and other abuses of confidential information on the Internet, and individual information is used for terrorism-related investigatory or law enforcement purposes, it is not surprising that some people are choosing not to participate in censuses and surveys, even when the questions appear noninvasive and the data are collected for statistical purposes under a pledge of confidentiality.[72] Statistical agencies need to devote resources to understanding respondents' privacy and confidentiality concerns.

Several laws address privacy and confidentiality concerns government-wide. For example, the 2002 E-Government Act requires agencies to publish privacy impact assessments (PIAs) whenever "initiating a new collection of information [that] includes any information in an identifiable form." The purpose of PIAs is to ensure there is no collection, storage, access, use, or dissemination of identifiable information that is not both needed and permitted. The 2015 Federal Cybersecurity Enhancement Act enjoins agencies to use sophisticated computer network scanning technology to detect and prevent hacking of records by unauthorized individuals. This enhanced privacy protection is now referred to in statistical agencies' confidentiality protection pledges (see Appendix A).

The challenge is to devise effective strategies for communicating these and other privacy and confidentiality policies and practices to respondents. Such strategies appear to be more necessary—and more challenging—than ever before.

PROTECTING AND RESPECTING THE AUTONOMY OF HUMAN RESEARCH PARTICIPANTS

Collecting data from individuals for research purposes with federal funds falls under a series of regulations, principles, and best practices that the federal government has developed over a period of more than 50 years (see National Research Council, 2003b, 2014c). The pertinent regulations, which have been adopted by 11 departments and 6 agencies, are known as the "Common Rule" (45 CFR 46). The Common Rule regulations (most recently revised in January 2017 to take effect in January 2018) require that researchers adequately protect the privacy of human participants and

[72]For a literature review of public opinion on privacy in the wake of the September 11, 2001, terrorist attacks, see National Research Council (2008a:App. M).

maintain the confidentiality of data collected from them, minimize the risks to participants from the data collection and analysis, select participants equitably with regard to the benefits and risks of the research, and seek the informed consent of individuals to participate (or not) in the research.[73] Under the regulations, most federally funded research involving human participants must be reviewed by an independent institutional review board (IRB) to determine that the design meets ethical requirements for protection.[74]

Not all federal statistical agencies consider all of their information collections to be subject to IRB review. They should, nonetheless, strive to incorporate the spirit of the Common Rule in the design and operation of all activities that involve data collection from individual respondents.

For activities for which IRB approval is required for data collection, an agency should work proactively with the IRB to determine how best to apply the regulations in ways that do not unnecessarily inhibit participant responses. For example, implied consent is typically used for mail and telephone surveys of the general population: written documentation does not generally provide any added protection to the respondent and could reduce participation. Statistical agencies should seek ways—such as sending an advance letter—to furnish information to potential respondents that will help them make an informed decision about whether to participate. Such information should include the planned uses of the data and their benefits to the public. Statistical agencies should then respect the contributions of respondents by compiling the data and making them accessible to users in convenient forms. A statistical agency has an obligation to publish statistical information from the data it collects unless it finds the results invalid, in which case the agency should publish an explanation of why results from a data collection were not of sufficient quality to be made available.

RESPECTING THE PROVIDERS OF ALTERNATIVE DATA

An essential element of moving to a new paradigm of using multiple data sources for federal statistics is for a statistical agency to develop procedures and practices that respect the time and other constraints of organizations that provide alternative (nonsurvey) data. Statistical agencies should not assume the cooperation of either other federal agencies or nonfederal providers.

[73] "Federal Policy for the Protection of Human Subjects," 82 *Federal Register* 7149 (January 19, 2017) (available: https://www.federalregister.gov/d/2017-01058 [April 2017]); see also Appendix A.

[74] For information about the Common Rule and certification of IRBs by the Office for Human Research Protections in the U.S. Department of Health and Human Services, see http://www.hhs.gov/ohrp [April 2017].

In working with federal agencies that hold useful administrative records, statistical agencies should plan to cooperate, communicate, and coordinate with them on a continuing basis, as urged in Hendriks (2012). The process should begin with a properly developed memorandum of understanding (see U.S. Office of Management and Budget, 2014a:App. B). Through a continuing relationship of mutual respect and trust, a statistical agency can better understand the strengths and limitations of a custodial agency's data and help identify and implement improvements in the data that are useful to both agencies.[75]

An important consideration in using administrative records is whether informed consent of the individuals or organizations that provided their information to the custodial agency is required. The U.S. Office of Management and Budget (2014a:App. A) identifies situations in which the statistical use of administrative records may qualify under the "routine use" exception of the Privacy Act or may qualify once the statistical agency has developed or modified an appropriate System of Records Notice (see Appendix A). In some instances, it may be necessary to obtain new consent of the original data providers.

[75]Working with nonfederal providers will also require relationships of mutual respect and trust. The specific issues for such relationships, including whether payment may be required, are likely to be provider-specific and so are not discussed here. A coordinated approach on the part of OMB and the statistical agencies would be helpful.

Practice 8: Protection of the Confidentiality of Data Providers' Information

To earn the respect and trust of individual respondents and other data providers, such as businesses and other government agencies, a statistical agency must be able to offer a credible pledge of confidentiality for the information it collects for statistical purposes. Providers must trust that the data they share with a statistical agency will neither be made available for any administrative, regulatory, law enforcement, or other targeted proceeding that might harm individuals or organizations nor be hacked or otherwise intruded on by unauthorized people inside or outside the statistical agency.

A credible pledge of confidentiality for individual and organizational responses is considered essential to encourage high response rates and accuracy of responses from survey participants.[76] Moreover, if individual participants have been assured of confidentiality, disclosure of identifiable information about them would violate the principle of respect for persons even if the information is not sensitive and would not result in any social, economic, legal, or other harm (see Practice 7; National Research Council, 2003b:Ch. 5). For sensitive administrative data obtained from another government agency, there must be a credible pledge of confidentiality in a properly formulated memorandum of understanding or other authorizing document.

[76] Reviews of research on how confidentiality and privacy concerns may affect response rates include Hillygus et al. (2006) and National Research Council (1979, 2004a:Ch. 4, 2013a:Ch. 1). Not all statistically useful data are collected under a pledge of confidentiality—see "Definition of a Federal Statistical Agency" in Part I above.

LEGAL PROTECTIONS

Some agencies, including the Census Bureau and the National Agricultural Statistics Service, have long had legislative protection for ensuring respondent confidentiality.[77] However, prior to the passage of the Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA), other agencies, including the Bureau of Labor Statistics, had to rely on strong statements of policy, legal precedents in court cases, or customary practices (see Gates, 2012; Norwood, 1995). Agencies that did not have legal protection for their practices were at risk of having their policies overturned by judicial interpretations of legislation or executive decisions that would require the agency to disclose identifiable data collected under a pledge of confidentiality.[78]

The passage of CIPSEA was a landmark event in the history of confidentiality protection for statistical data (see Appendix A).[79] Subtitle A provides a statutory basis for protecting the confidentiality of all federal data collected for statistical purposes under a confidentiality pledge, including but not limited to data collected by statistical agencies. CIPSEA states that individually identified information obtained under a confidentiality pledge cannot be disclosed to persons other than the agency's employees without the respondent's informed consent and the agency head's authorization and only when another law (e.g., Title 13 of the U.S. Code) does not prohibit the disclosure. It also provides penalties for employees who knowingly disclose identifiable statistical information (up to 5 years in prison, up to $250,000 in fines, or both). Principal statistical agencies and recognized statistical units may also designate contractors and outside researchers as "agents," who may have access to specified confidential information, such as microdata in a restricted access environment, if they agree to be subject to the penalties for disclosure.

Confidentiality protection for statistical data is now on a much firmer legal footing across the federal government than prior to CIPSEA, with one exception. Section 508 of the USA PATRIOT Act of 2001 (P.L. 107-56) amended the National Center for Education Statistics (NCES) Act of 1994 to allow the U.S. Attorney General (or an assistant attorney general) to apply to a court to obtain any "reports, records, and information (including individually identifiable information) in the possession" of NCES that

[77] For the Census Bureau, such legislation was first enacted in 1929 in Title 13 of the U.S. Code; for the National Agricultural Statistics Service, such provisions are in Title 7.

[78] For an example involving the Energy Information Administration, see National Research Council (1993b:185–186).

[79] CIPSEA was preceded by a Federal Statistical Confidentiality Order, issued by OMB on June 27, 1997. It told respondents who provide statistical information to specified agencies that their responses would be held in confidence and would not be used against them in any government action "unless otherwise compelled by law" (see Appendix A).

are considered relevant to an authorized investigation or prosecution of domestic or international terrorism. Section 508 also removed penalties for NCES employees who furnish individual records under this section. This exclusion for NCES has not been invoked, and its possible effect on survey response rates has not been tested, but its existence is not helpful for the mission of statistical agencies and the need for trust on the part of data providers.

Both the perception and reality of agencies' confidentiality protection may be affected by departmental initiatives to consolidate data processing and storage to bolster computer and network security in the federal government, improve the cost-effectiveness of information technology development and maintenance, and protect against cyberattacks. Such initiatives are required by, respectively, the 2002 Federal Information Security Management Act, the 2014 Federal Information Technology and Acquisition Reform Act, and the 2015 Federal Cybersecurity Enhancement Act (see Practice 2 and Appendix A). An effective statistical agency will work with its department on approaches to computer security that recognize the need for the agency to control the processing and storage of data collected for statistical purposes under a pledge that the data will be not accessible for other departmental purposes, such as regulation or enforcement.

CONFIDENTIALITY AND DATA ACCESS

Although confidentiality protection is essential for gaining and keeping trust with data providers, a statistical agency's fundamental mission is to disseminate information widely. Consequently, there is a tension between the goals of protection and access (see Practice 5). Agencies cannot guarantee zero risk of disclosure for public-use products. And even if all use were restricted to secure enclaves (which would not be desirable), there would still be the risk that an employee or agent might breach confidentiality inadvertently (or advertently) or that an agency's computer systems could be hacked. The challenge to statistical agencies is to devise appropriate methods and procedures to minimize the disclosure risk and to continually improve on methods as the threats to confidentiality change.

Computerized data processing has enabled statistical agencies for more than 60 years to make available a large volume of public-use products. Those products include detailed tabulations and public-use microdata samples that are safeguarded against disclosure by such basic methods as suppression of small cells and removal of obvious identifiers. The advent of distributed computing and the Internet greatly increased—and continues to increase—the disclosure risks for such products because of the potential for reidentifying individual respondents through data linkage with the vast

amounts of possibly related information on the web. Risks are increased when data for small groups are tabulated, when the same data are tabulated in a variety of ways, or when public-use microdata samples are released with highly detailed content, particularly when surveys are longitudinal and follow the same respondents over time. Risks are also increased when surveys include linked administrative data or collect biomarkers from blood samples or other physiological measures, as is increasingly being done (National Research Council, 2001a, 2010b).

Statistical agencies have responded to increased disclosure risks by pioneering not only more sophisticated techniques to further protect public-use products, but also procedures to restrict access (e.g., through secure enclaves) to qualified researchers and other agents when a public-use product has been deemed too risky to release (see Practice 5).[80] However, agencies have yet to develop formal approaches to confronting the ever more difficult challenges in this area. Such work can benefit from close attention to the work of computer scientists, who are developing conceptual frameworks for assessing disclosure risk, along with sophisticated privacy protective techniques for modifying datasets to preserve analytic capabilities for a given privacy guarantee. Although these techniques are not fully mature, they represent a means toward a more structured and less ad hoc way of measuring the effectiveness of alternative disclosure risk reduction procedures and implementing them as appropriate (see National Academies of Sciences, Engineering, and Medicine, 2017b:Ch. 5).

[80] For reviews of alternative procedures for providing data access while maintaining confidentiality protection, see National Academies of Sciences, Engineering, and Medicine (2017b:Ch. 5) and National Research Council (1993b, 2000a, 2003b, 2005b, 2009d, 2010f).

Practice 9: Commitment to Quality and Professional Standards of Practice

There is a wide range of actions a statistical agency should undertake to ensure the quality of its products and practices:

- develop strong staff expertise in the subject areas relevant to the agency's mission, in the theory and practice of statistics, and in data collection, processing, analysis, and dissemination techniques;
- keep abreast of and use modern statistical theory and sound statistical and computational practice in all technical work;
- publish and implement formal quality standards;
- maintain quality assurance programs to improve data quality and the processes of compiling, editing, and analyzing data;
- develop an understanding of the validity and accuracy of the agency's data and how to convey the resulting measures of quality (both uncertainty and bias) in comprehensible ways to users;
- document concepts, definitions, and data collection methods and discuss possible sources of error in data releases to the public (see Practice 4); and
- develop continuing relationships with appropriate professional organizations in statistics and relevant subject-matter areas.

The best guarantee of high-quality data is a strong professional staff, which includes experts in the subject-matter fields covered by the agency's program, experts in statistical methods and techniques, and experts in data collection, processing, and other operations (see Practice 11). A major function of an agency's leadership is to strike a balance among these staff

and promote working relationships that make the agency's program as productive as possible, with each group of experts contributing to the work of the others.

An effective statistical agency keeps up to date on developments in theory and practice that may be relevant to its program—for example, new techniques for imputing missing data (see, e.g., National Research Council, 2004a:App. F, 2010e) or for combining data from more than one source and estimating error in the resulting statistics (see National Academies of Sciences, Engineering, and Medicine, 2017b:Ch. 5); new technologies for data collection, processing, and dissemination; and new kinds of and uses for paradata (see, e.g., National Research Council, 2013a). Paradata, not only from data collection processes, but also from tracking how users work with an agency's data products on its website, can help improve methods in each of these important areas.

An effective statistical agency is also alert to social and economic changes that may call for changes in the concepts or methods used in particular datasets.[81] The need for change often conflicts with the need for comparability with past data series, and the latter need can easily dominate consideration of proposals for change. Agencies have the responsibility to manage this conflict by initiating more relevant series or revising existing series to improve quality while providing information to compare old and new series, such as was done when BLS revised the treatment of owner-occupied housing in the Consumer Price Index.[82]

An effective statistical agency devotes resources to developing, implementing, and inculcating standards for data quality and professional practice. Although a long-standing culture of data quality contributes to professional practice, an agency should also seek to develop and document standards through an explicit process. The existence of explicit standards and guidelines, regularly reviewed and updated, facilitates training of new in-house staff and contractors' staffs. *Statistical Policy Directive No. 2* (U.S. Office of Management and Budget, 2016a), on survey standards and guidelines, is helpful in that it covers every aspect of a survey from planning

[81] Reviews of concepts underlying important statistical data series, which have identified areas in which change is needed, include: poverty (National Research Council, 1995, 2005c); cost-of-living and price indexes (National Research Council, 2002); "satellite" accounts for nonmarket activities—e.g., home production, volunteerism (National Research Council, 2005a); food insecurity and hunger (National Research Council, 2006a); usual residence in the decennial census and surveys (National Research Council, 2006c); disability (National Research Council, 2009b); health satellite accounts (National Research Council, 2010a); medical care economic risk and burden (National Research Council and Institute of Medicine, 2012); happiness, suffering, and other dimensions of experience (National Research Council, 2013b); civic engagement and social cohesion (National Research Council, 2014b); and innovation (National Research Council, 2014a:Ch. 4; National Academies of Sciences, Engineering, and Medicine, 2017a).

[82] See, for example, Gillingham and Lane (1982).

through data release (see Appendix A).[83] It recommends that agencies develop additional, more detailed standards that focus on their specific statistical activities.[84]

To ensure the quality of its data collection programs and data releases, an effective statistical agency has not only formal quality assurance programs (e.g., well-developed methods for detecting outliers and other errors in raw data and errors from editing and other data processing steps), but also mechanisms and processes for obtaining both inside and outside reviews (see Practice 12). Such reviews should address various aspects of an agency's operations, including the soundness of the data collection and estimation methods and the completeness of the documentation of the methods used and the error properties of the data. For individual publications and reports, formal processes are needed that incorporate review by agency technical experts and, as appropriate, by technical experts in other agencies and outside the government.[85]

Finally, an effective statistical agency builds strong ties with relevant professional associations. It encourages professional staff to participate in relevant associations to refresh their human capital and develop networks of experts from other statistical agencies, academia, and the private sector. It also uses professional associations as one source of advice on ways to keep its data collection programs as relevant, accurate, timely, and cost-effective as possible.

[83]Data quality guidelines of statistical agencies in other countries are also helpful (see, e.g., Statistics Canada, 2009; United Kingdom Office for National Statistics, 2007).

[84]For examples, see National Center for Education Statistics (2012) and Bureau of Transportation Statistics (2005). These standards work within but go well beyond the broad data quality guidelines adopted by statistical agencies in response to the 2000 Information Quality Act (see Appendix A).

[85]See Appendix A for a description of OMB guidelines for peer review of scientific information; reviews at a program or agency-wide level are considered in Practice 12.

Practice 10: An Active Research Program

A STATISTICAL AGENCY SHOULD HAVE a research program that is relevant to its activities. Because a small agency may not be able to afford an appropriate research program, agencies should collaborate and share research results and methods (see Practice 13). Agencies can also augment their staff resources for research by using outside experts.

At least two major components should be part of a statistical agency's research program: (1) research on the substantive issues for which the agency's data are compiled, taking care not to take policy positions; and (2) research to evaluate and improve statistical methods and operational procedures, such as data processing flow. In addition, research should be conducted to understand how an agency's information is used, both inside and outside the government, for policy analysis, decision making, and public understanding (see Practice 6).

Research on data uses and users can contribute to future improvements in the concept and design of data collections and the format of data products. For example, public-use files of statistical microdata were developed in response to the analytic needs of government and academic researchers. Beginning with an understanding of the variety of uses and users of an agency's data, more in-depth research on the policy uses of an agency's information might, for example, explore the use of data in microsimulation and other economic models that are used in decision making (see National Research Council, 1991a,b, 1997a, 2000b, 2001b, 2003a, 2010d).

SUBSTANTIVE RESEARCH AND ANALYSIS

A statistical agency should include staff with responsibility for conducting objective substantive analyses of the data that the agency compiles, such as analyses that assess trends over time or compare population groups. Substantive analyses provided by an agency should be relevant to policy by addressing topics of public interest and concern. However, such analyses should not include positions on policy options or be designed to reflect any particular policy agenda (see Martin, 1981; Norwood, 1975; Triplett, 1991).

The existence and output of an analytical staff can contribute not only to the knowledge base in the applicable subject areas, but also to the credibility, relevance, accuracy, timeliness, and cost-effectiveness of the agency's data collection programs. Benefits that a strong subject-matter staff bring to a statistical agency include:

- Agency analysts are able to understand the need for and purposes of the data from a statistical program and how the data will be used. Such information must be available to refine the design and methods an agency is using to produce the data.
- Agency analysts have access to the complete microdata and so are better able than outside analysts to understand and describe the limitations of the data for analytic purposes and to identify errors or shortcomings in the data that can lead to subsequent improvements.
- Substantive research by agency analysts can benefit from and help reinforce an agency's credibility through its commitment to openness and maintaining independence from political influence.
- Substantive research can assist in formulating an agency's data program, suggesting changes in priorities, concepts, and needs for new data or discontinuance of outmoded or little-used series.

An agency's subject-matter analysts should be encouraged and have ample opportunity to build networks with analysts in other agencies, academia, the private sector, other countries, and relevant international organizations. Analysts should also be encouraged and have ample opportunity to present their work at relevant conferences and in working papers and refereed journal articles. The goal is for the agency to have widely recognized expertise in the subject areas in its mission.

The leaders of a statistical agency should take steps to ensure that the agency's subject-matter analysts and its methodological and operational staff are able to interact in a constructive manner. Overcoming barriers to communication is essential so that insights from subject-matter analysis can be translated effectively into improved data collection program design, methodology, and operations.

RESEARCH ON METHODOLOGY AND OPERATIONS

It is important for statistical agencies to be innovative in the methods used for data collection, processing, estimation, analysis, and dissemination, with the goal of improving data accuracy and timeliness and operational efficiency and reducing respondent burden. Careful evaluation of new methods is required to assess their benefits and costs in comparison with current methods and to determine effective implementation strategies, including the development of methods for bridging time series before and after a change in procedures.

Research on methodology and operational procedures must be ongoing. Currently, some of the important topics for research include:

- determining best uses of paradata to optimize costs and timeliness of data collection and estimation and accuracy of results (see National Research Council, 2013a);
- addressing challenges for computer-assisted interviews, which have included lengthy times to implement questionnaire changes and difficulty in providing adequate documentation of questionnaire content and pathways (see National Research Council, 2003c);
- understanding and minimizing mode effects on quality when obtaining data in two or more different ways (Internet, mail, telephone, and face-to-face response; see National Research Council, 2007b);
- improving the adequacy of the documentation of Internet data products and guidance for users with a wide range of analytical skills and understanding (see National Research Council, 2012);
- developing new methods of confidentiality protection (see National Academies of Sciences, Engineering, and Medicine, 2017b); and
- accelerating the use of multiple data sources by developing measures of error for alternate sources and identifying optimal ways to combine them to achieve such goals as reducing burden and costs and improving accuracy and timeliness, recognizing that it is likely not possible to achieve improvements on all dimensions at once (see Practices 3 and 9).

With regard to conceptualization and measurement of error, to the extent possible, statistical agencies should work to adapt the concept of total survey error, which has guided the design of probability surveys, to nonsurvey data sources. The total survey error framework includes bias (nonsampling error) and variance (sampling error). It can be adapted to administrative records alone or in combination with surveys, so long as the statistical agency can obtain sufficient information on sources of error in the records (e.g., coverage). There are nontraditional data sources for which

measuring error alone or in combination with surveys will be difficult, if not infeasible, and for which it will be necessary to label any statistics as experimental (see Practice 3).

It is noteworthy that many current practices in statistical agencies were developed through research they conducted or obtained from other agencies. Federal statistical agencies, frequently in partnership with academic researchers, pioneered the use of statistical probability sampling, the national economic accounts, input-output models, and other analytic methods. The U.S. Census Bureau pioneered the use of computers for processing the census. Several statistical agencies use academic principles of cognitive psychology—a research strand dating back to the early 1980s (see National Research Council, 1984)—to improve the design of questionnaires, the clarity of data presentation, and the ease of use of electronic data collection and dissemination tools. History has shown repeatedly that methodological and operations research can lead to large productivity gains in statistical activities at relatively low cost (see, e.g., Citro, 2016; National Research Council, 2010c).

An effective statistical agency actively partners with the academic community for methodological research. It also seeks out academic and industry expertise for improving data collection, processing, and dissemination operations. For example, a statistical agency can learn techniques and best practices for improving software development processes from computer scientists (see National Research Council, 2003c, 2004d). An effective agency also learns from and contributes to methodological research of statistical agencies in other countries and relevant international organizations.

Statistical agency management should take steps to ensure that methodological research staff are able to interact constructively with operational staff so that improvements to operations can be readily identified and implemented. Agency leaders should also strongly support methodological research and feasibility testing for major data collection programs, through such means as a methods panel that is operated in parallel with the agency's main program. This kind of testing is essential so that a program does not become locked into methods and procedures that are increasingly out of date and, at the same time, to assess new methods in a test environment before they are put into production.

Practice 11: Professional Advancement of Staff

A STATISTICAL AGENCY should recruit, develop, and support professional staff who are committed to the highest standards of quality work, professional practice, and professional ethics. To develop and maintain a high-caliber staff, a statistical agency needs to recruit qualified people with relevant skills for efficient and effective operations, including subject-matter experts in fields relevant to its mission (e.g., demographers, economists), statistical methodologists who specialize in data collection and analysis, and other skilled staff (e.g., computer specialists). Having sufficient in-house staff with the required types of expertise is as critical as having adequate budget resources for enabling a statistical agency to carry out its mission.

To retain and make the most effective use of its staff, an agency should provide opportunities for work on challenging projects in addition to more routine, production-oriented assignments. An agency's personnel policies, supported with sufficient resources, should enable staff to extend their technical capabilities through appropriate professional and developmental activities, such as attendance and participation in professional meetings, participation in relevant training programs, rotation of assignments, and involvement in collaborative activities with other statistical agencies.

An agency should also seek opportunities to reinforce the commitment of its staff to ethical standards of practice. Such standards are the foundation of an agency's credibility as a source of relevant, accurate, and timely information obtained through fair treatment of data providers and data users.

An effective federal statistical agency has personnel policies that encourage the development and retention of a strong professional staff who

are committed to the highest standards of quality work for their agency and in collaboration with other agencies. There are several key elements of such policies:

- The required levels of technical and professional qualifications for positions in the agency are identified, and the agency adheres to these requirements in recruitment and staff development. Position requirements take account of the different kinds of technical and other skills, such as supervisory skills, that are necessary for an agency to have a full range of qualified staff. Qualified staff include: statisticians and survey methodologists; experts in relevant subject-matter areas; experts in leading-edge technologies and methods for information storage, processing, estimation, data quality assessment, confidentiality protection, dissemination, and curation; and experts in management of complex, technical operations.
- Continuing technical education and training, appropriate to the needs of their positions, is provided to staff through in-house training programs and opportunities for external education and training.
- Position responsibilities are structured to ensure that staff have the opportunity to participate, in ways appropriate to their experience and expertise, in research and development activities to improve quality of data and cost-effectiveness of agency operations.
- Professional activities, such as publishing in refereed journals and presentations at conferences, are encouraged and recognized. Such presentations should include technical work in progress, with appropriate disclaimers. Participation in relevant statistical and other scientific associations, including leadership positions, is encouraged to promote interactions with researchers and methodologists in other organizations that can advance the state of the art. Such participation is also a mechanism for disseminating information about an agency's programs, including the sources and limitations of the data provided (see Practice 4). Guidance from the Office of Science and Technology Policy issued in 2010 stresses the importance of participation in professional activities as a means of ensuring a culture of scientific integrity in federal agencies (see Appendix A).
- Interaction with other professionals inside and outside the agency is fostered through opportunities to participate in technical advisory committee meetings, establish and be active in relevant listservs and blogs, interact with contract researchers and research consultants on substantive matters, interact with visiting fellows and staff detailed from other agencies, take assignments with other relevant statistical, policy, or research organizations, and regularly receive new assignments within the agency.

- Participation in cross-agency collaboration efforts, such as the Federal Committee on Statistical Methodology and its subcommittees, is supported. Such participation not only benefits the professional staff of an agency, but also contributes to improving the work of the statistical system as a whole (see Practice 13).
- Accomplishment is rewarded by appropriate recognition and by affording opportunities for further professional development. The prestige and credibility of a statistical agency is enhanced by the professional visibility of its staff, which may include establishing high-level nonmanagement positions for highly qualified technical experts.

An effective statistical agency considers carefully the costs and benefits—both monetary and nonmonetary—of using contractor organizations, not only for data collection, as most agencies do, but also to supplement in-house staff in other areas, such as carrying out methodological research.[86] Outsourcing can have benefits, such as: providing experts in areas in which the agency is unlikely to be able to attract highly qualified in-house staff (e.g., some information technology functions), enabling an agency to handle an increase in its workload that is expected to be temporary or that requires specialized skills, and allowing an agency to learn from best industry practices. However, outsourcing can also have costs, including that agency staff become primarily contract managers and less qualified as technical experts and leaders in their fields.

An effective statistical agency maintains and develops a sufficiently large number of in-house staff, including mathematical statisticians, survey researchers, subject-matter specialists, and information technology experts, who are qualified to analyze the agency's data and to plan, design, carry out, and evaluate its core operations, so that the agency maintains the integrity of its data and its credibility in planning and fulfilling its mission. At the same time, statistical agencies should maintain and develop staff with the expertise necessary for effective technical and administrative oversight of contractors.

An effective statistical agency has policies and practices to instill the highest possible commitment to professional ethics among its staff, as well as procedures for monitoring contractor compliance with ethical standards. When an agency comes under pressure to act against its principles—for example, if it is asked to disclose confidential information for an enforcement purpose or to support an inaccurate interpretation of its data—it must be

[86]Only the Bureau of Labor Statistics and the Census Bureau maintain their own interviewing staffs. The National Agricultural Statistics Service contracts with the National Association of State Departments of Agriculture for field interviewing staff, and other agencies contract with the Census Bureau or private survey contractors.

able to rely on its staff to resist such actions as contrary to the ethical principles of their profession.

An effective agency ensures that its staff are aware of and have access to such statements of professional practice as those of the American Association for Public Opinion Research (2015), the American Statistical Association (2016), and the International Statistical Institute (2010), as well as to the agency's own policies and practices regarding such matters as the protection of confidentiality, respect for privacy, and standards for data quality. It endeavors in other ways to ensure that its staff are fully cognizant of the ethics that must guide their actions in order for the agency to maintain its credibility as a source of objective, reliable information for use by all.

Practice 12: A Strong Internal and External Evaluation Program

Statistical agencies should have regular, ongoing evaluations of their major statistical programs and their overall portfolio of programs. Reviews of major data collection programs and their components should consider how to produce relevant, accurate, and timely data in the most cost-effective manner possible and whether there are ways to improve cost-effectiveness by combining data from multiple sources.[87] Reviews of an agency's portfolio should consider ways to reduce duplication, fill gaps, and adjust priorities so that the overall portfolio is as relevant as possible to the information needs of policy makers and the public. Such evaluations should include internal reviews by staff and external reviews by independent groups.

Statistical agencies that fully follow practices on using multiple data sources (Practice 3), openness (Practice 4), wide dissemination of data (Practice 5), commitment to quality and professional standards (Practice 9), and an active research program (Practice 10) will likely be in a good position to make continuous assessments of and improvements in the relevance, quality, and efficiency of their data collection systems. Yet even the best

[87] OMB issued a proposed addendum, Section 10: Performance Review, to *Statistical Policy Directive No. 4* in October 2016 for public comment (see Appendix A). The comment period has closed, and comments are under review. The addendum would require statistical agencies and recognized statistical units to submit annual performance reviews for their key statistical products, focused on such aspects as accuracy of the data, completeness of documentation, and timeliness (see Appendix A). *Statistical Policy Directive No. 3* (U.S. Office of Management and Budget, 1985) already requires statistical agencies to examine accuracy and timeliness of key economic indicators on a 3-year cycle.

functioning agencies will benefit from an explicit program of internal and independent external evaluations to obtain fresh perspectives.

EVALUATING QUALITY, RELEVANCE, EFFICIENCY

Evaluation of data quality for a continuing survey or any kind of data collection program begins with regular monitoring of quality indicators that are readily available to users. For surveys, such monitoring includes unit and item response rates, population coverage rates, and measures of sampling error. In addition, in-depth assessment of quality on a wide range of dimensions—including sampling and nonsampling errors across time and among population groups and geographic areas—needs to be undertaken on a periodic basis and the results made public (see Practices 4 and 9, and National Research Council, 2007b).

Research on methods to improve data quality may cover such areas as alternative methods for imputing values for missing data, alternative question wordings to reduce respondent reporting errors (based on cognitive methods), and alternative sources of data and ways for combining them to enhance quality. Methods for such research may include the use of "methods panels" (small samples for which experiments are conducted by using alternative procedures and questionnaires), matching with administrative records, and simulations of sensitivity to alternative procedures. The goal of the research should be the development of better methods that are feasible and cost-effective.

In ongoing programs for which it is disruptive to implement improvements on a continuing basis, a common practice is to undertake major research and development activities at intervals of 5, 10, or more years. Agencies should ensure, however, that the intervals between major research and development activities do not become so long that data collection programs deteriorate in quality, relevance, and efficiency. Regular, well-designed program evaluations, with adequate budget support, are key to ensuring that data collection programs do not deteriorate. Having a set schedule for research and development efforts will enable data collection managers to ensure that the quality and usefulness of their data are maintained and help prevent that increasingly less optimal procedures are locked in over time.

In addition to quality, relevance of an agency's data collection programs needs to be assessed. The question of relevance is whether the agency is "doing the right thing" in contrast to whether the agency is "doing things right." Relevance should be assessed not only for particular programs or closely related sets of programs, but also for an agency's complete portfolio

in order to assist it in making the best choices among program priorities given the available resources.

Communicating closely with stakeholders and important user constituencies—through such means as regular meetings, workshops, conferences, and other activities—is important to ensuring relevance (see Practice 6). Including other federal statistical colleagues in this communication, both as users and as collaborators, can be valuable.

Statistical agencies commonly find it difficult to discontinue or scale back a particular data series, even when it has largely outlived its usefulness relative to other series, because of objections by users who have become accustomed to it. In the face of limited resources, however, discontinuing a series is preferable to across-the-board cuts in all programs, which would reduce the accuracy and usefulness of both the more relevant and less relevant data series. Regular internal and external reviews and a documented priority-setting process or framework can help an agency not only reassess its priorities, but also develop the justification and support for changes to its portfolio.

Finally, statistical agencies should review their programs for efficiency and cost-effectiveness.[88] Federal statistics as a public good represent a legitimate call on public resources, and statistical agencies in turn are properly called on to analyze the costs of their programs on a continuing basis to ensure the best return possible on tax dollars. For this purpose, statistical agencies should develop complete, informative models for evaluating costs of current procedures and possible alternatives and follow best practice for design of statistical production processes. One excellent guide to best practices is the Generic Statistical Business Process Model of the United Nations Economic Commission for Europe. First developed in 2008 and most recently updated in 2013 (version 5), this model is designed to enable statistical agencies to describe production processes in a coherent way, compare processes within and among organizations, and make better decisions on production systems and allocation of resources (UNECE High-Level Group for the Modernisation of Official Statistics, 2013).

[88] "Efficiency" is generally defined as an ability to avoid waste (of materials, energy, money, time) in producing a specified output. "Cost-effectiveness" connotes a broader, comparative look at inputs and outputs to assess the most advantageous combination. ("Cost-benefit" analysis attempts to add monetary values to outputs.) In the context of federal statistical programs, cost-effectiveness analysis would assess the costs of conducting a program for different combinations of desired characteristics, such as improved accuracy or timeliness and reduced burden on respondents.

TYPES OF REVIEWS

Regular program reviews should include a mixture of internal and external evaluation. Agency staff should set goals and timetables for internal evaluations that involve staff who do not regularly work on the program under review. Independent external evaluations should also be conducted on a regular basis, the frequency of which should depend on the importance of the data, how quickly the phenomena being measured change, and how quickly respondent behavior and data collection technology that may adversely affect a program change.

In a world in which people and organizations are increasingly less willing to respond to surveys, it becomes increasingly urgent to have more frequent evaluations to determine whether there are alternative data sources to surveys with better quality. Agencies should seek outside reviews not only of specific programs, but also of program priorities and quality practices across their entire portfolio.

External reviews can take many forms. They may include recommendations from advisory committees that meet at regular intervals (typically, every 6 months). However, advisory committees should never be the sole source of outside review because the members of such committees rarely have the opportunity to become deeply familiar with agency programs. External reviews can also take the form of a "visiting committee," following the model of the National Science Foundation (NSF),[89] or a special committee or panel established by a relevant professional association or other recognized group.[90]

[89] For links to evaluations of NSF programs, see https://www.nsf.gov/od/oia/activities/cov/ [April 2017].

[90] Examples include an evaluation of the National Center for Education Statistics (NCES) by the National Institute of Statistical Sciences (2017) and numerous evaluations by the National Research Council (e.g., National Research Council, 2009a, which reviewed the statistical programs of the Bureau of Justice Statistics).

Practice 13: Coordination and Collaboration with Other Statistical Agencies

The U.S. FEDERAL STATISTICAL SYSTEM consists of many agencies in different departments, each with its own mission and subject-matter focus (see the history section in Part I and Appendix B). Yet these agencies have a common interest in serving the public need for credible, relevant, accurate, and timely information gathered as efficiently and fairly as possible. Moreover, needed information may often span the mission areas of more than one statistical agency: for example, both the Bureau of Labor Statistics and the National Center for Education Statistics have programs that relate to education and employment outcomes of the population. Consequently, statistical agencies should not and do not conduct their activities in isolation. An effective statistical agency actively seeks opportunities to conduct research and carry out other activities in collaboration with other statistical agencies to enhance the value of its own information and that of the system as a whole. Such collaboration is essential for smaller statistical agencies with limited resources and equally important for larger agencies so that they do not overlook useful innovations outside their own agency.

When possible and appropriate, federal statistical agencies should collaborate not only with each other, but also with policy, research, and program agencies in their departments, with state and local statistical agencies, and with foreign and international statistical agencies. Such collaborations can serve many purposes, including: standardization of concepts, measures, and classifications (see, e.g., National Research Council,

2004b,c, and Appendix A); augmentation of available information for cross-national and subnational comparisons (see, e.g., National Research Council, 2000c,d); identification of useful new data sources and data products; and improvements in many aspects of statistical program design and methods.

ROLE OF THE U.S. OFFICE OF MANAGEMENT AND BUDGET

The responsibility for coordinating statistical work in the federal government is specifically assigned to the Office of Information and Regulatory Affairs (OIRA) in the U.S. Office of Management and Budget (OMB) (see Appendix A). The Statistical and Science Policy (SSP) Office in OIRA, typically with the assistance of interagency committees, reviews concepts of interest to more than one agency (e.g., gender and race/ethnicity classifications) and oversees the development and periodic revision of standard classification systems (of industries, metropolitan areas, etc.). It also considers methods and data sources that should be widely adopted, such as the statistical use of administrative records, and works to establish common goals and norms on major statistical issues, such as confidentiality protection. SSP staff also consult with other parts of OMB on statistical budgets and, by reviewing statistical information collections as well as the statistical programs of the government as a whole, identify data gaps, programs that may be duplicative, and areas in which interagency cooperation might lead to greater efficiency and added utility of data.[91] SSP also is responsible for coordinating U.S. participation in international statistical activities.[92]

SSP established and contributes to the interagency Federal Committee on Statistical Methodology (FCSM), which issues guidelines and recommendations on statistical issues common to a number of agencies, typically by working through subcommittees.[93] FCSM also hosts research conferences and statistical policy seminars that facilitate professional interaction and development.[94] SSP encourages the Committee on National Statistics at the National Academies of Sciences, Engineering, and Medicine to serve as an

[91]SSP annually prepares a compilation of information on the full range of federal statistical programs; the latest volume, known as the "Blue Book," is for fiscal 2017 (U.S. Office of Management and Budget, 2017).

[92]SSP was formerly the Statistical Policy Office: it was renamed to reflect added responsibilities with respect to the 2001 Information Quality Act standards and guidelines, OMB's guidance on peer review planning and implementation, and evaluations of science underlying proposed regulatory actions (see Appendix A).

[93]See https://fcsm.sites.usa.gov/ [April 2017].

[94]Presentations from the most recent FCSM research conference and statistical policy seminar, respectively, are available at: https://fcsm.sites.usa.gov/reports/research/2015-research/ [April 2017] and http://www.copafs.org/seminars/fcsm2014policy.aspx [April 2017].

independent adviser and reviewer of federal statistical activities. Finally, SSP chairs the Interagency Council on Statistical Policy as a mechanism for statistical agency heads to work together with OMB to coordinate federal statistical programs for the common good (see Appendix B).[95]

FORMS OF INTERAGENCY COLLABORATION

Interagency collaboration and coordination takes many forms, some multilateral, some bilateral. Many result from common interests in specific subject areas, such as economic statistics, statistics on people with disabilities, or statistics on children or the elderly. Current interagency collaborative efforts include a systemwide initiative to facilitate the statistical uses of administrative records and other nonsurvey data sources;[96] a multiagency program for research on survey methodology administered by the National Science Foundation's Division of Social and Economic Sciences on behalf of the federal statistical agencies; a collaboration of the Bureau of Labor Statistics and the Census Bureau to update the supplemental poverty measure; and continued work on internationally comparable measures of disability led by the Washington Group on Disability Statistics established under the United Nations Statistical Commission (for a description, see U.S. Office of Management and Budget, 2017:73–80).

A common bilateral arrangement is an agreement of a program agency to provide administrative data to a statistical agency to use as a sampling frame, a source of classification information, a summary compilation to check (and possibly revise) preliminary sample results, and a source with which to improve imputations for survey nonresponse, reduce variability in estimates for small geographic areas, or substitute for survey questions. The Census Bureau, for example, uses Schedule C tax information from the Internal Revenue Service in place of surveys for millions of nonemployer businesses. Such practices improve statistical estimates, reduce costs, and eliminate duplicate requests for information from the same respondents.

In other arrangements, federal statistical agencies engage in cooperative data collection with state statistical agencies to let one collection system satisfy the needs of both. A number of such joint systems have been developed, notably by the Bureau of Labor Statistics, the National Agricultural Statistics Service, the National Center for Education Statistics, and the National Center for Health Statistics.

Another example of a joint arrangement is one in which a statistical agency contracts with another to conduct a survey, compile special

[95]The ICSP was given a statutory basis in the 1995 reauthorization of the Paperwork Reduction Act; see Appendix A.

[96]This activity is facilitated by a new FCSM Subcommittee on Administrative, Alternative, and Blended Data (see Practice 3).

tabulations, or develop models. Such arrangements make use of the special skills of the supplying agency and facilitate the use of common concepts and methods. The Census Bureau conducts many surveys for other agencies; both the National Center for Health Statistics and the National Agricultural Statistics Service receive funding from other agencies in their departments to support their survey work; and the National Center for Science and Engineering Statistics receives funding from agencies in other departments to support several of its surveys (see U.S. Office of Management and Budget, 2017:Table 2).

The major federal statistical agencies are also concerned with the international comparability of statistics. Under the leadership of SSP, they contribute to the deliberations of the United Nations Statistical Commission, the Organisation for Economic Co-operation and Development, and other international organizations; participate in the development of international standard classifications and systems; support educational activities that promote improved statistics in developing countries; and learn from and contribute to the work of established statistical agencies in other countries in such areas as survey methodology, record linkage, confidentiality protection techniques, and data quality standards.

Two continuing collaborative efforts in the United States for providing statistical information to the public in a broad area of interest are the Federal Interagency Forum on Aging-Related Statistics (Aging Forum) and the Federal Interagency Forum on Child and Family Statistics (Child Forum). The Aging Forum was established in the mid-1980s by the National Institute on Aging, with the National Center for Health Statistics and the Census Bureau, to improve the quality and usefulness of federal data on aging and to inform the public, policy makers, and researchers about trends for this important population. It was reorganized in 1998 to include 6 new agencies, and it currently includes 16 statistical and program agencies.[97] The Aging Forum produces an indicators chart book (first published in 2000) (see Federal Interagency Forum on Aging-Related Statistics, 2016).

The Child Forum was formalized in a 1994 executive order to foster collaboration in the collection and reporting of federal data on children and families. Its membership currently includes 23 statistical and program agencies, and its chart books (e.g., Federal Interagency Forum on Child and Family Statistics, 2016) describe the condition of America's children.[98] Those conditions include changing population and family characteristics, the environment in which children are living, and indicators of well-being in the areas of economic security, health, behavior, social environment, and education.

[97] See https://agingstats.gov/ [April 2017].
[98] See https://www.childstats.gov/ [April 2017].

Clearly, no single statistical or program agency alone could have produced reports of those two forums. Working together in this way, federal statistical agencies contribute to presenting data in a form that is more relevant to policy concerns and thereby strengthen the statistical system overall.

CHALLENGES AND REWARDS FOR COLLABORATION

Collaborative activities, such as sharing and integrating data compiled by different statistical and program agencies, standardizing concepts and measures, reducing unneeded duplication, and working together on methodological challenges, invariably require effort to overcome differences in agency missions and operations. Yet with constrained budgets and increasing demands for more relevant, accurate, and timely statistical information, collected at reduced costs and burden, the importance of proactive collaboration and coordination among statistical agencies cannot be overstated. To achieve the most effective integration of their work for the public good, agencies must be willing to take a long view, to strive to accommodate each other, and to act as partners in the development of statistical information for public use. The rewards of effective collaboration can be not only data that are more efficiently obtained, of higher quality, and more relevant to policy concerns, but also a stronger, more effective statistical system as a whole.

References for Part III

All URL addresses valid as of April 2017. *Please note that, as of February 2017, U.S. Office of Management and Budget documents previously on the main whitehouse.gov site were relocated to an archived site. Wherever possible, a citation is provided to a stable site such as federalregister.gov.*

American Association for Public Opinion Research (2015). *AAPOR Code of Professional Ethics and Practices*. Revised 11/30/2015. Oakbrook Terrace, IL: American Association for Public Opinion Research. Available: https://www.aapor.org/Standards-Ethics/AAPOR-Code-of-Ethics.aspx.

American Statistical Association (2016). *Ethical Guidelines for Statistical Practice*. Alexandria, VA: American Statistical Association. Available: http://www.amstat.org/ASA/Your-Career/Ethical-Guidelines-for-Statistical-Practice.aspx.

Biemer, P., D. Trewin, H. Bergdahl, and J. Lilli (2014). A system for managing the quality of official statistics. *Journal of Official Statistics 30*(3), 381–442. Available: https://doi.org/10.2478/jos-2014-0022.

Brackstone, G. (1999). Managing data quality in a statistical agency. *Survey Methodology 25*(2), 139–149. Available: https://unstats.un.org/unsd/dnss/docViewer.aspx?docID=194.

Bradburn, N. M. and C. J. E. Fuqua (2010). Indicators and the federal statistical system: An essential but fraught partnership. *The ANNALS of the American Academy of Political and Social Science 630*, 89–108.

Brick, J. M. and D. Williams (2013). Explaining rising nonresponse rates in cross-sectional surveys. *The ANNALS of the American Academy of Political and Social Science 645*(1), 36–59. Available: http://journals.sagepub.com/doi/full/10.1177/0002716212456834.

Bureau of Transportation Statistics (2005). *BTS Statistical Standards Manual*. Washington, DC: U.S. Department of Transportation. Available: https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/subject_areas/statistical_policy_and_research/bts_statistical_standards_manual/index.html.

Chakrabarty, R. and G. Torres (1996). *The American Housing Survey—A Quality Profile*. Current Housing Reports No. H121/95-1. Office of Policy Development and Research, U.S. Department of Housing and Urban Development; U.S. Census Bureau, U.S. Department of Commerce. Washington, DC: U.S. Government Printing Office. Available: https://www.census.gov/content/dam/Census/programs-surveys/ahs/publications/h12195-1.pdf.

Citro, C. F. (2014). From multiple modes for surveys to multiple data sources for estimates. *Survey Methodology 40*(2), 137–161. Available: http://www.statcan.gc.ca/pub/12-001-x/2014002/article/14128-eng.htm.

Citro, C. F. (2016). The US federal statistical system's past, present, and future. *Annual Review of Statistics and Its Application 3*(1), 347–373. Available: http://dx.doi.org/10.1146/annurev-statistics-041715-033405.

Daas, P. J. H., S. J. L. Ossen, M. Tennekes, and J. M. S. Burger (2012). *Evaluation and Visualization of the Quality of Administrative Records Used for Statistics*. Paper for the European Conference on Quality in Official Statistics, Athens, Greece. Department of Methodology and Development. The Hague: Statistics Netherlands. Available: http://www.pietdaas.nl/beta/pubs/pubs/Q2012_Session23_paper.pdf.

de Leeuw, E. and W. de Heer (2002). Trends in household survey nonresponse: A longitudinal and international comparison. Pp. 41–54 in *Survey Nonresponse*. R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J.A. Little, eds. New York: John Wiley & Sons. Available: https://www.researchgate.net/publication/44161946_Trends_in_Household_Survey_Nonresponse_A_Longitudinal_and_International_Comparison.

Doyle, P., J. Lane, J. Theeuwes, and L. Zayatz (2001). *Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies*. North Holland, Netherlands: Elsevier Science.

Duncan, J. W. and W. C. Shelton (1978). *Revolution in United States Statistics, 1926–1976*. Office of Federal Statistical Policy and Standards, U.S. Department of Commerce. Washington, DC: U.S. Government Printing Office. Available: https://catalog.hathitrust.org/Record/007886111/Home.

Energy Information Administration (1996). *Residential Energy Consumption Survey Quality Profile*. Washington, DC: U.S. Department of Energy. Available: https://digital.library.unt.edu/ark:/67531/metadc666723/.

Eurostat (2000). *Assessment of the Quality in Statistics*. Working Group on Assessment of Quality in Statistics, Luxembourg, April 4–5, 2000. Doc. Eurostat/A4/Quality/00/General/Standard. Luxembourg: Eurostat.

Available: http://www.unece.org/fileadmin/DAM/stats/documents/2000/11/metis/crp.3.e.pdf.

Federal Committee on Statistical Methodology (1978). *An Error Profile: Employment as Measured by the Current Population Survey*. Statistical Policy Working Paper 3 (NTIS PB86-214269/AS). Subcommittee on Nonsampling Errors. Washington, DC: U.S. Department of Commerce. Available: https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/04/spwp3.pdf.

Federal Committee on Statistical Methodology (2001). *Measuring and Reporting Sources of Error in Surveys*. Statistical Policy Working Paper 31 (NTIS PB2001104329). Subcommittee on Measuring and Reporting the Quality of Survey Data. Washington, DC: U.S. Office of Management and Budget. Available: https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/04/spwp31.pdf.

Federal Committee on Statistical Methodology (2005). *Report on Statistical Disclosure Limitation Methodology*. Statistical Policy Working Paper 22. Originally prepared in 1994 by Subcommittee on Disclosure Limitation Methodology. Revised by Confidentiality and Data Access Committee. Washington, DC: U.S. Office of Management and Budget. Available: https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/04/spwp22.pdf.

Federal Committee on Statistical Methodology (2009). *Profiles in Success of Statistical Uses of Administrative Data*. Statistical Policy Working Paper 45. Subcommittee on Statistical Uses Of Administrative Records. Washington, DC: U.S. Office of Management and Budget. Available: https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/04/StatisticalUsesofARData.pdf.

Federal Committee on Statistical Methodology (2013). *Data Quality Assessment Tool for Administrative Data*. Statistical Policy Working Paper 46. Subcommittee on Statistical Uses of Administrative Records. Washington, DC: U.S. Office of Management and Budget. Available: https://s3.amazonaws.com/sitesusa/wp-content/uploads/sites/242/2014/04/DataQualityAssessmentTool.pdf.

Federal Interagency Forum on Aging-Related Statistics (2016). *Older Americans 2016: Key Indicators of Well-Being*. Washington, DC: U.S. Government Printing Office. Available: https://agingstats.gov/docs/LatestReport/Older-Americans-2016-Key-Indicators-of-WellBeing.pdf.

Federal Interagency Forum on Child and Family Statistics (2016). *America's Children in Brief: Key National Indicators of Well-Being, 2016*. Washington, DC: U.S. Government Printing Office. Available: https://www.childstats.gov/pdf/ac2016/ac_16.pdf.

Gates, G. W. (2012). Confidentiality. Pp. 93–95 in *Encyclopedia of the U.S. Census: From the Constitution to the American Community Survey (ACS)* (2nd edition). M.J. Anderson, C.F. Citro, and J.J. Salvo, eds. Washington, DC: CQ Press.

Gillingham, R. and W. Lane (1982). Changing the treatment of shelter costs for homeowners in the CPI. *Monthly Labor Review 105*(6), 9–14. Available: https://www.bls.gov/opub/mlr/1982/06/art2full.pdf.

Hendriks, C. (2012). Input data quality in register based statistics: The Norwegian experience. In *Proceedings of the 2012 Joint Statistical Meetings (San Diego, CA)*, Alexandria, VA, pp. 1472–1480. American Statistical Association.

Hillygus, D. S., N. H. Nie, K. Prewitt, and H. Pals (2006). *The Hard Count: The Political and Social Challenges of Census Mobilization*. New York: Russell Sage Foundation. Available: http://www.jstor.org/stable/10.7758/9781610442886.

Interagency Technical Working Group on Developing a Supplemental Poverty Measure (2010). Observations from the Interagency Technical Working Group on Developing a Supplemental Poverty Measure. Available: https://www.census.gov/hhes/povmeas/methodology/supplemental/research/SPM_TWGObservations.pdf.

International Statistical Institute (2010). *Declaration on Professional Ethics*. Voorburg, The Netherlands: International Statistical Institute. Available: https://www.isi-web.org/index.php/news-from-isi/296-declarationprofessionalethics-2010uk.

Kalton, G., M. Winglee, S. Krawchuk, and D. Levine (2000). *Quality Profile for SASS Rounds 1–3: 1987–1995, Aspects of the Quality of Data in the Schools and Staffing Surveys (SASS)*. Washington, DC: National Center for Education Statistics, U.S. Department of Education. Available: https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2000308.

Lohr, S. and T. Raghunathan (2017). Combining survey data with other data sources. *Statistical Science 32*(3), 293–312. Available: https://projecteuclid.org/euclid.ss/1494489817.

Manski, C. F. (2015). Communicating uncertainty in official economic statistics: An appraisal fifty years after Morgenstern. *Journal of Economic Literature 53*(3), 631–653. Available: https://www.aeaweb.org/articles?id=10.1257/jel.53.3.631.

Martin, M. E. (1981). Statistical practice in bureaucracies. *Journal of the American Statistical Association 76*(373), 1–8. Available: http://www.tandfonline.com/doi/abs/10.1080/01621459.1981.10477593.

Muth, M. K., M. Sweitzer, D. Brown, K. Capogrossi, S. Karns, D. Levin, A. Okrent, P. Siegel, and C. Zhen (2016). *Understanding IRI Household-Based and Store-Based Scanner Data*. Technical Bulletin No. TB-1942. Washington, DC: Economic Research Service, U.S. Department of Agriculture. Available: https://www.ers.usda.gov/publications/pub-details/?pubid=47636.

National Academies of Sciences, Engineering, and Medicine (2016). *Reducing Respondent Burden in the American Community Survey: Proceedings of a Workshop*. T.J. Plewes, rapporteur. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/23639.

National Academies of Sciences, Engineering, and Medicine (2017a). *Advancing Concepts and Models for Measuring Innovation: Proceedings of a Workshop*. C.D. Mackie, rapporteur. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/23640.

National Academies of Sciences, Engineering, and Medicine (2017b). *Innovations in Federal Statistics Using New Data Sources While Respecting Privacy*. Panel on Improving Federal Statistics for Policy and Social Science Research Using Multiple Data Sources and State-of-the-Art Estimation Methods, R.M. Groves and B.A. Harris-Kojetin, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/24652.

National Center for Education Statistics (2012). *Statistical Standards of the National Center for Education Statistics*. First issued in 1987, revised in 1992 and 2002, and updated in 2012. Washington, DC: U.S. Department of Education. Available: https://nces.ed.gov/statprog/2012/.

National Institute of Statistical Sciences (2017). *Integrity, Independence and Innovation: The Future of NCES*. Report of an Expert Panel. Washington, DC: National Institute of Statistical Sciences. Available: https://niss.org/news/niss-panel-report-released-integrity-independence-and-innovation-future-nces.

National Research Council (1976). *Setting Statistical Priorities*. Panel on Methodology for Statistical Priorities, Committee on National Statistics. Washington, DC: National Academy Press. Available from the Committee on National Statistics.

National Research Council (1979). *Privacy and Confidentiality as Factors in Survey Nonresponse*. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/19845.

National Research Council (1984). *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*. Report of the Advanced Research Seminar on Cognitive Aspects of Survey Methodology, T.B. Jabine, M.L. Straf, J.M. Tanur, and R. Tourangeau, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/930.

National Research Council (1985). *Sharing Research Data*. Committee on National Statistics, S.E. Fienberg, M.E. Martin, and M.L. Straf, eds. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/2033.

National Research Council (1986). *Creating a Center for Education Statistics: A Time for Action*. Panel to Evaluate the National Center for Education Statistics, D.B. Levine, ed. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/19230.

National Research Council (1991a). *Improving Information for Social Policy Decisions: The Uses of Microsimulation. Volume I: Review and Recommendations*. Panel to Evaluate Microsimulation Models for Social Welfare Programs, C.F. Citro and E.A. Hanushek, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/1835.

National Research Council (1991b). *Improving Information for Social Policy Decisions: The Uses of Microsimulation. Volume II: Technical Papers*. Panel to Evaluate Microsimulation Models for Social Welfare Programs, C.F. Citro and E.A. Hanushek, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/1853.

National Research Council (1993a). *The Future of the Survey of Income and Program Participation*. Panel to Evaluate the Survey of Income and Program Participation, C.F. Citro and G. Kalton, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/2072.

National Research Council (1993b). *Private Lives and Public Policies: Confidentiality and Accessibility of Government Statistics*. Panel on Confidentiality and Data Access, G.T. Duncan, T.B. Jabine, and V.A. de Wolf, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/2122.

National Research Council (1995). *Measuring Poverty—A New Approach*. Panel on Poverty and Family Assistance, C.F. Citro and R.T. Michael, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/4759.

National Research Council (1997a). *Assessing Policies for Retirement Income: Needs for Data, Research, and Models*. Panel on Retirement Income Modeling, C.F. Citro and E.A. Hanushek, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/5420.

National Research Council (1997b). *The Bureau of Transportation Statistics: Priorities for the Future*. Panel on Statistical Programs and Practices of the Bureau of Transportation Statistics, C.F. Citro and J.L. Norwood, eds. Committee on National Statistics and Transportation Research Board. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/5809.

National Research Council (2000a). *Improving Access to and Confidentiality of Research Data: Report of a Workshop*. C.M. Mackie and N. Bradburn, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/9958.

National Research Council (2000b). *Measuring the Science and Engineering Enterprise: Priorities for the Division of Science Resource Studies*. Committee to Assess the Portfolio of the Division of Science Resources Studies of NSF, Committee on National Statistics and Office of Scientific and Engineering Personnel. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/9775.

National Research Council (2000c). *Small-Area Estimates of School-Age Children in Poverty: Evaluation of Current Methodology*. Panel on Estimates of Poverty for Small Geographic Areas, C.F. Citro and G. Kalton, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/6427.

National Research Council (2000d). *Small-Area Income and Poverty Estimates: Priorities for 2000 and Beyond*. Panel on Estimates of Poverty for Small Geographic Areas, C.F. Citro and G. Kalton, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/9957.

National Research Council (2001a). *Cells and Surveys: Should Biological Measures Be Included in Social Science Research?* C.E. Finch, J.W. Vaupel, and K. Kinsella, eds. Committee on Population. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/9995.

National Research Council (2001b). *Evaluating Welfare Reform in an Era of Transition*. Panel on Data and Methods for Measuring the Effects of Changes in Social Welfare Programs, R.A. Moffitt and M. Ver Ploeg, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/10020.

National Research Council (2002). *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes*. Panel on Conceptual, Measurement, and Other Statistical Issues in Developing Cost-of-Living Indexes, C. Schultze and C.M. Mackie, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/10131.

National Research Council (2003a). *Estimating Eligibility and Participation for the WIC Program: Final Report*. Panel to Evaluate the USDA's Methodology for Estimating Eligibility and Participation for the WIC Program, M. Ver Ploeg and D.M. Betson, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/10804.

National Research Council (2003b). *Protecting Participants and Facilitating Social and Behavioral Sciences Research*. Panel on Institutional Review Boards, Surveys, and Social Science Research, C.F. Citro, D.R. Ilgen, and C.B. Marrett, eds. Committee on National Statistics and Board on Behavioral, Cognitive, and Sensory Sciences. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/10638.

National Research Council (2003c). *Survey Automation: Report and Workshop Proceedings*. Oversight Committee for the Workshop on Survey Automation, D.L. Cork, M.L. Cohen, R.M. Groves, and W. Kalsbeek, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/10695.

National Research Council (2004a). *The 2000 Census: Counting Under Adversity*. Panel to Review the 2000 Census, C.F. Citro, D.L. Cork, and J.L. Norwood, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/10907.

National Research Council (2004b). *Eliminating Health Disparities: Measurement and Data Needs*. Panel on DHHS Collection of Race and Ethnic Data, M. Ver Ploeg and E. Perrin, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/10979.

National Research Council (2004c). *Measuring Research and Development Expenditures in the U.S. Economy.* Panel on Research and Development Statistics at the National Science Foundation, L.D. Brown, T.J. Plewes, and M.A. Gerstein, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11111.

National Research Council (2004d). *Reengineering the 2010 Census: Risks and Challenges*. Panel on Research on Future Census Methods, D.L. Cork, M.L. Cohen, and B.F. King, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/10959.

National Research Council (2005a). *Beyond the Market: Designing Nonmarket Accounts for the United States.* Panel to Study the Design of Nonmarket Accounts, K.G. Abraham and C.M. Mackie, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11181.

National Research Council (2005b). *Expanding Access to Research Data: Reconciling Risks and Opportunities*. Panel on Data Access for Research Purposes, Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11434.

National Research Council (2005c). *Experimental Poverty Measures: Summary of a Workshop*. J. Iceland, rapporteur. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11166.

National Research Council (2006a). *Food Insecurity and Hunger in the United States: An Assessment of the Measure*. Panel to Review USDA's Measurement of Food Insecurity and Hunger, G.S. Wunderlich and J.L. Norwood, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11578.

National Research Council (2006b). *Improving Business Statistics Through Interagency Data Sharing: Summary of a Workshop*. C. Kuebler and C.M. Mackie, rapporteurs. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11738.

National Research Council (2006c). *Once, Only Once, and in the Right Place: Residence Rules in the Decennial Census*. Panel on Residence Rules in the Decennial Census, D.L. Cork and P.R. Voss, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11727.

National Research Council (2007a). *Understanding Business Dynamics: An Integrated Data System for America's Future*. Panel on Measuring Business Formation, Dynamics, and Performance, J. Haltiwanger, L.M. Lynch, and C.M. Mackie, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11844.

National Research Council (2007b). *Using the American Community Survey: Benefits and Challenges*. Panel on the Functionality and Usability of Data from the

American Community Survey, C.F. Citro and G. Kalton, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11901.

National Research Council (2008a). *Protecting Individual Privacy in the Struggle Against Terrorists—A Framework for Program Assessment*. Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and Other National Goals, Committee on Law and Justice, Committee on National Statistics, and Computer Science and Telecommunications Board. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12452.

National Research Council (2008b). *Using the American Community Survey for the National Science Foundation's Science and Engineering Workforce Statistics Program*. Panel on Assessing the Benefits of the American Community Survey for the NSF Division of Science Resources Statistics, Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12244.

National Research Council (2009a). *Ensuring the Quality, Credibility, and Relevance of U.S. Justice Statistics*. Panel to Review the Programs of the Bureau of Justice Statistics, R.M. Groves and D.L. Cork, eds. Committee on National Statistics and Committee on Law and Justice. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12671.

National Research Council (2009b). *Improving the Measurement of Late-Life Disability in Population Surveys: Beyond ADLs and IADLs, Summary of a Workshop*. G.S. Wunderlich, rapporteur. Committee on National Statistics and Committee on Population. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12740.

National Research Council (2009c). *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. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12564.

National Research Council (2009d). *Protecting Student Records and Facilitating Education Research: A Workshop Summary*. M. Hilton, rapporteur. Committee on National Statistics and Center for Education. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12514.

National Research Council (2009e). *Reengineering the Survey of Income and Program Participation*. Panel on the Census Bureau's Reengineered Survey of Income and Program Participation, C.F. Citro and J.K. Scholz, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12715.

National Research Council (2010a). *Accounting for Health and Health Care: Approaches to Measuring the Sources and Costs of Their Improvement*. Panel to Advance a Research Program on the Design of National Health Accounts, Committee on

National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12938.

National Research Council (2010b). *Conducting Biosocial Surveys: Collecting, Storing, Accessing, and Protecting Biospecimens and Biodata*. Panel on Collecting, Storing, Accessing, and Protecting Biological Specimens and Biodata in Social Surveys, R.M. Hauser, M. Weinstein, R. Pool, and B. Cohen, eds. Committee on National Statistics and Committee on Population. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12942.

National Research Council (2010c). *Envisioning the 2020 Census*. Panel on the Design of the 2010 Census Program of Evaluations and Experiments, L.D. Brown, M.L. Cohen, D.L. Cork, and C.F. Citro, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12865.

National Research Council (2010d). *Improving Health Care Cost Projections for the Medical Population: Summary of a Workshop*. G. Wunderlich, rapporteur. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12985.

National Research Council (2010e). *The Prevention and Treatment of Missing Data in Clinical Trials*. Panel on Handling Missing Data in Clinical Trials, Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12955.

National Research Council (2010f). *Protecting and Accessing Data from the Survey of Earned Doctorates: A Workshop Summary*. T.J. Plewes, rapporteur. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12797.

National Research Council (2012). *Communicating Science and Engineering Data in the Information Age*. Panel on Communicating National Science Foundation Science and Engineering Information to Data Users, Committee on National Statistics and Computer Science and Telecommunications Board. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/13282.

National Research Council (2013a). *Nonresponse in Social Science Surveys: A Research Agenda*. Panel on a Research Agenda for the Future of Social Science Data Collection, R. Tourangeau and T.J. Plewes, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/18293.

National Research Council (2013b). *Subjective Well Being: Measuring Happiness, Suffering, and Other Dimensions of Experience*. Panel on Measuring Subjective Well-Being in a Policy-Relevant Framework, A.A. Stone and C.M. Mackie, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/18548.

National Research Council (2014a). *Capturing Change in Science, Technology, and Innovation: Improving Indicators to Inform Policy*. Panel on Developing Science, Technology, and Innovation Indicators for the Future, R.E. Litan, A.W. Wyckoff, and K.H. Fealing, eds. Committee on National Statistics and Board on Science, Technology, and Economic Policy. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/18606.

National Research Council (2014b). *Civic Engagement and Social Cohesion: Measuring Dimensions of Social Capital to Inform Policy*. Panel on Measuring Social and Civic Engagement and Social Cohesion in Surveys, K. Prewitt, C.D. Mackie, and H. Habermann, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/18831.

National Research Council (2014c). *Proposed Revisions to the Common Rule for the Protection of Human Subjects in the Behavioral and Social Sciences*. Committee on Revisions to the Common Rule for the Protection of Human Subjects in Research in the Behavioral and Social Sciences, Board on Behavioral, Cognitive, and Sensory Sciences, Committee on National Statistics, and Committee on Population. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/18614.

National Research Council and Institute of Medicine (2012). *Medical Care Economic Risk: Measuring Financial Vulnerability from Spending on Medical Care*. Panel on Measuring Medical Care Risk in Conjunction with the New Supplemental Poverty Income Measure, M.J. O'Grady and G.S. Wunderlich, eds. Committee on National Statistics and Board on Health Care Services. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/13525.

Norwood, J. L. (1975). Should those who produce statistics analyze them? How far should analysis go? An American view. *Bulletin of the International Statistical Institute [Proceedings of the 40th Session] 46*, 420–432.

Norwood, J. L. (1995). *Organizing to Count: Change in the Federal Statistical System*. Washington, DC: The Urban Institute Press.

Statistics Canada (2009). *Statistics Canada Quality Guidelines* (5th ed.). Ottawa: Statistics Canada. Available: https://unstats.un.org/unsd/dnss/docs-nqaf/Canada-12-539-x2009001-eng.pdf.

Triplett, J. (1991). The federal statistical system's response to emerging data needs. *Journal of Economic and Social Measurement 17*(3–4), 155–201. Available: http://content.iosPress.com/articles/journal-of-economic-and-social-measurement/jem17-3-4-03.

UNECE High-Level Group for the Modernisation of Official Statistics (2013, December). Generic statistical business process model (GSBPM), version 5.0. Maintained by Modernisation Committee on Standards. Available: http://www1.unece.org/stat/platform/display/metis/The+Generic+Statistical+Business+Process+Model.

United Kingdom Office for National Statistics (2007). *Guidelines for Measuring Statistical Quality*. London: Office for National Statistics. Available: https://unstats.un.org/unsd/dnss/docs-nqaf/UK-Guidelines_Subject.pdf.

U.S. Census Bureau (1998). *Survey of Income and Program Participation Quality Profile 1998* (3rd ed.). SIPP Working Paper No. 230. Washington, DC: U.S. Department of Commerce. Available: https://www.census.gov/library/working-papers/1998/demo/SEHSD-WP1998-11.html.

U.S. Office of Management and Budget (1985). Statistical Policy Directive No. 3: Compilation, release, and evaluation of principal federal economic indicators. 50 *Federal Register* 38932 (September 25, 1985). Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/omb/inforeg/statpolicy/dir_3_fr_09251985.pdf.

U.S. Office of Management and Budget (2008). Statistical Policy Directive No. 4: Release and dissemination of statistical products produced by federal statistical agencies. 73 *Federal Register* 12625 (March 7, 2008). Available: https://www.federalregister.gov/d/E8-4570.

U.S. Office of Management and Budget (2014a). Guidance for providing and using administrative data for statistical purposes. OMB Memorandum M-14-06. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2014/m-14-06.pdf.

U.S. Office of Management and Budget (2014b). Statistical Policy Directive No. 1: Fundamental responsibilities of federal statistical agencies and recognized statistical units. 79 *Federal Register* 71609 (December 2, 2014). Available: https://www.federalregister.gov/d/2014-28326.

U.S. Office of Management and Budget (2016a). Statistical Policy Directive No. 2: Standards and guidelines for statistical surveys; Addendum: Standards and guidelines for cognitive interviews. 81 *Federal Register* 70586 (October 12, 2006). Available: https://www.federalregister.gov/d/2016-24607.

U.S. Office of Management and Budget (2016b, May 4). Supplemental guidance on the implementation of M-15-14 "Management and oversight of federal information technology"—Applying FITARA common baseline to statistical agencies and units. Memorandum from Tony Scott, Federal Chief Information Officer, and Howard Shelanski, Administrator, Office of Regulatory Affairs. Available: https://management.cio.gov/assets/docs/FITARA_Guidance_Statistical_Agencies_and_Units_OMB.pdf.

U.S. Office of Management and Budget (2017). *Statistical Programs of the United States Government: Fiscal Year 2017*. Statistical and Science Policy Office, Office of Information and Regulatory Affairs. Washington, DC: U.S. Government Printing Office. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/information_and_regulatory_affairs/statistical-programs-2017.pdf.

Appendixes

– A –

Legislation and Regulations That Govern Federal Statistics

This appendix summarizes the major legislation and agency regulations and guidance that govern the operations of the federal statistical system as a whole, which is under the regulatory authority of the U.S. Office of Management and Budget (OMB). The descriptions are organized into three main categories: (1) the Paperwork Reduction Act and related OMB statistical policy documents and other guidance; (2) confidentiality and privacy protection; and (3) information quality, peer review, performance evaluation, scientific integrity, and transparency. See Appendix B for descriptions of the overall structure of the system and the principal statistical agencies. Note that each agency is governed by additional legislation and guidance particular to its department and itself.

Most of the legislation, regulations, and guidance summarized herein pertains to OMB, which plays a critical role in oversight of the federal government's widely dispersed statistical operations. The oversight role dates to 1939, when the functions of a Central Statistical Board, created in 1933, were transferred to the then-named Bureau of the Budget (see Anderson, 2015; Duncan and Shelton, 1978; Norwood, 1995). Recent legislation and guidance address such system-wide issues as confidentiality protection and privacy of respondents, data quality (including peer review prior to dissemination), efficiency of operations, and scientific integrity and transparency.

AUTHORITY OF THE U.S. OFFICE OF MANAGEMENT AND BUDGET (OMB) OVER FEDERAL STATISTICS

The 1980 Paperwork Reduction Act, as Reauthorized and Amended in 1995, and Associated 2006 Implementation Guidance

The Paperwork Reduction Act (PRA) of 1980 (P.L. 96-511 and codified at 44 USC 3501 and following; reauthorized and amended in 1986 by P.L. 99-500 and 1995 by P.L. 104-13) is the legal foundation for the modern statistical coordination and management mission of OMB. It establishes OMB's review power over federal statistical agencies and myriad other agencies throughout the federal government that collect information from individuals and organizations. This review power covers both burden imposed by information collection and methods and practices for data collection and dissemination.

Background, 1933–1980

The PRA's origins trace back to Executive Order 6226, signed by Franklin D. Roosevelt in July 1933, which established a Central Statistical Board to "appraise and advise upon all schedules of all Government agencies engaged in the primary collection of statistics required in carrying out the purposes of the National Industrial Recovery Act, to review plans for tabulation and classification of such statistics, and to promote the coordination and improvement of the statistical services involved." Members of the board were appointed by the relevant cabinet secretaries. The board was established in law for a 5-year period in 1935. Its functions were transferred to the Bureau of the Budget (itself established in 1921) in 1939, when the Budget Bureau was transferred to the Executive Office of the President.

The 1942 Federal Reports Act represented another milestone by codifying the authority for the Budget Bureau to coordinate and oversee the work of federal statistical agencies. Most famously, it provided that no federal agency could collect data from 10 or more respondents without approval of the budget director. (Data collections by contractors on behalf of federal agencies are covered by this provision, although data collections by government grantees are generally not covered.) The 1950 Budget and Accounting Procedures Act (31 USC 1104(d)) further strengthened the statistical coordinating and improvement role of OMB, giving OMB authorization to promulgate regulations and orders governing statistical programs throughout the federal government.

The statistical policy function continued in the budget office in the Executive Office of the President when the Budget Bureau became the Office of Management and Budget in 1970. However, in 1977, the statistical policy staff was split into two groups: one group remained in OMB to handle

the paperwork clearance and review function for statistical agencies; the other group was moved to the U.S. Department of Commerce to address statistical policy and standards issues (Executive Order 12013, October 7, 1977).[1]

Paperwork Reduction, 1980–Present

The overarching goal of the 1980 PRA was to reduce the burden of filling out federal forms by businesses and individuals. It created a new Office of Information and Regulatory Affairs (OIRA) within OMB, which was charged to reduce the combined burden imposed by regulatory agencies and administrative and statistical program agencies. The PRA required OIRA to engage in long-range planning to improve federal statistical programs; review statistical budgets; coordinate government statistical functions; establish standards, classifications, and other guidelines for statistical data collection and dissemination; and evaluate statistical program performance. In furtherance of that work, Executive Order 12318 (August 21, 1981) revoked the 1977 order and moved the statistical policy office from the Department of Commerce and under OIRA; the 1986 reauthorization of the PRA required the appointment of a chief statistician at OMB to carry out the statistical policy functions (100 Stat. 1783-337).[2] In the 1995 reauthorization and extensive revision of the PRA, two of the most significant provisions added for statistical policy were: (1) codifying the Interagency Council on Statistical Policy (ICSP), chaired by the chief statistician; and (2) updating the legislation to reflect the chief statistician's role in coordinating U.S. participation in international statistical activities.

Survey Clearance Process under the PRA

In January 2006 the OMB Statistical and Science Policy Office released *Guidance on Agency Survey and Statistical Information Collections—Questions and Answers When Designing Surveys for Information Collections*.[3] The guidance is a set of 81 questions and answers that attempts to demystify the OMB clearance process (required by the PRA) for surveys and other statistical information collections. Its purpose is to explain OMB's review process, assist agencies in strengthening their supporting statements

[1]Seven months later, a pair of *Federal Register* notices (43 *Federal Register* 19260 and 19308) formally transferred the content of and responsibility for various regulatory circulars on federal statistical activities to the Commerce Department—at which time they were designated "statistical policy directives" for the first time.

[2]Consequent to congressional hearings, the Reagan administration first appointed a chief statistician in 1983.

[3]Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/omb/inforeg/pmc_survey_guidance_2006.pdf [April 2017].

for information collection requests, and provide advice for improving information collection designs.

The *Guidance* covers such topics as its purpose; submission of information collection requests (ICRs, often called clearance packages) to OMB; scope of the information collection (e.g., calculation of burden hours on respondents); choice of methods; sampling; modes of data collection; questionnaire design and development; statistical standards; informing respondents about their participation and the confidentiality of their data; response rates and incentives; analysis and reporting; and studies using stated preference methods (which ask respondents about the use or non-use value of a good in order to obtain willingness-to-pay estimates relevant to benefit or cost estimation). The *Guidance* includes a glossary of terms and ICR supporting statement instructions.

The document outlines the statutory timing and process requirements for all statistical information collection requests in order to obtain OMB approval (which is indicated by an OMB control number on an approved survey questionnaire). After an agency has developed a draft information collection plan and instrument, the agency must publish a 60-day notice for public comment in the *Federal Register* and have the draft survey instrument available for the public to review. Following the 60 days, the agency may submit its clearance package to OMB. With that submission, the agency must place a second notice in the *Federal Register*, allowing a 30-day public comment period and notifying the public that OMB approval is being sought and that comments may be submitted to OMB. This notice runs concurrent with the first 30 days of OMB review—half of OMB's total of 60 days after receipt of the clearance package to make its decision to approve or disapprove or to instruct the agency to make a substantive change to its proposed collection. Generally, agencies need to allow 6 months to complete the entire process, including development of a clearance package, public comment, and agency, departmental, and OMB review.

In recent years, OMB has issued several memoranda to clarify particular interpretations and applications of the PRA to agency activities.[4] Topics covered include an overview of PRA requirements, PRA implications of social media and web-based interactive technologies, the use of generic clearances, options for streamlining the PRA process for scientific research, a fast-track process for qualitative customer service delivery feedback, and answers to PRA questions related to challenges and prizes.

[4] All the memoranda can currently be found at https://obamawhitehouse.archives.gov/omb/inforeg_infocoll [April 2017].

OMB Statistical Policy Documents

Historically, OMB has issued guidance and standards on a variety of topics via "circulars," which are expected to have a continuing effect of 2 or more years.[5] Early statistical policy guidance was issued in this form. After the statistical policy function was moved to the U.S. Department of Commerce in 1977, numerous circulars were revised and reissued as "statistical policy directives" (43 *Federal Register* 19260) so as not to cause confusion with other OMB circulars. Since the statistical policy function was moved back to OMB in 1981, terminology has varied, with some policy documents called "directives" and others "standards" or "classification." The process of issuing or revising a directive or standard involves expert review, agency consultation, and public comment.

For convenience, five statistical policy directives issued by the OMB Statistical and Science Policy Office that are currently in effect are described first—these are directives nos. 1, 2, 3, 4, and 14. Four statistical standards and classifications—for industries, occupations, metropolitan and micropolitan statistical areas, and race and ethnicity—are described next, along with a standard for products that is under construction. The industry and product classifications are developed jointly with Canada and Mexico.

Statistical Policy Directive No. 1—Fundamental Responsibilities of Federal Statistical Agencies and Recognized Statistical Units

OMB issued its latest statistical policy directive on December 2, 2014 (U.S. Office of Management and Budget, 2014b). The directive was labeled no. 1 because of its foundational importance. (The original no. 1 was combined with no. 2 as described below.) The directive cites relevant documents issued by OMB (e.g., other statistical policy directives) and by the Office of Science and Technology Policy, together with *Principles and Practices for a Federal Statistical Agency* (National Research Council, 2013), the European Statistics Code of Practice (European Statistical System Committee, 2011), and the *Fundamental Principles of Official Statistics* (United Nations Statistical Commission, 2014), as contributing "to an integrative framework guiding the production of Federal statistics, encompassing design, collection, processing, editing, compilation, storage, analysis, release, and dissemination" (U.S. Office of Management and Budget, 2014b:71611). The directive notes, however (U.S. Office of Management and Budget, 2014b:71612):

> Although these principles and policies provide a common foundation for core statistical agency functions, their actual implementation in

[5]See https://www.whitehouse.gov/omb/information-for-agencies/circulars [April 2017]. Guidance that is of a transitory nature or requires a one-time action by federal departments is issued in the form of "bulletins."

> the form of standards and practices can involve a wide range of managerial and technical challenges. Therefore, to support agency decision-making in a manner that fosters statistical quality, OMB developed this Statistical Policy Directive. This Directive provides a unified articulation of Federal statistical agency responsibilities.

The directive goes on to articulate four statistical agency responsibilities: (1) produce and disseminate relevant and timely information; (2) conduct credible and accurate statistical activities; (3) conduct objective statistical activities; and (4) protect the trust of information providers by ensuring the confidentiality and exclusive statistical use of their responses. The directive states (U.S. Office of Management and Budget, 2014b:71614):

> Federal statistical agencies and recognized statistical units must adhere to these responsibilities and adopt policies, best practices, and appropriate procedures to implement them. *Federal departments must enable, support, and facilitate Federal statistical agencies and recognized statistical units as they implement these responsibilities.* [emphasis added]

Statistical Policy Directive No. 2—Standards and Guidelines for Statistical Surveys

OMB issued *Standards and Guidelines for Statistical Surveys* in September 2006 as a combined update and revision of *Statistical Policy Directive No. 1, Standards for Statistical Surveys*, and *Statistical Policy Directive No. 2, Publication of Statistics*. The new document (U.S. Office of Management and Budget, 2006), which is now Directive No. 2, includes 20 standards and one or more associated guidelines for every aspect of survey methodology from planning through data release:

1. survey planning,
2. survey design,
3. survey response rates,
4. pretesting survey systems,
5. developing sampling frames,
6. required notification to potential survey respondents,
7. data collection methodology,
8. data editing,
9. nonresponse analysis and response rate calculation,
10. coding,
11. data protection,
12. evaluation,
13. developing estimates and projections,
14. analysis and report planning,
15. inference and comparisons,

16. review of information products,
17. releasing information,
18. data protection and disclosure avoidance for dissemination,
19. survey documentation, and
20. documentation and release of public-use microdata.

On October 12, 2016, OMB issued a notice in the *Federal Register* of a final decision to add an *Addendum: Standards and Guidelines for Cognitive Interviews* to Directive No. 2 (U.S. Office of Management and Budget, 2016a). This addendum recognizes the important role that qualitative cognitive interviewing techniques play in the design of effective survey questions (see National Research Council, 1984, 2006:Ch. 8).

Statistical Policy Directive No. 3—Compilation, Release, and Evaluation of Principal Federal Economic Indicators

OMB first issued Directive No. 3 in the 1970s and strengthened it in 1985.[6] Its purpose is clearly stated (U.S. Office of Management and Budget, 1985:38932):

> [This directive] designates statistical series that provide timely measures of economic activity as Principal Economic Indicators and requires prompt release of these indicators by statistical agencies in a politically-neutral manner. The intent of the directive is to preserve the time value of such information, strike a balance between timeliness and accuracy, prevent early access to information that may affect financial and commodity markets, and preserve the distinction between the policy-neutral release of data by statistical agencies and their interpretation by policy officials.

Each September OMB issues the *Schedule of Release Dates for Principal Federal Economic Indicators* for the subsequent calendar year.[7] At present, the following agencies issue one or more of the 38 principal economic indicators:

- Bureau of Economic Analysis (5 indicators, including gross domestic product [GDP], personal income and outlays, corporate profits)
- Bureau of Labor Statistics (7 indicators, including the employment situation, Consumer Price Index [CPI])
- Census Bureau (13 indicators, including new residential construction, monthly retail trade)
- Energy Information Administration (natural gas storage)

[6] Norwood (2016) recounts the history of threats to the integrity of economic indicators that necessitated the directive's issuance and updating.

[7] Available: https://obamawhitehouse.archives.gov/omb/inforeg_statpolicy [April 2017].

- Federal Reserve Board (4 indicators, including money stock measures, consumer installment credit)
- Foreign Agricultural Service (world agricultural production)
- National Agricultural Statistics Service (6 indicators, including agricultural prices, grain production)
- World Agricultural Outlook Board (world agricultural supply and demand)

Statistical Policy Directive No. 4—Release and Dissemination of Statistical Products Produced by Federal Statistical Agencies

OMB issued Directive No. 4, which essentially covers all statistical releases other than those specified in Directive No. 3, in 2008 (U.S. Office of Management and Budget, 2008). It includes not only statistical information released in printed reports or on the Internet, but also statistical press releases, which describe or announce a statistical data product. Statistical press releases are the sole responsibility of the relevant statistical agency. Each fall statistical agencies must issue a schedule of when they expect each regular or recurring product to be released and give timely notification of any change to the published schedule.

On October 17, 2016, OMB issued a notice in the *Federal Register* requesting comments on a proposed addendum to Directive No. 4, which would constitute *Section 10: Performance Review* (U.S. Office of Management and Budget, 2016b). Comments were due December 1, 2016. As proposed, the addendum, which incorporates language from Directive No. 3, would require each statistical agency and recognized statistical unit to submit an annual performance review of the production and dissemination of its key statistical products to OMB. Key products would be defined by the agency in consultation with OMB. Reviews would address for each product (U.S. Office of Management and Budget, 2016b:71541–71542):

> (a) The accuracy and reliability of the series, e.g., the magnitude and direction of all revisions, the performance of the series relative to established benchmarks, and the proportion and effect of nonresponses or responses received after the publication of preliminary estimates; (b) the accuracy, completeness, and accessibility of documentation describing the methods used in compiling and revising the product; (c) the agency's performance in meeting its established release schedule and the prompt release objective of this Directive; (d) the agency's ability to avoid disclosure prior to the scheduled release time; (e) any additional issues (such as periodicity, electronic access, etc.) that the Administrator for Information and Regulatory Affairs specifies in writing to the agency at least 6 months in advance of the scheduled submission date.

OMB would include a summary of the year's evaluations in its annual report to Congress. If enacted, this addendum would represent the first formal process by OMB for performance review specifically of statistical products, with the exception that Directive No. 3 requires that the key economic indicators (see above) be evaluated every 3 years.[8]

Statistical Policy Directive No. 14—Definition of Poverty for Statistical Purposes

OMB first issued standards for the definition of poverty in 1969. It adopted the existing poverty thresholds (first specified by Mollie Orshansky of the Social Security Administration in 1963 and used by the Office of Economic Opportunity) for different categories of families defined by size, number of children, gender of the family head, and farm-nonfarm residences. (One change from Orshansky's specification was that the farm thresholds were raised from 70% to 85% of the nonfarm thresholds.) For most family types, the thresholds represented the costs of a minimally adequate diet multiplied by three to allow for all other expenses.

The 1969 directive specified that the thresholds would be updated each year for the change in the Consumer Price Index (instead of the cost of the Economy Food Plan as in prior years) and compared with families' total regular money income as measured in the Current Population Survey. The directive was promulgated as *Statistical Policy Directive No. 14* in 1978, when the statistical policy function was briefly housed in the Department of Commerce (U.S. Department of Commerce, 1978:19269); minor modifications were made to the thresholds beginning in 1982 (the nonfarm thresholds were used for all families, the thresholds for male- and female-headed families were averaged, and the largest family size category was raised from 7 to 9 people).[9] No further changes have been made to the official thresholds or definition of countable resources, although major socioeconomic changes in the United States and in income support policies have made the official poverty concept increasingly unable to inform assessments of policy effectiveness for different population groups (e.g., refunds from the Earned Income Tax Credit are not counted in the resource measure).

With input from the observations of an Interagency Technical Working Group on Developing a Supplemental Poverty Measure (2010), the U.S.

[8]Statistical agencies were subject to earlier performance review programs (see, e.g., National Research Council, 2009:101), but these were not devised specifically for statistical products. Agencies continue to be subject to the legislation and guidance summarized in the section below on "Information Quality, Peer Review, Performance Evaluation, Scientific Integrity, and Transparency," which apply to the entire executive branch.

[9]Available: http://www2.census.gov/prod2/popscan/p60-133.pdf (page 9) [April 2017].

Census Bureau released a new supplemental poverty measure (SPM) in fall 2011 (referencing poverty in calendar 2010), using thresholds developed by the U.S. Bureau of Labor Statistics.[10] The thresholds and definition of countable resources as money and near-money disposable income for the SPM were derived in large part from the recommendations from the report of a Committee on National Statistics panel, *Measuring Poverty: A New Approach* (National Research Council, 1995). The SPM, which is designed to be a useful tool for policy evaluation, is issued annually as is the official measure.

North American Industry Classification System (NAICS)

The NAICS was developed by the United States, Canada, and Mexico to provide a common, contemporary classification system for economic production activity following the enactment of the North American Free Trade Agreement (NAFTA). NAICS, which is a substantial revision of its predecessor, the Standard Industrial Classification (SIC), was first issued in 1997. (The SIC was originally issued in various sections in 1938–1940 and revised on an irregular basis between 1940 and its last iteration in 1987.) Interagency and country working groups (under the aegis of OMB in the United States) have the opportunity to update NAICS every 5 years for years ending in 2 and 7 so that it keeps up reasonably well with changes in the structure of industrial activity in the three countries. NAICS was most recently updated for use beginning in 2017.[11]

North American Product Classification System (NAPCS)

NAPCS is intended to be a comprehensive, market- or demand-based, hierarchical classification system for products or outputs (goods and services) that: (a) is not industry-of-origin based but can be linked to the NAICS industry structure; (b) is consistent across the three North American countries; and (c) promotes improvements in the identification and classification of service products across international classification systems, such as the Central Product Classification System of the United Nations. NAPCS responds to the problem that a business establishment can only have one NAICS code, even though it may produce more than one

[10]The latest supplemental poverty measure statistics for 2015 were published in September 2016. Available: https://www.census.gov/library/publications/2016/demo/p60-258.html [April 2017].

[11]See "North American Industry Classification System—Revision for 2017," 81 *Federal Register* 52584 (August 8, 2016). Available: https://www.federalregister.gov/d/2016-18774 [April 2017]. See also "North American Industry Classification System (NAICS)—Updates for 2017," 80 *Federal Register* 46480 (August 4, 2015). Available: https://www.federalregister.gov/d/2015-19022 [April 2017].

product. It is also the case that the same product can be produced by more than one industry.

NAPCS has been under development since 1998, beginning with exploratory efforts to develop classifications for the services sector. At present a NAPCS for 2017 is in beta testing by U.S. statistical agencies; the plan is to update NAPCS every 5 years on the same cycle as NAICS. For more information, see Economic Classification Policy Committee of the United States (2003), in which the three countries agreed that the objectives and principles articulated in sections A through C of that paper define the purposes of NAPCS and the operational guidelines for creating it.

Standard Occupational Classification (SOC)

The SOC is used by federal statistical agencies to classify workers into occupational categories for collecting, tabulating, and disseminating data.[12] The first SOC was published in 1977 in an effort to standardize the collection of occupational data by multiple agencies. It was revised in 1980 but not universally adopted until an interagency process under the aegis of OMB further revised it in 1998 for use in the 2000 decennial census and surveys conducted in the following decade. Work to revise the 2000 SOC was completed in time for its use in 2010 for the American Community Survey (ACS), which provides occupational data in place of the decennial census "long form" sample, and other surveys. The next planned revision of the SOC is scheduled for 2018 and every 10 years thereafter.

Metropolitan Statistical Areas (MSAs)

For more than 60 years, the OMB Metropolitan Area Classification Program has provided standards for delineating areas that are "metropolitan" in nature for use throughout the federal government. In general, such an area has a population nucleus plus one or more adjacent communities that have a high degree of interaction with the nucleus. The usefulness of standardizing these classifications became clear in the 1940s, and the Bureau of the Budget (the predecessor to OMB) led an effort to develop what were then called "standard metropolitan areas" for use in 1950 census publications. For censuses from 1960 through 2000, OMB revised as appropriate the definitional criteria for metropolitan areas before each census and, on the basis of those citeria, issued an updated list of recognized areas after each census.[13]

[12] Available: http://www.bls.gov/soc [April 2017].

[13] Issues of rural area classification were discussed at a Committee on National Statistics' workshop sponsored by the Economic Research Service, U.S. Department of Agriculture, in April 2015 (National Academies of Sciences, Engineering, and Medicine, 2016).

The definitional criteria issued before the 2000 census marked a major revision to the coverage of the program. *Standards for Defining Metropolitan and Micropolitan Statistical Areas* defined not only metropolitan statistical areas, but also, for the first time, micropolitan areas.[14] Metropolitan areas are those with a central urbanized core of 50,000 or more people in one or more counties; micropolitan areas are those with a central urbanized core of 10,000 or more people in one or more counties. The list of metropolitan and micropolitan areas using the 2000 criteria was initially issued in 2003 and updated annually through 2008 by OMB on the basis of the Census Bureau's population estimates.[15] Two years later, OMB issued *2010 Standards for Delineating Metropolitan and Micropolitan Statistical Areas*, which largely continued the criteria adopted for the 2000 standards.[16] Areas based on these standards, using data from the 2010 census and the American Community Survey, were announced in 2013 and updated in 2015.[17]

Beginning with the 2010 census, the revision and updating process was changed to reflect the availability of needed commuting and employment information from the continuous ACS. Under the changed process, OMB will issue as often as annually a list of newly recognized areas by using Census Bureau population estimates; in addition, on the basis of ACS and census data, OMB may revise the criteria for area delineation every 5 years instead of every 10 years. The first such revision is planned to occur in 2018, using population estimates and ACS data on commuting and employment for 2011–2015. Input to the OMB decisions is provided by an interagency Metropolitan Area Standards Review Committee.

Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity

The first standards on this topic (originally labeled as *Statistical Policy Directive No. 15*) were issued in 1977. *Race and Ethnic Standards for Federal Statistics and Administrative Reporting* specified a minimum set of racial and ethnic categories for reporting of race and ethnicity on federal surveys and in administrative records systems. It recommended either two separate questions, one on ethnicity (Hispanic or non-Hispanic) and one on race (white, black, Asian or Pacific Islander, American Indian or Alaska Native),

[14]See 65 *Federal Register* 82228 (December 27, 2000). Available: https://www.federalregister.gov/d/00-32997 [April 2017].

[15]See, for example, *Update of Statistical Area Definitions and Guidance on Their Uses*. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/omb/bulletins/fy2009/09-01.pdf [April 2017].

[16]See 75 *Federal Register* 37245. Available: https://www.federalregister.gov/d/2010-15605 [April 2017].

[17]See https://obamawhitehouse.archives.gov/sites/default/files/omb/bulletins/2015/15-01.pdf [April 2017].

or a combined question that included Hispanic as a category. The U.S. decennial census has historically included additional categories under the two-question format.

Following an intensive research, testing, and consultation process, OMB issued a revised set of standards in 1997.[18] The updated standards retain a two-question format, includes separate categories for Asians and for Native Hawaiian and other Pacific Islanders, emphasizes self-identification, and allows respondents to select more than one racial category.

The 2010 census included several experimental panels to test different strategies that incorporated alternative wording and format for the questions on race and ethnicity, including a combined race and ethnicity question. Analysis of the results led to an important finding that the combined question improved reporting.[19] Additional research was conducted in subsequent years, including a National Content Test (NCT) in 2015 for 2020 census planning.[20]

On September 30, 2016, OMB issued a request for comments on a "possible limited revision" of *Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity*.[21] Comments were requested within 30 days on the possibility of allowing the use of a combined race and ethnicity question, adding a "Middle Eastern or North African (MENA)" category, and some other changes in terminology. On March 1, 2017, OMB asked for comments within 60 days on the interim proposals of the Federal Interagency Working Group for Research on Race and Ethnicity, which took account of the comments received on the September 30 notice.[22]

---

[18] Available: https://obamawhitehouse.archives.gov/omb/fedreg_1997standards/ [April 2017].

[19] Available: http://www.census.gov/2010census/pdf/2010_Census_Race_HO_AQE.pdf [April 2017].

[20] This site describes the design of the NCT: http://www.census.gov/content/dam/Census/programs-surveys/decennial/2020-census/2015_census_tests/nct/2015-nct-omb-package.pdf [April 2017]; limited results are included in the March 3, 2017, interim report of the Federal Interagency Working Group for Research on Race and Ethnicity: https://www.whitehouse.gov/the-press-office/2017/03/03/federal-register-notice-interim-report-omb-march-03-2017 [March 2017].

[21] See 81 *Federal Register* 67398. Available: https://www.federalregister.gov/d/2016-23672 [April 2017].

[22] See 82 *Federal Register* 12242. Available: https://www.federalregister.gov/d/2017-03973 [March 2017].

Other OMB Guidance for Federal Statistics

2014 Guidance for Providing and Using Administrative Data for Statistical Purposes (M-14-06)

Recognizing the informational value and potential efficiencies to be achieved by using already collected federal administrative data for federal statistics, the OMB director issued M-14-06 on February 14, 2014 (U.S. Office of Management and Budget, 2014a). The intent of this memorandum is to provide agencies with "guidance for addressing the legal, policy, and operational issues that exist with respect to using administrative data for statistical purposes." First, it "calls for departmental and agency leadership to: (i) foster greater collaboration between program and statistical offices; (ii) develop strong data stewardship policies and practices around the statistical use of administrative data; (iii) require the documentation of quality control measures and key attributes of important administrative datasets; and (iv) require the designation of responsibilities and practices through the use of agreements amongst these offices." Second, it "encourages Federal departments and agencies to promote the use of administrative data for statistical purposes . . . [while] continu[ing] to fully protect the privacy and confidentiality afforded to the individuals, businesses, and institutions providing the data." Third, it "provides some 'best practice' tools, including detailed guidance on the interaction of the Privacy Act requirements and the use of administrative data for statistical purposes, as well as a model interagency agreement for . . . sharing data for statistical purposes." Fourth, it "requires each department/agency to report to OMB, within 120 days . . . on its progress in implementing this Memorandum."

2015 Improving Statistical Activities through Interagency Collaboration (M-15-15)

Issued by the administrator of the OMB Office of Information and Regulatory Affairs (OIRA) on July 8, 2015 (U.S. Office of Management and Budget, 2015), this memorandum "strongly encourages the Federal statistical agencies and units, and their parent Departments, to build interagency collaboration that will help the Federal statistical community more effectively meet the information needs of the 21st century." The memorandum cites examples of successful interagency collaboration (including within and across departments). It also describes available tools for collaboration, such as use of the 1933 Economy Act, which authorizes departments and agencies to buy goods and services from each other, and the new Category Management model for federal contracting. This model could, for example, facilitate obtaining a single license for government-wide

use of statistical software in accordance with the 2014 Federal Information Technology Acquisition Reform Act (see below).

CONFIDENTIALITY AND PRIVACY PROTECTION

Protecting the confidentiality of individual information collected under a confidentiality pledge—whether from individuals, households, businesses, or other organizations—is a bedrock principle of federal statistics. Federal statistical agencies also strive to respect the privacy of individual respondents through such means as limiting the collection of information to that which is necessary for an agency's mission (see Principle 3 and Practices 7 and 8). Respect for privacy has a history in federal legislation and regulation that extends back many decades; so, too, does protection of confidentiality, except that not all federal agencies were covered.[23] With the passage of CIPSEA in 2002 (see below), a firm legislative foundation was established for confidentiality protection of statistical data government-wide.

1974 Privacy Act

The Privacy Act of 1974 (P.L. 93-579, as amended; codified at 5 USC 552a) is a landmark piece of legislation that grew out of concerns about the implications of computers, credit bureaus, proposals for national databanks, and the like on personal privacy. The act states in part (5 USC 552a(b)):

> No agency shall disclose any record which is contained in a system of records by any means of communication to any person, or to another agency, except pursuant to a written request by, or with the prior written consent of, the individual to whom the record pertains, unless disclosure of the record [is subject to one or more of 12 listed conditions.]

The defined conditions for disclosure of personal records without prior consent include use for statistical purposes by the Census Bureau, for statistical research or reporting when the records are to be transferred in a form that is not individually identifiable, for routine uses within a U.S. government agency, for preservation by the National Archives and Records Administration "as a record which has sufficient historical or other value to warrant its continued preservation by the United States Government," for law enforcement purposes, for congressional investigations, and for other administrative purposes.

[23]For example, Title 13 of the U.S. Code, providing for confidentiality protection for economic and population data collected by the U.S. Census Bureau, dates back to 1929; in contrast, the Bureau of Labor Statistics had no legal authority for its policies and practices of confidentiality protection until the passage of CIPSEA in 2002 (see National Research Council, 2003:119–121).

The Privacy Act mandates that every federal agency have in place an administrative and physical security system to prevent the unauthorized release of personal records; it also mandates that every agency publish in the *Federal Register* one or more system of records notices (SORNs) for newly created and revised systems of records that contain personally identifiable information as directed by OMB.[24] SORNs are to describe not only the records and their uses by the agency, but also procedures for storing, retrieving, accessing, retaining, and disposing of records in the system.[25]

1991 Federal Policy for the Protection of Human Subjects, 45 Code of Federal Regulations (CFR) 46, Subpart A ("Common Rule"), as Revised in 2017

The 1991 Common Rule regulations, promulgated by the U.S. Department of Health and Human Services (DHHS)[26] and signed onto by nine other cabinet departments and seven independent agencies (in their own regulations), represent the culmination of a series of DHHS regulations dating back to the 1960s (see Practice 7 and National Research Council, 2003:Ch. 3). The regulations are designed to protect individuals whom researchers wish to recruit for research studies funded by the federal government, which include surveys and other kinds of statistical data collection.[27] The regulations require that researchers obtain informed consent from prospective participants, minimize risks to participants, balance risks and benefits appropriately, select participants equitably, monitor data collection to ensure participant safety (where appropriate), and protect participant privacy and maintain data confidentiality (where appropriate). Institutional review boards (IRBs) at universities and other organizations and agencies, registered with DHHS, review research protocols to determine whether they qualify for exemption from or are subject to IRB review and, if the latter, whether the protocol satisfactorily adheres to the regulations. Some federal statistical agencies are required to submit data collection protocols to an IRB for approval; other agencies

[24]See OMB Circular A-130, *Management of Government Information Resources*, Appendix I, revised 1996. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf [April 2017].

[25]For an example of SORNs for a statistical agency, see https://www.census.gov/about/policies/privacy/sorn.html [April 2017].

[26]See: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/ [April 2017]. In addition to Subpart A of 45 CFR 46, DHHS and some other departments and agencies have signed onto Subparts B, C, and D, which pertain to pregnant women, human fetuses, and neonates; prisoners; and children, respectively.

[27]Of those departments with statistical units, all signed onto the Common Rule with the exception of the Departments of Labor and the Treasury.

maintain exemption from IRB review but follow the principles and spirit of the regulations.

An Advance Notice of Proposed Rulemaking, issued in 2011, proposed changes to the Common Rule, including revisions to the provisions for confidentiality protection.[28] A Notice of Proposed Rulemaking, which indicated responses to the extensive comments on the ANPRM, was issued in 2015; it, too, included a comment period.[29] A final rule was published January 19, 2017,[30] which is to take effect January 19, 2018 (for cooperative research involving more than one institution, the effective date is January 20, 2020). Some of the changes from the 1991 version of the Common Rule are:

- The U.S. Department of Labor became a signatory to the Common Rule; consequently, only one department that houses a federal statistical agency (U.S. Department of the Treasury) is not a signatory.
- Provisions to exempt research with human participants from IRB review were modified and enlarged, and, where appropriate, IRB review is to be focused on the adequacy of confidentiality protection.
- To assist IRBs in determining the adequacy of confidentiality protection, the Secretary of DHHS, after consultation with OMB and other federal signatories, is to issue guidance on which procedures qualify. Moreover, the federal signatories, consulting with appropriate experts, are to reexamine within 1 year and every 4 years thereafter the meaning of "identifiable private information" and identify analytic techniques that can produce same.
- Provisions are added for "broad" consent for storage, maintenance, and secondary research use of identifiable private information or biospecimens.

1997 Order Providing for the Confidentiality of Statistical Information

This order, issued in 1997, was designed to bolster the confidentiality protections afforded by statistical agencies or units (as listed in the order), some of which lacked legal authority to back up their confidentiality protection.[31] CIPSEA (see next section) placed confidentiality protection

[28] See 76 *Federal Register* 44512 (July 26, 2011). Available: https://www.federalregister.gov/d/2011-18792 [April 2017]. See also National Research Council (2014).

[29] See 80 *Federal Register* 53933 (September 8, 2015). Available: https://www.federalregister.gov/d/2015-21756 [April 2017].

[30] See 82 *Federal Register* 7149 (January 19, 2017). Available: https://www.federalregister.gov/d/2017-01058 [April 2017].

[31] See 62 *Federal Register* 35044 (June 27, 1997). Available: https://www.federalregister.gov/d/97-16934 [April 2017].

for statistical information on a strong legal footing across the entire federal government.

Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA) and Associated 2007 Implementation Guidance

The Confidential Information Protection and Statistical Efficiency Act (CIPSEA) (Title V of the E-Government Act of 2002, P.L. 107-347) is landmark legislation to strengthen the statistical system with regard to confidentiality protection and data sharing. Enactment of CIPSEA was the culmination of more than 30 years of efforts to standardize and bolster legal protections for data collected solely for statistical purposes by federal agencies while permitting limited sharing of individually identifiable business information among three statistical agencies for efficiency and quality improvement. CIPSEA has two subtitles, covering confidentiality and sharing data, respectively.

Subtitle A of CIPSEA, Confidential Information Protection

Subtitle A strengthens and extends statutory confidentiality protection for all statistical data collections of the U.S. government. Prior to CIPSEA, such protection was governed by a patchwork of laws applicable to specific agencies, judicial opinions, and agencies' practices. For all data furnished by individuals or organizations to an agency under a pledge of confidentiality for exclusively statistical purposes, Subtitle A provides that the data will be used only for statistical purposes and will not be disclosed in identifiable form to anyone not authorized by the title. It makes knowing and willful disclosure of confidential statistical data a class E felony with fines up to $250,000 and imprisonment for up to 5 years.

Subtitle A pertains not only to surveys, but also to collections by a federal agency for statistical purposes from nonpublic administrative records (e.g., confidential state government agency records). Data covered under Subtitle A are not subject to release under a Freedom of Information Act request.

Subtitle B of CIPSEA, Statistical Efficiency

Subtitle B permits the Bureau of Economic Analysis (BEA), the Bureau of Labor Statistics (BLS), and the Census Bureau to share individually identifiable business data for statistical purposes. The subtitle has three main purposes: (1) to reduce respondent burden on businesses; (2) to improve the comparability and accuracy of federal economic statistics by permitting these three agencies to reconcile differences among sampling frames, business classifications, and business reporting; and (3) to increase understanding of the U.S. economy and improve the accuracy of key national

indicators, such as the National Income and Product Accounts. Several data-sharing projects have been initiated under Subtitle B.

The subtitle does not permit sharing among BEA, BLS, and the Census Bureau of any individually identifiable tax return data that originate from the Internal Revenue Service (IRS). This limitation currently blocks some kinds of business data sharing, such as those for sole proprietorships, which are important for improving the efficiency and quality of business data collection by statistical agencies. For tax return information, data sharing is limited to a small number of items for specified uses by a small number of specific agencies (under Title 26, Section 6103 of the U.S. Code, and associated Treasury Department regulations, as modified in the 1976 Tax Reform Act). The law provides access to specific tax return items by the Census Bureau for use in its population estimates program and economic census and survey programs, by the National Agricultural Statistics Service for conducting the Census of Agriculture, by the Congressional Budget Office for long-term models of the Social Security and Medicare programs, and by BEA for producing the National Income and Product Accounts. (Prior to the 1976 act, the President could issue an executive order authorizing access to tax records.) The governing statute would have to be modified to extend sharing of tax return items to agencies not specified in the 1976 legislation.

A proposal for legislation to expand access to IRS information for limited statistical purposes has been developed through interagency discussions: it would authorize the Bureau of Labor Statistics to receive limited business data from the Census Bureau (comingled with business tax information) for the purpose of synchronizing the two agencies' business lists. It would also authorize BEA to receive business tax information for partnerships and sole proprietors with receipts exceeding a yet-to-be-determined threshold. Such access would allow BEA to improve the measurement of income and international transactions in the national accounts. This proposal has yet to be introduced as a bill in Congress.

Associated 2007 CIPSEA Implementation Guidance

OMB is charged to oversee and coordinate the implementation of CIPSEA; after a thorough interagency development and coordination process, OMB released final guidance for CIPSEA in 2007 (U.S. Office of Management and Budget, 2007). The guidance, which pertains to both Subtitles A and B, covers such topics as the steps that agencies must take to protect confidential information; wording of confidentiality pledges in materials that are provided to respondents; steps that agencies must take to distinguish any data or information they collect for nonstatistical purposes and to provide proper notice to the public of such data; and

ways in which agents (e.g., contractors, researchers) may be designated to use individually identifiable information for analysis and other statistical purposes and be held legally responsible for protecting the confidentiality of that information.

A key provision of the CIPSEA guidance defines statistical agencies and units, which are the only federal agencies that may assign agent status for confidentiality protection purposes to contractors, researchers, or others. The guidance defines a statistical agency or unit as "an agency or organizational unit of the executive branch whose activities are predominantly the collection, compilation, processing, or analysis of information for statistical purposes." A total of 16 agencies are currently so recognized (see Appendix B):

- Of the 16 agencies, 12 were enumerated in OMB's 1997 confidentiality order and carried over into the CIPSEA implementation guidance: they are all principal statistical agencies and thereby members of the Interagency Council on Statistical Policy (see Appendix B). The Office of Research, Evaluation, and Statistics in the U.S. Social Security Administration is also a principal statistical agency and an ICSP member but was not named in the order or the guidance; it later applied for recognition as a statistical unit (see below).
- In 2007, OMB recognized two more units that applied for designation under the procedures outlined in the guidance: the Center for Behavioral Health Statistics and Quality in the Office of Applied Studies in the Substance Abuse and Mental Health Services Administration, U.S. Department of Health and Human Services, and the Microeconomic Surveys Section of the Board of Governors of the Federal Reserve.
- Since 2007, OMB has recognized two additional units: the Office of Research, Evaluation, and Statistics and the National Animal Health Monitoring System Program Unit in the Animal and Plant Health Inspection Service, U.S. Department of Agriculture.

E-Government Act of 2002, Section 208, and Associated 2003 Implementation Guidance

Section 208 of the E-Government Act of 2002 (P.L. 107-347) requires federal agencies to conduct a privacy impact assessment (PIA) whenever an agency develops or obtains information technology that handles individually identifiable information or whenever the agency initiates a new collection of individually identifiable information.[32] The assessment

[32]Section 208 also mandates that OMB lead interagency efforts to improve federal information technology and use of the Internet for government services.

is to be made publicly available and cover topics such as what information is being collected and why, with whom the information will be shared, what provisions will be made for informed consent regarding data sharing, and how the information will be secured. Typically, PIAs cover not only privacy issues, but also confidentiality, integrity, and availability issues.[33] OMB was required to issue guidance for development of the assessments, which was done in a September 26, 2003, memorandum (M-03-22) from the OMB director to the heads of executive agencies and departments.[34]

Section 208, together with Title III, FISMA (see below), and Title V, CIPSEA (see above), of the 2002 E-Government Act are the latest in a series of laws beginning with the Privacy Act of 1974 (see above) that govern access to individual records maintained by the federal government (see also "Federal Cybersecurity Enhancement Act of 2015," below).

Federal Information Security Management Act of 2002 (FISMA)

The Federal Information Security Management Act (FISMA) was enacted in 2002 as Title III of the E-Government Act of 2002 (P.L. 107-347). The act was meant to bolster computer and network security in the federal government and affiliated parties (such as government contractors) by mandating yearly audits.

FISMA imposes a mandatory set of processes that must be followed for all information systems used or operated by a federal agency or by a contractor or other organization on behalf of a federal agency. These processes must follow a combination of Federal Information Processing Standards documents, the special publications issued by the National Institute of Standards and Technology (SP-800 series), and other legislation pertinent to federal information systems, such as the Privacy Act of 1974 and the Health Insurance Portability and Accountability Act of 1996.

The first step is to determine what constitutes the "information system" in question. There is no direct mapping of computers to an information system; rather, an information system can be a collection of individual computers put to a common purpose and managed by the same system owner. The next step is to determine the types of information in the system and categorize each according to the magnitude of harm that would result if the system suffered a compromise of confidentiality, integrity, or availability. Succeeding steps are to develop complete system documentation, conduct

[33] See, e.g., the available privacy impact assessments prepared by the Census Bureau at https://www.census.gov/about/policies/privacy/pia/list_of_available_pias.html [April 2017].

[34] Available: https://obamawhitehouse.archives.gov/omb/memoranda_m03-22/ [April 2017].

a risk assessment, put appropriate controls in place to minimize risk, and arrange for an assessment and certification of the adequacy of the controls.

FISMA affects federal statistical agencies directly in that each agency must follow the FISMA procedures for its own information systems. In addition, some departments are taking the position that all information systems in a department constitute a single information system for purposes of FISMA: those departments are taking steps to require that statistical agencies' information systems and personnel be incorporated into a centralized department-wide system.

2014 Federal Information Technology Acquisition Reform Act (FITARA) and Associated 2015 and 2016 Implementation Guidance

The Federal Information Technology Acquisition Reform Act (FITARA) was enacted on December 19, 2014, to respond to such federal information technology (IT) challenges as duplicate IT spending among and within agencies; difficulty in understanding the cost and performance of IT investments; and inability to benchmark IT spending between federal and private-sector counterparts. FITARA has four major objectives: (1) strengthening the authority over and accountability for IT costs, performance, and security of agency chief information officers (CIOs); (2) aligning IT resources with agency missions and requirements; (3) enabling more effective planning for and execution of IT resources; and (4) providing transparency about IT resources across agencies and programs. It requires agencies (defined as cabinet departments and independent agencies) to pursue a strategy of consolidation of agency data centers, charges agency CIOs with the responsibility for implementing FITARA, and charges the U.S. Government Accountability Office with producing quarterly scorecards to assess how well agencies are meeting the FITARA objectives.

The director of OMB issued implementation guidance for FITARA, M-15-14, *Management and Oversight of Federal Information Technology*, on June 20, 2015.[35] This memorandum explicitly stated that agencies must implement the FITARA guidance to ensure that information acquired under a pledge of confidentiality solely for statistical purposes is used exclusively for those purposes. It also provided a "Common Baseline for IT Management," which lays out FITARA responsibilities of CIOs and other agency officials, such as the chief financial officer and program officials. On May 4, 2016, the federal CIO and the administrator of OIRA, both in OMB, jointly issued *Supplemental Guidance on the Implementation of M-15-14 "Management and Oversight of Federal Information Technology"—Applying FITARA Common*

[35] Available: https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2015/m-15-14.pdf [April 2017].

Baseline to Statistical Agencies and Units (U.S. Office of Management and Budget, 2016c). This supplemental guidance posed questions for CIOs and other officials, including statistical agency heads, to address when implementing FITARA for statistical agency programs. The questions refer to the fundamental responsibilities of federal statistical agencies outlined in *Statistical Policy Directive No. 1* (see above), which include confidentiality protection and meeting deadlines for key statistics.

Federal Cybersecurity Enhancement Act of 2015

The Federal Cybersecurity Enhancement Act of 2015 is Title II, Subpart B, of the Cybersecurity Act of 2015, which was attached as a rider to the Consolidated Appropriations Act of 2016, and so became law when the appropriations bill was signed on December 18, 2015, and became P.L. 114-113. The impetus for Title II, Subpart B, is the efforts of the U.S. Department of Homeland Security (DHS), dating back to 2003, to deploy systems for detection and prevention of intrusions ("hacking") into federal government information networks (see Latham & Watkins, 2016:3). As of the end of 2015, this technology, known as EINSTEIN, covered only 45 percent of federal network access points. The act requires DHS to "make [EINSTEIN] available" to all federal agencies within 1 year, and thereafter requires all agencies to "apply and continue to utilize the capabilities" across their networks.

The technology, currently in version E3A, has been welcomed by federal statistical agencies, but agencies initially were concerned about a DHS interpretation of the act that would allow DHS staff to monitor traffic on agency networks and follow up on actual or likely intrusions. Such surveillance by DHS staff could lead to violations of agencies' pledges to protect the confidentiality of information provided by individual respondents for statistical purposes, which state that only statistical agency employees or sworn agents can see such information. Ultimately, DHS retained its surveillance authority, and statistical agencies modified their confidentiality pledges. As described in a *Federal Register* notice from the U.S. Census Bureau (other statistical agencies have issued similar notices):[36]

> DHS and Federal statistical agencies, in cooperation with their parent departments, have developed a Memorandum of Agreement for the installation of Einstein 3A cybersecurity protection technology to monitor their Internet traffic and have incorporated an associated Addendum on Highly Sensitive Agency Information that provides additional protection and enhanced security handling of confidential

[36] "Agency Information Collection Activities; Request for Comments; Revision of the Confidentiality Pledge Under Title 13 United States Code, Section 9," 81 *Federal Register* 94321 (December 23, 2016). Available: https://www.federalregister.gov/d/2016-30959 [April 2017].

> statistical data. However, many current Title 13, U.S.C. and similar statistical confidentiality pledges promise that respondents' data will be seen only by statistical agency personnel or their sworn agents. Since it is possible that DHS personnel could see some portion of those confidential data in the course of examining the suspicious Internet packets identified by Einstein 3A sensors, statistical agencies need to revise their confidentiality pledges to reflect this process change.

The Bureau of Labor Statistics led an interagency research program to test revised wording with samples of respondents, and agencies revised their pledges accordingly. As an example, the Census Bureau's revised pledge, for example, provided in 81 *Federal Register* 94321 (December 23, 2016), states:

> The U.S. Census Bureau is required by law to protect your information. The Census Bureau is not permitted to publicly release your responses in a way that could identify you. Per the Federal Cybersecurity Enhancement Act of 2015, your data are protected from cybersecurity risks through screening of the systems that transmit your data.

INFORMATION QUALITY, PEER REVIEW, PERFORMANCE EVALUATION, SCIENTIFIC INTEGRITY, AND TRANSPARENCY

Since 2000, there has been heightened interest in the Congress and the executive branch regarding the quality of scientific evidence, including federal statistics. Legislation and guidance from OMB and the Office of Science and Technology Policy (OSTP) have addressed concerns of information quality, performance evaluation, scientific integrity, and transparency.

The Information Quality Act of 2000 and Associated 2002 Implementation Guidance

The Information Quality Act of 2000 (P.L. 106-554) directed OMB to issue government-wide guidelines that "provide policy and procedural guidance to Federal agencies for ensuring and maximizing the quality, objectivity, utility, and integrity of information (including statistical information) disseminated by Federal agencies." It also required federal agencies to develop their own implementing procedures, including "administrative mechanisms allowing affected persons to seek and obtain correction of information maintained and disseminated by the agency." After a public comment period, OMB issued government-wide guidelines on February 22, 2002.[37]

[37] See 67 *Federal Register* 8452 (February 22, 2002). Available: https://www.federalregister.gov/d/R2-59 [April 2017].

2002 Federal Statistical Agency Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Disseminated Information

A few months after OMB issued implementation guidance in February 2002 for the 2000 Information Quality Act (see above), 13 principal statistical agencies issued a notice outlining a common approach to the development and provision of guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of disseminated information.[38] The notice directed people to the websites of each agency for more information and to learn how to comment on draft guidelines. Each agency then finalized its own guidelines.[39] The information quality framework developed by the agencies was followed in the 2006 revision of OMB's standards and guidelines for statistical surveys (see *Statistical Policy Directive No. 2* above).

2004 OMB Final Information Quality Bulletin for Peer Review

Consistent with the 2000 Information Quality Act (see above), OMB developed guidance for federal agencies with regard to seeking peer review of the policy-relevant scientific information an agency disseminates. After two rounds of public comment, OMB issued the *Final Information Quality Bulletin for Peer Review* on December 16, 2004 (hereafter referred to as the *Bulletin*): it requires federal agencies to conduct a peer review of "influential scientific information" before the information is released to the public.[40] "Influential scientific information" is defined as "scientific information the agency reasonably can determine will have or does have a clear and substantial impact on important public policies or private sector decisions" (U.S. Office of Management and Budget, 2005:2667). The *Bulletin* allows agencies discretion to select the type of peer review process most appropriate for a given scientific information product. Research reports and nonroutine collections by statistical agencies that can be considered "influential scientific information" are covered under the guidelines, but "routine statistical information released by federal statistical agencies (e.g., periodic demographic and economic statistics) and the analysis of these data to compute standard indicators and trends (e.g., unemployment and poverty rates)" are excluded (U.S. Office of Management and Budget, 2005:2674).

[38] See 67 *Federal Register* 38467 (June 4, 2002). Available: https://www.federalregister.gov/d/02-13892 [April 2017].

[39] See, e.g., https://www.census.gov/about/policies/quality/guidelines.html [April 2017].

[40] Available: https://obamawhitehouse.archives.gov/omb/memoranda_fy2005_m05-03/ [April 2017].

2010 Government Performance and Results Modernization Act

The Government Performance and Results Modernization Act of 2010, which supersedes the Performance Assessment Rating Tool and the Government Performance and Results Act of 1993, was signed into law on January 4, 2011.[41] It requires performance assessment of government programs for purposes of evaluating agency performance and improvement. In carrying out the provisions of the act, the director of OMB coordinates with agencies to develop the federal government performance plan. The act requires all federal agencies, with few exceptions, to establish performance indicators to be used in measuring or assessing progress toward their identified performance goals and an objective, quantifiable, and measurable means by which to compare actual program results with these established performance goals. Additionally, each agency must describe how it will ensure the accuracy and reliability of the data used, including validation of measures, data sources, required level of accuracy, data limitations, and management of those limitations.

The broad scope of agencies affected by this act, and the use of the act in making budgetary decisions based on measured achievement toward program goals, has fostered added focus among many agencies on how to collect high quality data and produce sound government statistics with which to conduct rigorous program evaluation. The proposed addendum, issued in 2016, to *Statistical Policy Directive No. 4* (see above), prescribes a program of annual performance reviews for federal statistical products.

2010 Office of Science and Technology Policy (OSTP) Memorandum on Scientific Integrity

In a memorandum on scientific integrity issued March 9, 2009, President Obama stated:[42]

> The public must be able to trust the science and scientific process informing public policy decisions. Political officials should not suppress or alter scientific or technological findings and conclusions. If scientific and technological information is developed and used by the Federal Government, it should ordinarily be made available to the public. To the extent permitted by law, there should be transparency in the preparation, identification, and use of scientific and technological information in policymaking. The selection of scientists and technology professionals for positions in the executive branch should be based on

[41] Available: http://www.gpo.gov/fdsys/pkg/PLAW-111publ352/pdf/PLAW-111publ352.pdf [April 2017].

[42] Available: http://www.whitehouse.gov/the_press_office/Memorandum-for-the-Heads-of-Executive-Departments-and-Agencies-3-9-09 [April 2017].

their scientific and technological knowledge, credentials, experience, and integrity.

The President directed OSTP to develop a strategy to ensure scientific integrity in government decision making. In response, the OSTP director issued a memorandum on December 17, 2010, that called for executive departments and agencies to develop policies to "ensure a culture of scientific integrity," "strengthen the actual and perceived credibility of Government research," "facilitate the free flow of scientific and technological information, consistent with privacy and classification standards," and "establish principles for conveying scientific and technological information to the public."[43] The memorandum included guidance on the selection of candidates for scientific positions, independent peer review, whistleblower protections, promoting access to scientific and technological information in online open formats, and agency communications. It also provided guidance on public communications, use of federal advisory committees, professional development of government scientists and engineers, and implementation.

Each statistical agency is covered by its department's scientific integrity policies. In addition, the principal statistical agencies developed a *Statement of Commitment to Scientific Integrity* that documents in a single place their response to the OSTP memorandum. The statement articulates how the *Principles and Practices for a Federal Statistical Agency, Fourth Edition* (National Research Council, 2009), various OMB statistical policy directives and standards, and each agency's information quality guidelines together form "the foundation for achieving and maintaining scientific integrity within and among the principal statistical agencies."[44]

2013 OSTP Memorandum on Increasing Access to the Results of Scientific Research

On February 23, 2013, the OSTP director issued a memorandum for heads of executive departments and agencies on "Increasing Access to the Results of Scientific Research."[45] Citing the importance of scientific research for driving improvements in "areas such as health, energy, the environment, agriculture, and national security," the memorandum outlined the administration's commitment to:

[43] Available: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/scientific-integrity-memo-12172010.pdf [April 2017].

[44] See http://www.census.gov/content/dam/Census/about/about-the-bureau/policies_and_notices/scientificintegrity/Scientific_Integrity_Statement_of_the_Principal_Statistical_Agencies.pdf [April 2017].

[45] Available: https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf [April 2017].

> ensuring that, to the greatest extent and with the fewest constraints possible and consistent with law and the objectives set out [in the memorandum], the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community. Such results include peer-reviewed publications and digital data.

The memorandum directed federal agencies with over $100 million in annual research and development expenditures to develop a plan "to support increased public access to the results of research funded by the Federal Government. This includes any results published in peer-reviewed scholarly publications that are based on research that directly arises from Federal funds . . . ." The memorandum further directed agencies to develop plans for ensuring archiving of and access to data underlying federally funded research and the associated documentation or metadata. The memorandum listed various topics to be covered in each agency's plan, including that the plan be posted on the agency's website and provide for protection of the confidentiality of individual respondents' information.

REFERENCES FOR APPENDIX A

All URL addresses valid as of April 2017. *Please note that, as of February 2017, U.S. Office of Management and Budget documents previously on the main whitehouse.gov site were relocated to an archived site. Wherever possible, a citation is provided to a stable site such as federalregister.gov.*

Anderson, M. J. (2015). *The American Census: A Social History* (3rd ed.). New Haven, CT: Yale University Press.

Duncan, J. W. and W. C. Shelton (1978). *Revolution in United States Statistics, 1926–1976*. Office of Federal Statistical Policy and Standards, U.S. Department of Commerce. Washington, DC: U.S. Government Printing Office. Available: https://catalog.hathitrust.org/Record/007886111/Home.

Economic Classification Policy Committee of the United States (2003). Overview of NAPCS objectives, guidance, and implementation strategy and goals: A United States perspective. NAPCS discussion paper prepared for the Trilateral Steering Group Meeting, Washington, DC, May 5–9, 2003. Available: http://www.census.gov/eos/www/napcs/papers/overviewobj.pdf.

European Statistical System Committee (2011). *European Statistics Code of Practice for the National and Community Statistical Authorities*. Luxembourg: Eurostat. Available: http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-32-11-955.

Interagency Technical Working Group on Developing a Supplemental Poverty Measure (2010). Observations from the Interagency Technical Working Group on Developing a Supplemental Poverty Measure. Available: https://www.census.gov/hhes/povmeas/methodology/supplemental/research/SPM_TWGObservations.pdf.

Latham & Watkins (2016). What you need to know about the Cybersecurity Act of 2015. *Client Alert Commentary* No. 1927 (February 18). Available: https://www.lw.com/thoughtLeadership/lw-Cybersecurity-Act-of-2015.

National Academies of Sciences, Engineering, and Medicine (2016). *Rationalizing Rural Area Classifications for the Economic Research Service: Workshop Summary*. G.S. Wunderlich, rapporteur. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://www.nap.edu/catalog/21843/rationalizing-rural-area-classifications-for-the-economic-research-service-a. doi: 10.17226/21843.

National Research Council (1984). *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*. Report of the Advanced Research Seminar on Cognitive Aspects of Survey Methodology, T.B. Jabine, M.L. Straf, J.M. Tanur, and R. Tourangeau, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/930.

National Research Council (1995). *Measuring Poverty—A New Approach*. Panel on Poverty and Family Assistance, C.F. Citro and R.T. Michael, eds. Committee on National Statistics. Washington, DC: National Academy Press. Available: https://doi.org/10.17226/4759.

National Research Council (2003). *Protecting Participants and Facilitating Social and Behavioral Sciences Research*. Panel on Institutional Review Boards, Surveys, and Social Science Research, C.F. Citro, D.R. Ilgen, and C.B. Marrett, eds. Committee on National Statistics and Board on Behavioral, Cognitive, and Sensory Sciences. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/10638.

National Research Council (2006). *Once, Only Once, and in the Right Place: Residence Rules in the Decennial Census*. Panel on Residence Rules in the Decennial Census, D.L. Cork and P.R. Voss, eds. Committee on National Statistics. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/11727.

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. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/12564.

National Research Council (2013). *Principles and Practices for a Federal Statistical Agency, Fifth Edition*. Committee on National Statistics, C.F. Citro and M.L. Straf, eds.. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/18318.

National Research Council (2014). *Proposed Revisions to the Common Rule for the Protection of Human Subjects in the Behavioral and Social Sciences*. Committee on Revisions to the Common Rule for the Protection of Human Subjects in Research in the Behavioral and Social Sciences, Board on Behavioral, Cognitive, and Sensory Sciences, Committee on National Statistics, and Committee on Population. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/18614.

Norwood, J. L. (1995). *Organizing to Count: Change in the Federal Statistical System*. Washington, DC: The Urban Institute Press.

Norwood, J. L. (2016). Politics and federal statistics. *Statistics and Public Policy 3*(1), 1–8. American Statistical Association Distinguished Statistician Lecture, 2000. Available: http://dx.doi.org/10.1080/2330443X.2016.1241061.

United Nations Statistical Commission (2014, March 3). *Fundamental Principles of Official Statistics*. New York: United Nations. Endorsed by the General Assembly as Resolution A/RES/68/261. Available: http://unstats.un.org/unsd/dnss/gp/FP-New-E.pdf.

U.S. Department of Commerce (1978). Transfer of responsibility for certain statistical standards from OMB to Commerce. 43 *Federal Register* 19260 (May 4, 1978).

U.S. Office of Management and Budget (1985). Statistical Policy Directive No. 3: Compilation, release, and evaluation of principal federal economic indicators. 50 *Federal Register* 38932 (September 25, 1985). Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/omb/inforeg/statpolicy/dir_3_fr_09251985.pdf.

U.S. Office of Management and Budget (2005). Final information quality bulletin for peer review. 70 *Federal Register* 2664 (January 14, 2005). Available: https://www.federalregister.gov/d/05-769.

U.S. Office of Management and Budget (2006). Statistical Policy Directive No. 2: Standards and guidelines for statistical surveys. 71 *Federal Register* 55522 (September 22, 2006). Available: https://www.federalregister.gov/d/06-8044.

U.S. Office of Management and Budget (2007). Implementation guidance for Title V of the E-Government Act, Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA). 72 *Federal Register* 33362 (June 15, 2007). Available: https://www.federalregister.gov/d/E7-11542.

U.S. Office of Management and Budget (2008). Statistical Policy Directive No. 4: Release and dissemination of statistical products produced by federal statistical agencies. 73 *Federal Register* 12625 (March 7, 2008). Available: https://www.federalregister.gov/d/E8-4570.

U.S. Office of Management and Budget (2014a). Guidance for providing and using administrative data for statistical purposes. OMB Memorandum M-14-06. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2014/m-14-06.pdf.

U.S. Office of Management and Budget (2014b). Statistical Policy Directive No. 1: Fundamental responsibilities of federal statistical agencies and recognized statistical units. 79 *Federal Register* 71609 (December 2, 2014). Available: https://www.federalregister.gov/d/2014-28326.

U.S. Office of Management and Budget (2015). Improving statistical activities through interagency collaboration. OMB Memorandum M-15-15. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2015/m-15-15.pdf.

U.S. Office of Management and Budget (2016a). Statistical Policy Directive No. 2: Standards and guidelines for statistical surveys; Addendum: Standards and guidelines for cognitive interviews. 81 *Federal Register* 70586 (October 12, 2006). Available: https://www.federalregister.gov/d/2016-24607.

U.S. Office of Management and Budget (2016b). Statistical Policy Directive No. 4 Addendum: Release and dissemination of statistical products produced by federal statistical agencies and recognized statistical units [adding section 10, performance review]. 81 *Federal Register* 71538 (October 17, 2016). Available: https://www.federalregister.gov/d/2016-25049.

U.S. Office of Management and Budget (2016c, May 4). Supplemental guidance on the implementation of M-15-14 "Management and oversight of federal information technology"—Applying FITARA common baseline to statistical agencies and units. Memorandum from Tony Scott, Federal Chief Information Officer, and Howard Shelanski, Administrator, Office of Regulatory Affairs. Available: https://management.cio.gov/assets/docs/FITARA_Guidance_Statistical_Agencies_and_Units_OMB.pdf.

– B –

Organization of the Federal Statistical System

OVERVIEW

This appendix begins with an overview of the U.S. statistical system as a whole, including structure and budget. It then briefly summarizes the statistical functions of the U.S. Office of Management and Budget (OMB), the principal statistical agencies and other recognized statistical units, and a selection of major statistical programs housed or sponsored by other agencies.

Structure

The United States has a highly decentralized statistical system in contrast with other developed countries (see Norwood, 1995). Essentially, the system grew by adding separate agencies whenever the need for objective empirical information on a particular aspect of the economy, society, or environment came to the fore (see "Brief History of the U.S. Federal Statistical System" in Part I). Periodic recommendations from presidential commissions and other initiatives to consolidate one or more of the principal statistical agencies have never been adopted.

The 13 principal statistical agencies, which sit on the Interagency Council on Statistical Policy (ICSP), are: the Bureau of Economic Analysis (BEA); Bureau of Justice Statistics (BJS); Bureau of Labor Statistics (BLS); Bureau of Transportation Statistics (BTS); Census Bureau; Economic Research Service

(ERS); Energy Information Administration (EIA); National Agricultural Statistics Service (NASS); National Center for Education Statistics (NCES); National Center for Health Statistics (NCHS); National Center for Science and Engineering Statistics (NCSES); Office of Research, Evaluation, and Statistics (ORES); and Statistics of Income Division (SOI).[1]

At the center of the federal statistical system is the OMB Office of Information and Regulatory Affairs (OIRA), which includes the Statistical and Science Policy (SSP) Office headed by the chief statistician, a senior executive civil service position. OIRA also includes clearance officers who review individual survey and other information collection requests from most agencies: SSP staff clear information requests from many of the principal statistical agencies and consult with the OIRA desk officers for the other agencies. SSP staff also consult with other parts of OMB that are responsible for recommending budgets for statistical agencies.

The chief statistician chairs the ICSP, whose 14 member agencies are in 10 cabinet departments and 2 independent agencies. In addition to the ICSP members, there are about 115 other executive branch agencies that conduct substantial statistical activities defined by SSP as $500,000 or more in annual spending (see Figure I.1 in Part I; see also U.S. Office of Management and Budget, 2017:Table 1).

There are some important federal agencies that have statistical activities that are not included in the OMB annual compilation because they are not part of the executive branch. These agencies include the Congressional Budget Office, which develops and applies projection models for the budgetary impact of current and proposed federal programs; the Federal Reserve Board, which compiles the widely used Flow of Funds report and other statistical series and periodically conducts the Survey of Consumer Finances (see "Other Recognized Statistical Units," below); and the U.S. Government Accountability Office, which uses statistical data in evaluations of government programs.

Budget

For fiscal 2016, OMB estimated the government provided $7.2 billion in direct funding for federal statistical activities of $500,000 or more, in 128 agencies (U.S. Office of Management and Budget, 2017:Table 1). "Statistical activities" are defined by U.S. Office of Management and Budget (2017:3)

[1]The chief statistician of the United States, who chairs the ICSP, added the National Center for Veterans Analysis and Statistics (NCVAS) as a rotating member of the ICSP in January 2017, under the authority of the 1995 Paperwork Reduction Act (see Appendix A). NCVAS replaced the Office of Environmental Information (OEI), which had been the rotating member for many years. Neither of these two agencies is a principal statistical agency or a recognized statistical unit as defined below; they are described in "Other Statistical Programs" below.

to include a wide variety of functions, including not only survey and census design, data collection, and dissemination, but also data analysis, forecasting, and modeling. The total, which includes $0.8 billion in direct funding for the 2020 decennial census, covers programs carried out by the 13 designated principal statistical agencies and 115 other policy, research, and programmatic agencies.[2]

"Direct funding" covers congressional appropriations to an agency. Some agencies (e.g., the Census Bureau) carry out statistical activities for other agencies on a cost-reimbursable basis. The funding for these activities is allocated to the sponsoring agency and not to the data collection agency. OMB's annual compilation of statistical programs generally includes the entire budget for each of the 13 principal agencies; other agencies determine which parts of their budgets should be included according to the OMB definition of statistical activities.

In fiscal 2016, the 13 principal agencies accounted for 38 percent of the total budget authority for statistical activities, excluding the 2020 census; with the 2020 census authority included, they accounted for 43 percent.

Not all of the work of the principal agencies is carried out in-house. For fiscal 2017, OMB estimated that 43 percent of the total budget authority of the 13 agencies (including the decennial census) would be used to purchase statistical services from other organizations (U.S. Office of Management and Budget, 2017:Table 2). These purchases would include reimbursing state and local governments for administrative records (e.g., birth and death records provided to the National Center for Health Statistics and unemployment insurance establishment data provided to the Bureau of Labor Statistics); reimbursing private organizations for systems development, data collection, and data analysis services; and reimbursing other federal agencies, principally the Census Bureau, for data collection and other services.

In dollar terms, the Census Bureau, the National Center for Education Statistics, the Bureau of Labor Statistics, and the National Center for Health Statistics planned to allocate the largest amounts of their fiscal 2017 budgets to purchasing statistical services; by percentage of budget authority, the National Center for Education Statistics, the National Center for Health Statistics, the National Center for Science and Engineering Statistics, and the Bureau of Justice Statistics planned to be the largest users of purchased services. These patterns have remained roughly constant over the past decade (see U.S. Office of Management and Budget, 2001:Table 3).

[2]In the total number of policy, research, and program agencies, every institute or center of the National Institutes of Health is treated as a separate agency.

U.S. OFFICE OF MANAGEMENT AND BUDGET

The 1995 reauthorization of the Paperwork Reduction Act of 1980 and other legislation give OIRA the authority to approve all agency information collection requests, including all survey and other statistical information requests. OIRA also reviews all proposed economically significant federal regulations (those estimated to have greater than $100 million impact on the economy).

SSP, the chief statistician's office, establishes statistical policies and standards, identifies priorities for improving programs, evaluates statistical programs for compliance with OMB guidance, reviews statistical agency budgets, approves information collections for many of the principal statistical agencies, provides guidance to OIRA desk officers who review statistical information requests from other federal agencies, and coordinates U.S. participation in international statistical activities.[3] It currently has a staff of five professionals, who focus on statistical issues, typically augmented by 2–3 professional staff on short-term details to OMB from a statistical agency.

As required by the Paperwork Reduction Act, SSP annually issues *Statistical Programs of the United States Government* (the "blue book"; see U.S. Office of Management and Budget, 2017). It also prepares a chapter each year in the *Analytical Perspectives* volume of the President's budget, which provides a cross-cutting analysis of the budget requests for the principal statistical agencies (see "Strengthening Federal Statistics" in U.S. Office of Management and Budget, 2016). SSP also sponsors the Federal Committee on Statistical Methodology and other interagency collaborations, such as the Interagency Forum on Aging-Related Statistics and the Interagency Forum on Child and Family Statistics (see Practice 13).

Appendix A provides background information on the Paperwork Reduction Act, statistical policy directives issued by the chief statistician's office, and other legislation and OMB documents that affect the U.S. statistical system.

PRINCIPAL STATISTICAL AGENCIES

This section covers the 13 principal statistical agencies identified by OMB;[4] they are all members of the ICSP. The information includes origins, authorizing legislation or other authority, status of agency head (presidential appointee, career senior executive service official), and budget and full-time permanent staffing levels in 2016 (it is drawn largely from agency websites;

[3]See https://obamawhitehouse.archives.gov/omb/inforeg_statpolicy [April 2017].

[4]See https://obamawhitehouse.archives.gov/omb/inforeg_statpolicy/bb-principal-statistical-agencies-recognized-units [April 2017]. OMB defines these agencies as having "statistical work as their primary mission."

see also U.S. Office of Management and Budget, 2017:Tables 1 and 2). The agencies are discussed in alphabetical order.

Bureau of Economic Analysis

The Bureau of Economic Analysis (BEA; see https://bea.gov/) is part of the Economics and Statistics Administration in the U.S. Department of Commerce (as is the Census Bureau), headed by the Under Secretary for Economic Affairs. The BEA director is a career senior executive service appointee, and the agency has a full-time staff of about 499 people and had direct funding in fiscal 2016 of $105.1 million.

BEA's history traces back to 1820 when the Secretary of the Treasury was directed by Congress to compile and publish statistics on U.S. foreign commerce. Three 20th-century predecessors of BEA were all located in the Department of Commerce: the Bureau of Statistics (1903–1912); the Bureau of Foreign and Domestic Commerce (1912–1945); and the Office of Business Economics (1945–1972).

BEA produces statistics on the performance of the nation's economy. Although it collects some source data, it primarily compiles data from the Census Bureau, the Bureau of Labor Statistics, and other agencies as input to estimating the National Income and Product Accounts (NIPAs), which include estimates of the nation's gross domestic product (GDP) and related measures. GDP, which was recognized by the Department of Commerce as its greatest achievement of the 20th century in a December 2009 ceremony, has major influence on U.S. financial markets.

Since the NIPAs were first developed in the aftermath of the Great Depression, BEA has extended its estimates to cover a wide range of economic activities for the nation, regions, and industries and also for the nation's position in the world economy. BEA also produces "satellite accounts" in such areas as health care, travel and tourism, and arts and culture production. Satellite accounts provide a framework for testing alternative economic assumptions without disruption to BEA's core economic accounts.

Bureau of Justice Statistics

The Bureau of Justice Statistics (BJS; see http://www.bjs.gov/), in the U.S. Department of Justice (DOJ), was formally established by the Justice Systems Improvement Act of 1979 (P.L. 96-157). It inherited statistical functions that had previously been vested in an office of the Law Enforcement Assistance Administration (which had been established in 1968).

BJS is housed in the DOJ's Office of Justice Programs (OJP), which also contains the National Institute of Justice (a research agency) and

other agencies that are primarily focused on providing grant and technical assistance to state and local governments and law enforcement agencies. BJS's director is a presidential appointee (not requiring Senate confirmation—a change as of August 2012) and reports to the assistant attorney general for OJP. BJS has a full-time staff of about 54 and had direct funding in fiscal 2016 of $50.2 million.

The centerpiece of BJS's data collections is the National Crime Victimization Survey (originally the National Crime Survey), which has served as one of the nation's principal measures of crime (particularly crime not reported to police) since its full-scale implementation in 1972. Data collection for most BJS surveys is conducted by the Census Bureau or private contractors, and BJS estimated that 59 percent of its anticipated fiscal 2017 budget authority would be spent on purchased services (U.S. Office of Management and Budget, 2017:Table 2).

BJS publishes annual statistics on criminal victimization, populations under correctional supervision, law enforcement management and administration, case processing in the state and federal courts, and sexual violence in prisons under the 2003 Prison Rape Elimination Act. Its periodic data series cover the administration of law enforcement agencies and correctional facilities, prosecutorial practices and policies, state court case processing, felony convictions, criminal justice expenditure and employment, civil case processing in state courts, and special studies on other criminal justice topics.

Bureau of Labor Statistics

The Bureau of Labor Statistics (BLS; https://www.bls.gov/) is an agency of the U.S. Department of Labor. It is responsible for the production of some of the nation's most sensitive and important economic data, including unemployment statistics and consumer and producer price indexes, which are closely watched by the public, Congress, other federal agencies, state and local governments, businesses, and labor organizations. The BLS commissioner is a presidential appointee, subject to Senate confirmation, and serves for a fixed term of 4 years. BLS has a full-time staff of about 2,148 people and had direct funding in fiscal 2016 of $609 million.

The history of the BLS dates back to 1884, when the Bureau of Labor was established in the Interior Department to collect information about employment and labor. It was made an independent (subcabinet) agency by the Department of Labor Act in 1888; it was made part of the Department of Commerce and Labor (as the Bureau of Labor) in 1903, and it was transferred to the newly created Department of Labor in 1913.

BLS programs use a variety of data collection methods and sources. Certain wage, benefit, employment, and price data are collected by BLS staff

in offices throughout the country, who contact employers, households, and businesses directly. BLS also has contractual arrangements with various state agencies to collect much of the data it publishes on employment and workplace safety and health. Its contractual arrangements with the Census Bureau support collection of data for several programs, including the Current Population Survey (the source of monthly unemployment statistics) and the Consumer Expenditure Survey (the source of the market baskets for the Consumer Price Index [CPI]). Some BLS data, such as those for the various national longitudinal surveys, are collected by private contractors. Finally, certain BLS data, such as information on work stoppages, come from secondary sources.

BLS's surveys, indexes, and statistics fall into four main categories:

1. consumer expenditures and prices, including the CPI, the producer price index, and U.S. import and export prices indexes;
2. the labor force, including monthly data on employment from households and business establishments, monthly and periodic data on unemployment, time use, job openings and labor turnover, occupational employment and projections of trends, and longitudinal data on the work experience of cohorts of the population;
3. compensation and working conditions, including the employment cost index, workplace injuries and fatalities, employee benefits, and occupational requirements; and
4. productivity.

Bureau of Transportation Statistics

The Bureau of Transportation Statistics (BTS; https://www.bts.gov/) is under the Office of the Assistant Secretary of Transportation for Research and Technology (OST-R) in the U.S. Department of Transportation. OST-R also includes the Intelligent Transportation Systems Joint Program Office; the Office of Research, Development, and Technology; the Transportation Safety Institute; and the Volpe National Transportation Systems Center.

BTS's director is a career senior executive service appointee who reports to the assistant secretary. Prior to 2004, the director was a presidential appointee with a fixed term of 4 years who reported directly to the secretary of the department. BTS has a full-time staff of about 84 and had direct funding of $26 million in fiscal 2016.

BTS was established by the 1991 Intermodal Surface Transportation Efficiency Act and began operations in late 1992. It was moved to the newly created Research and Innovative Technology Administration (RITA) by the Norman Y. Mineta Research and Special Programs Improvement Act of

2004. BTS moved with the rest of RITA to OST-R in 2014. The 2015 Fixing America's Surface Transportation (FAST) Act (P.L. 114-94) authorized the reorganization of BTS and strengthened its ability to produce statistical products free of political influence.

In regard to independence, Section 6017 of the FAST Act specified that the BTS director did not need the approval of the department for data collection or analysis or for the substance of any statistical data product or press release. The act charged the BTS director with a "significant role" in allocation of the BTS budget, hiring, and grant and contract awards, with the exception that the secretary was to direct external support functions, such as coordination of activities involving BTS and other departmental administrations. Finally, the act charged the departmental chief information officer to consult with the BTS director to ensure that information technology decisions protected the confidentiality of BTS statistical information in accordance with the Confidential Information Protection and Statistical Efficiency Act (CIPSEA).

Prior to the establishment of BTS, statistical programs of the Department of Transportation focused exclusively on specific modes of transportation (highways, airlines, railroads, etc.). the exception was the first 10 years of the department's existence (1967–1977), when the Office of the Secretary funded intermodal surveys on commodity flows and long-distance personal transportation. BTS is charged to produce an annual report on transportation statistics, develop intermodal data on commodity and passenger flows, administer the National Transportation Library, and carry out other functions to ensure that the department, the states, and other federal agencies have available comprehensive information on the nation's transportation systems. BTS also operates the Office of Airline Information, which was transferred to it from the now-defunct Civil Aeronautics Board in 1995. The 2015 FAST Act added a new Port Performance Freight Statistics Program to BTS's portfolio. BTS contracts with the Census Bureau for major surveys.

Census Bureau

The Census Bureau (see http://www.census.gov/) is part of the Economics and Statistics Administration in the U.S. Department of Commerce (as is BEA), headed by the Under Secretary for Economic Affairs. It conducts population and economic censuses and a wide array of surveys.

The first censuses were conducted by U.S. marshals under the authority of the Secretary of State. Beginning in 1850, a separate census office was established each decade to supervise the census. In 1902 a permanent Census Bureau was established; it was made part of the new Department of Commerce and Labor in 1903, and it moved to the newly created Department

of Commerce in 1913. Title 13 of the U.S. Code includes the major legal provisions related to the Census Bureau, including strict provisions for protecting the confidentiality of population and business information.

The director of the Census Bureau is appointed by the President with Senate confirmation for a fixed 5-year term that can be renewed once (to begin in years ending in 2 and 7).[5] The bureau has about 6,240 full-time staff and had direct funding in fiscal 2016 of $1,368.4 million, which included $829.8 million for the upcoming 2020 census.

The major periodic activity of the Census Bureau is the decennial population census, which in 2010 consisted of basic questions on age, sex, race, Hispanic origin, relationship to household head, and housing tenure (own, rent). The Census Bureau, under the decennial program, also conducts the continuous American Community Survey, which includes questions previously part of a long-form sample in the decennial census. Other population data products of the Census Bureau include annual population and housing estimates (developed in cooperation with state and local governments), estimates of poverty, median income, and health insurance coverage using statistical models for small areas, and geographic products based on the Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) system.

The Census Bureau also has a large portfolio of censuses and surveys about businesses, nonprofit organizations, and federal, state, and local governments. In addition, it carries out reimbursable work for other agencies, primarily for the conduct of surveys, which include the Current Population Survey (for BLS), the American Housing Survey (for the Department of Housing and Urban Development), the Consumer Expenditure Survey (for BLS), the National Crime Victimization Survey (for BJS), and the National Health Interview Survey (for NCHS). The Census Bureau estimated it would conduct about $337 million of reimbursable work in fiscal 2017 (U.S. Office of Management and Budget, 2017:Table 2).

Economic Research Service

The Economic Research Service (ERS; http://www.ers.usda.gov/), along with the National Agricultural Statistics Service and two other agencies in the U.S. Department of Agriculture (USDA), reports to the Under Secretary for Research, Education, and Economics. The administrator of ERS is a career senior executive service appointee; the agency has a full-time staff of about 345 and had direct funding of $85.4 million in fiscal 2016.

The origins of ERS trace back to 1905, when USDA established the Office of Farm Management, which was renamed the Office of Farm Management

[5]The fixed term was signed into law in August 2012; previously, the director served at the pleasure of the President.

and Farm Economics in 1919. The office's research areas included farm organization, cost of production, farm labor, farm finance, land economics, agricultural history, and rural life studies. Several reorganizations took place, and in 1961, USDA created ERS with responsibility for conducting economic research and policy analysis that informs program and policy decisions throughout USDA. The agency's mission is to anticipate food, agricultural, agri-environmental, and rural development issues that are on the horizon and conduct peer-reviewed economic research so that research findings are available when issues require decisions by policy makers. As a statistical agency, ERS does not make recommendations: it designs its research to show the consequences of alternative policy or programmatic choices.

ERS is also the primary source of statistical indicators on food and agriculture, such as those that gauge the health of the farm sector (including farm income estimates and projections), assess the current and expected performance of the agricultural sector (including trade and productivity), measure food insecurity in the United States and abroad, and measure dimensions of food availability and access. ERS jointly funds two primary data collection efforts: (1) the Agricultural Resources Management Survey (ARMS) on farm household and business income and crop practices, also funded by the National Agricultural Statistics Service; and (2) the National Household Food Acquisition and Purchase Survey, which focuses on American households' food purchase and acquisitions behavior, also funded by USDA's Food and Nutrition Service.

Energy Information Administration

The Energy Information Administration (EIA; http://www.eia.gov/) is an agency of the U.S. Department of Energy (DOE); its administrator is a presidential appointee with Senate confirmation. EIA has a full-time staff of about 340 and had fiscal 2016 direct funding of $122 million.

EIA was created by Congress in 1977 as part of the newly established Department of Energy. Its mission is to provide policy-independent energy data, forecasts, and analyses in order to promote sound policy making, efficient markets, and public understanding regarding energy and its interaction with the economy and the environment. To assure EIA's independence, the Department of Energy Organization Act specifies that EIA's products are not subject to clearance by executive branch officials: in particular, the administrator does not need to obtain the approval of any other DOE official for data collection and analysis, and he or she does not need to "obtain the approval of any other officer or employee of the United States" before publishing energy data and analysis reports (42 USC 7135(d)).

Many EIA data products, such as weekly, monthly, and annual data on petroleum and natural gas supply, deal with specific industries; others contain data on all fuel types. EIA's mandatory energy supply surveys are conducted by private contractors, who survey energy producers, users, and transporters, and certain other businesses. Data on energy consumption are collected for households, commercial buildings, manufacturing, and transportation. Analyses prepared by EIA staff cover energy economics, technology, production, prices, distribution, storage, consumption, and environmental effects.

EIA forecasts cover all energy types and include supply, consumption, prices, and other factors. Short-term forecasts cover 1–2 years; 20-year projections are also developed and often serve as the baseline for independent analyses of policy proposals that are prepared by EIA at the request of Congress or the administration. More than three-quarters of EIA's resources are used for energy data collection and dissemination; the rest is used to support forward-looking forecasts, projections, and analyses.

National Agricultural Statistics Service

The National Agricultural Statistics Service (NASS; see https://www.nass.usda.gov/) is under the Under Secretary for Research, Education, and Economics in USDA (as is ERS). The administrator of NASS is a career senior executive service appointee. NASS has about 1,038 full-time staff and had direct funding of $168.4 million in fiscal 2016.

The foundation of NASS began with the establishment of USDA in 1862. Agricultural supply information was one of the purposes of the new department. The first official report on the condition of crops was issued in July 1863. NASS's mission of providing timely, accurate, and useful statistics continues today through its agriculture estimates and census of agriculture programs. In its agricultural estimates program, NASS provides the USDA forecasts and estimates for numerous commodities. The census of agriculture is conducted every 5 years and provides comprehensive information about the nation's agriculture down to the county level, which provides a foundation for farm policy among its many uses.

Slightly more than one-third of the agency's staff is located at its Washington, DC, headquarters; the rest of the staff is located at the National Operations Center near St. Louis, Missouri, and in 12 regional offices, each of which is responsible for the statistical work in several states. All field and telephone interviewing staff are obtained through contracting with the National Association of State Departments of Agriculture (NASDA). NASS researchers also collaborate with researchers, largely from land-grant universities and the National Institute of Statistical Sciences, to improve

statistical methodologies and practices of both the agricultural estimates and the census of agriculture programs.

NASS provides data services for many agencies inside and outside USDA. It collaborates with state departments of agriculture and land-grant universities to meet state, local, and national needs for agricultural statistics. Through cooperative agreements going back as far as 1917 and memoranda of understanding, NASS provides data collection and statistical services to other federal agencies, and it provides statistics to the public through trust fund agreements with private producer organizations when federal funding is inadequate.

NASS works with its regional field offices to carry out hundreds of surveys every year and prepares reports covering virtually every aspect of U.S. agriculture. Examples include production and supplies of food and fiber, prices paid and received by farmers, farm labor and wages, farm finances, chemical use, and changes in the demographic characteristics of U.S. producers.

National Center for Education Statistics

The National Center for Education Statistics (NCES; http://nces.ed.gov/) is part of the Institute of Education Sciences (IES) in the U.S. Department of Education; IES also includes three research and evaluation centers. The NCES commissioner is a presidential appointee for a fixed term of 6 years (not requiring Senate confirmation—a change as of August 2012). It has a full-time staff of about 93 people, and its direct funding was $332.6 million in fiscal 2016.

NCES's origins date back to 1867 when Congress established a Department of Education and gave it a primary mission of "collecting such statistics and facts as shall show the condition and progress of education in the several States and Territories, and of diffusing such information respecting the organization and management of schools and school systems and methods of teaching" (P.L. 39-73, 14 Stat. 434). The legislation also charged the department's commissioner to issue an annual report. However, only 2 years later the department was abolished, and an Office of Education was established in the U.S. Department of the Interior, where it remained through 1939. The Office of Education was part of the newly created Federal Security Agency from 1939 to 1953, when it was made part of the newly created U.S. Department of Health, Education, and Welfare. A separate Department of Education was reestablished in 1980.

A major function of the Office of Education throughout its history was the collection and publication of education statistics. NCES was established in 1965 as a staff office reporting to the Commissioner of Education. NCES received statutory authority in 1974; in 1980 it was made part of the

Office of Educational Research and Improvement, which in 2002 became the IES. Supporting the independence of NCES, the Education Sciences Reform Act of 2002, which created IES, stipulated that "each Commissioner [head of one of IES's constituent centers], except the Commissioner for Education Statistics, shall carry out such Commissioner's duties . . . under the supervision and subject to the approval of the Director" of IES (20 USC 9517(d)).

NCES has an extensive survey program, including longitudinal surveys that follow the educational experience of cohorts of the U.S. population from early childhood through adulthood, periodic surveys of adult literacy, and international studies of educational achievement. It also collects the "Common Core of Data" from administrative records of state and local K–12 educational agencies, and it collects data for the Integrated Postsecondary Education Data System. It regularly assesses the educational knowledge and achievement of primary and secondary school students in the National Assessment of Educational Progress. It also administers the Statewide Longitudinal Data Systems program, which provides grants to the states to develop longitudinal databases of student records for analyzing student performance and for identifying methods to improve achievement.

NCES contracts for a substantial portion of its work, including not only data collection, but also data analysis and preparation of reports. For fiscal 2017 (U.S. Office of Management and Budget, 2017:Table 2), 92 percent of its planned budget authority was to be spent on data and analysis from state agencies, the Census Bureau, and private contractors.

National Center for Health Statistics

The National Center for Health Statistics (NCHS; https://www.cdc.gov/nchs/index.htm) is part of the Centers for Disease Control and Prevention (CDC) in the U.S. Department of Health and Human Services (DHHS). The NCHS director is a career senior executive service appointee. It has a full-time staff of about 508 people and had direct funding of $160.4 million in fiscal 2016.

NCHS's roots lie in two formerly separate historical strands for the provision of national health statistics. The first strand is vital statistics on births, deaths, and other life events, which traces back to 1902, when Congress gave the newly created permanent Census Bureau the authority to establish registration areas to produce nationally comparable vital statistics by working with state agencies. This function was transferred in 1946 to the Federal Security Administration, which was folded into the new U.S. Department of Health, Education, and Welfare in 1953, which subsequently was split into two federal departments. The second strand is general

statistics on the nation's health, which were authorized in the 1956 National Health Survey Act.

NCHS was created in 1960 as the merger of the National Office of Vital Statistics and the National Health Survey Division; it was relocated every few years in DHHS until its last relocation in 1987, when it was made part of CDC. In 2005 it became one of three centers reporting to the newly created Coordinating Center for Health Information and Service in CDC. In 2013, further administrative reorganization placed NCHS within the new CDC Office of Public Health Scientific Services (78 *Federal Register* 70049, November 22, 2013).

NCHS has four major programs:

1. The National Health Interview Survey, in continuous operation since 1956, collects a wide range of information on self-reported health status and conditions and use of health care services by the population.
2. Several surveys collect information from health care providers, including nursing homes, hospitals, and outpatient facilities.
3. The National Health and Nutrition Examination Survey ascertains self-reported information on health and dietary intake and also, by use of mobile examining units, obtains extensive information from physical examinations and laboratory tests.
4. The nation's basic vital statistics are collected and maintained.

For fiscal 2017 (U.S. Office of Management and Budget, 2017:Table 2), 85 percent of NCHS's estimated budget authority would be used to purchase data collection and reporting services from state and local governments, the Census Bureau, and private contractors.

National Center for Science and Engineering Statistics

The National Center for Science and Engineering Statistics (NCSES; https://www.nsf.gov/statistics/) is part of the Social, Behavioral, and Economic Sciences Directorate of the National Science Foundation (NSF). Its director is a career senior executive service appointee. It has a full-time staff of about 50 people and had direct funding in fiscal 2016 of $58.2 million.

NCSES was formerly the Division of Science Resources Statistics and before that the Division of Science Resources Studies. It became NCSES with passage of the America COMPETES Reauthorization Act of 2010 (Section 505; 42 USC 1862), with an expanded mandate to serve as a "central Federal clearinghouse for the collection, interpretation, analysis, and dissemination of objective data on science, engineering, technology, and research and development."

NCSES's history began in 1950, when the newly created NSF was charged to maintain a register of scientific and technical personnel so that the nation would be able to mobilize the scientific and technical work force in the event of a major war. Although no longer required to maintain a complete register, NSF has continued (by the terms of its founding act, as amended) to have responsibility "to provide a central clearinghouse for the collection, interpretation, and analysis of data on scientific and engineering resources and to provide a source of information for policy formulation by other agencies of the Federal Government" (42 USC 1862). NSF also has a congressional mandate from 1980 to provide information on women and minorities in science and engineering.

The NSF mandates provide the basis for statistical programs in NCSES. The center is called on to support the collection of statistical data on research and development trends, the science and engineering workforce, U.S. competitiveness, and the condition and progress of the nation's science, technology, engineering, and mathematics (STEM) education; to support research using the data it collects and on methodologies in areas related to its work; and to support the education and training of researchers in the use of its own and other large-scale, nationally representative data sets. NCSES designs, supports, and directs a coordinated collection of periodic national surveys and performs a variety of other data collections and research, providing policymakers, researchers, and other decision makers with high quality data and analysis on research and development, innovation, the education of scientists and engineers, and the science and engineering workforce. To support its programs, 66 percent of NCSES's estimated budget authority in fiscal 2017 (U.S. Office of Management and Budget, 2017:Table 2) would be used to purchase data collection and other services from the Census Bureau and private contractors. NCSES also serves as staff to the National Science Board in producing the biennial congressionally mandated *Science and Engineering Indicators Report*, which uses data from all NCSES surveys.

Office of Research, Evaluation, and Statistics

The Office of Research, Evaluation, and Statistics (ORES; https://www.ssa.gov/policy/index.html) is located in the Social Security Administration (SSA). ORES reports to the SSA Deputy Commissioner for Retirement and Disability Policy. ORES is headed by an associate commissioner, who is a career senior executive service appointee; it has a full-time staff of about 66 people and had direct funding of $23.4 million in fiscal 2016.

SSA began as the Social Security Board in 1935; it became part of the Federal Security Agency in 1939, part of the Department of Health, Education, and Welfare in 1953, and part of DHHS in 1980; it regained

independent agency status in 1995. From the outset, SSA has had a research, statistics, and evaluation function.

ORES produces numerous recurring statistical publications about the Social Security and Supplemental Security Income (SSI) programs, such as the *Annual Statistical Supplement*. ORES also produces statistical publications about earnings and employment and other topics related to Social Security, such as the *Income of the Population 55 or Older* and the *Income of the Aged Chartbook*.

ORES conducts and sponsors research and evaluation on the effects of the Social Security and SSI programs and proposed changes in those programs on individuals, the economy, and program solvency. It develops and operates microsimulation models to assess the distributional effects of proposed reforms to the Social Security and SSI programs. ORES also conducts comparative analyses of social insurance systems in other countries. The research generated by ORES often is published in its in-house journal, the *Social Security Bulletin*. In addition, ORES funds two external research networks through cooperative agreements, the Retirement Research Consortium (RRC) and the Disability Research Consortium (DRC).[6] The RRC and DRC promote research on a wide range of topics related to Social Security retirement and disability policy at universities and think tanks.

Finally, ORES performs a significant data infrastructure function in support of policy research. ORES is responsible for working with outside research partners to create restricted use research datasets by linking survey and other external data to Social Security program data. ORES also supports epidemiologists by providing vital status data on subjects of health research.[7]

Statistics of Income Division

The Statistics of Income Division (SOI; see https://www.irs.gov/uac/tax-stats) is housed in the Office of Research, Analysis, and Statistics of the Internal Revenue Service (IRS) in the U.S. Department of the Treasury. The director is a career senior executive service appointee. SOI has a full-time staff of approximately 117 employees and had direct funding of $36.9 million in fiscal 2016.

SOI's history traces back to the enactment of authority to levy individual income taxes in the 16th amendment to the U.S. Constitution, which was ratified in 1913. Section 21 of the Revenue Act of 1916 mandated the annual "publication of statistics reasonably available with respect to the operation

[6]See https://www.ssa.gov/policy/rrc/ and https://www.ssa.gov/policy/drc/index.html [April 2017].

[7]See https://www.ssa.gov/policy/about/epidemiology.html [April 2017].

of the Income tax law" (39 Stat. 776); identical language is found in the current Internal Revenue Code (see 26 USC 6108).

SOI provides income, financial, and tax information data products to the user community that are based largely on individual and corporate tax returns and on returns filed by most tax-exempt organizations. It also provides periodic data derived from other returns and schedules, such as estate and gift taxes, foreign income and taxes, and gains and losses from sales of capital assets.

On written request, SOI tax return data are available to staff in the Department of the Treasury and the Congressional Joint Committee on Taxation for policy analysis and revenue estimation. SOI data are also available to the Congressional Budget Office for modeling Social Security and Medicare programs, but not for any other purpose. Selected tax return data are also available, under strict confidentiality protection provisions, for use by the Census Bureau, the Bureau of Economic Analysis, and the National Agricultural Statistics Service: the purposes of this access are for structuring censuses and national economic accounts and conducting related statistical activities authorized by law. (See discussion of the Confidential Information Protection and Statistical Efficiency Act of 2002 in Appendix A.)

OTHER RECOGNIZED STATISTICAL UNITS

Implementation guidance issued in 2007 for the 2002 Confidential Information Protection and Statistical Efficiency Act (CIPSEA) (see Appendix A) recognized 12 principal statistical agencies—the 13 agencies identified as principal statistical agencies that serve on the ICSP, with the exception of the Social Security Administration Office of Research, Evaluation, and Statistics.[8] It provided a mechanism by which other agencies or units can be recognized as statistical agencies or units for purposes of CIPSEA (U.S. Office of Management and Budget, 2007:33368):[9]

> Other agencies or units that wish to be recognized as statistical agencies or units for purposes of CIPSEA must send a request to the Chief Statistician at OMB. The request must come from the head of the agency or unit and have the concurrence of the larger organization within which

[8]CIPSEA adopted the designation of 12 agencies from the *Order Providing for the Confidentiality of Statistical Information*, issued by OMB in 1997 (see Appendix A). Why the 1997 order named 12 instead of 13 agencies is not clear.

[9]In addition to the requirements for designation as a statistical unit spelled out in the quoted paragraph, *Statistical Policy Directive No. 1* (U.S. Office of Management and Budget, 2014) applies to recognized statistical units in addition to principal statistical agencies. This directive requires a statistical agency or unit's department to recognize the agency or unit's independence (see Appendix A).

> the agency or unit resides. This request should include a statement of the organizational definition of the agency or unit, its mission, statistical activities, and any nonstatistical activities, and demonstrate that its activities are predominantly statistical. Statistical activities include the collection, compilation, processing, or analysis of information for the purpose of describing the characteristics of groups or making estimates concerning the whole or relevant groups, or components within, the economy, society, or the natural environment. Statistical activities also include the development of methods or resources that support these activities, such as measurement methods, models, statistical classifications, or sampling frames.

CIPSEA-recognized statistical agencies and units may bring data collection contractors and researchers under the CIPSEA umbrella by swearing them in as special agents. Such agents may have access to confidential data of the agency or unit for specified purposes (e.g., to conduct a survey for the agency or analyze data from a survey in a secure environment). To date, CIPSEA has recognized four statistical units: the Office for Research, Evaluation, and Statistics of SSA; the Center for Behavioral Health Statistics and Quality of the Substance Abuse and Mental Health Services Administration in the U.S. Department of Health and Human Services; the Microeconomic Surveys Section of the Federal Reserve Board; and the National Animal Health Monitoring System Program Unit of the Animal and Plant Health Inspection Service in the U.S. Department of Agriculture. The last three units are described below.

Center for Behavioral Health Statistics and Quality

The Center for Behavioral Health Statistics and Quality (CBHSQ; https://www.samhsa.gov/about-us/who-we-are/offices-centers/cbhsq) is the lead federal agency for behavioral health statistics. It is housed in the Substance Abuse and Mental Health Services Administration (SAMHSA) in DHHS. CBHSQ provides national leadership in behavioral health statistics and epidemiology; promotes basic and applied research in behavioral health data systems and statistical methodology; designs and carries out special data collection and analytic projects to examine issues for SAMHSA and other federal agencies; participates with other federal agencies in developing national health statistics policy; and consults and advises the SAMHSA administrator and the DHHS secretary on statistical matters.

CBHSQ conducts the continuing National Survey on Drug Use and Health (NSDUH), which is the nation's primary data system for collecting information on the incidence and prevalence of substance abuse and adverse health consequences associated with drug abuse from the civilian, noninstitutionalized population of the United States for people aged 12

and older. NSDUH (formerly called the National Household Survey on Drug Abuse) was fielded periodically from 1972 to 1990 and then annually beginning in 1991.

Other CBHSQ statistical programs include the Behavioral Health Services Information System and its associated surveys, which are the primary data sources for information on the nation's substance abuse treatment system and outcomes; the Drug Abuse Warning Network (DAWN), a public health surveillance system that monitors drug-related visits to hospital emergency departments, as well as drug-related deaths investigated by medical examiners and coroners; and other programs (see https://www.samhsa.gov/data/node/20).

Microeconomic Surveys Section

The Microeconomic Surveys Section of the Division of Research and Statistics of the Federal Reserve Board (https://www.federalreserve.gov/econresdata/rsmecs-staff.htm) conducts research in a variety of areas, including consumer finances, financial markets, general applied microeconomics, survey methodology, and other statistical methodology. The section has responsibilities for a number of the surveys conducted by the Board.

The section's principal survey is the triennial Survey of Consumer Finances (SCF), first conducted in 1983 and most recently conducted in 2013 (see https://www.federalreserve.gov/econres/scfindex.htm). The SCF ascertains detailed information on families' balance sheets, pensions, income, and demographic characteristics from an area probability sample of households supplemented by a list sample from federal income tax records of high-income families that hold disproportionately large amounts of assets. This design enables the SCF to provide information on the distribution of family income and assets that is unmatched by any other U.S. survey. Data from the SCF are widely used, both in analyses by other agencies, most notably the Federal Reserve Board, and in scholarly work at major economic research centers. The survey has contained a panel element over two periods: respondents to the 1983 survey were reinterviewed in 1986 and 1989, and respondents to the 2007 survey were reinterviewed in 2009.

National Animal Health Monitoring System Program Unit

The Animal Plant Health and Inspection Service of the U.S. Department of Agriculture established the National Animal Health Monitoring System Program Unit (NAHMS; https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/monitoring-and-surveillance/nahms/about) in 1983 to collect, analyze, and disseminate data on animal health, management,

and productivity across the United States. The NAHMS staff conduct national studies on the health and health management of U.S. domestic livestock and poultry populations. Studies are designed to meet the information needs of the industries associated with these commodities. Each animal group is studied at regular intervals, providing up-to-date and trend information needed to monitor animal health, support trade decisions, assess research and product development needs, answer questions for consumers, and set policy.

NAHMS data on producer biosecurity practices and animal movements (such as transport distances and frequencies) for multiple animal industries have been essential in establishing parameters for the North American Animal Disease Spread Model, which simulates the spread and control of highly contagious diseases in susceptible animals. Trend data collected between 1990 and 2006 by four NAHMS swine studies documented the decline and virtual elimination of trichinae in commercial market swine and showed a corresponding decrease in the use of production practices that contribute to trichinae.

NAHMS works with NASS to randomly select operations by size that represent the targeted national livestock and poultry populations. NASS personnel contact sampled producers to complete the questionnaire. If the study includes biological sampling, animal health professionals or the producers themselves collect and ship the samples to designated veterinary diagnostic laboratories. Producers then receive results of biological sample testing.

OTHER STATISTICAL PROGRAMS

This section briefly describes eight statistical programs that are conducted or sponsored by agencies of the federal government other than the principal statistical agencies and recognized statistical units. The programs were selected purposively to illustrate the breadth and depth of the federal government's statistical portfolio. They are in alphabetical order.

Health and Retirement Study

The Health and Retirement Study (HRS; see http://hrsonline.isr.umich.edu/) is a longitudinal panel survey with more than 20,000 respondents representing people ages 51 and older in the United States. It provides in-depth information on middle- and older-aged people's transitions from the workforce to retirement, savings behavior and pension plans, physical and cognitive health, disability, family structure, health care expenditures, and many other aspects of financial, social, physical, and mental well-being.

The HRS began in 1992 and currently introduces a new cohort of people ages 51–56 every 6 years. People in the sample are interviewed in-person or by telephone every 2 years. The HRS is conducted by the University of Michigan with support from the National Institute on Aging and the Social Security Administration.

The HRS has provided data for a wide range of path-breaking research studies, has made innovations in data collection methods, and has inspired similar efforts in many countries around the world. Similar panel surveys in other countries include the China Health and Retirement Longitudinal Study; the English Longitudinal Study of Ageing; the Japanese Study of Aging and Retirement; the Korean Longitudinal Study of Aging; the Longitudinal Aging Study in India; the Mexican Health and Aging Study; the New Zealand Health, Work and Retirement Survey; and the Survey of Health, Ageing, and Retirement in Europe.

Medical Expenditure Panel Survey

The Medical Expenditure Panel Survey (MEPS; http://meps.ahrq.gov/mepsweb/) is a statistical program of the Agency for Healthcare Research and Quality (AHRQ) in the U.S. Department of Health and Human Services. MEPS is the core health care expenditure survey in the United States, with a primary analytical focus directed to the topics of health care access, cost, and coverage. MEPS was designed to provide data for health care policy analysis and research. It was first conducted in 1977 and again in 1987 under different names and became a continuous survey in 1996.

MEPS consists of three interrelated parts: the household component, the medical provider component, and the insurance component. The household survey collects information from household members and their health care providers and employers in order to construct a complete picture of medical care use, expenditures, and health insurance coverage and reimbursements. Households are in a MEPS panel for five rounds of interviewing that cover 2 years so that patterns of medical care and expenditures can be observed over time; a new household panel begins every year. Data for the MEPS household and medical provider surveys are collected by private contractors; the household survey sample of about 14,000 households per year is drawn from the NCHS National Health Interview Survey. The MEPS insurance component collects data each year from a sample of about 30,000 private- and public-sector employers on the health insurance plans they offer their employees. The collected data include the number and types of private insurance plans offered (if any), premiums, contributions by employers and employees, eligibility requirements, benefits associated with these plans, and employer characteristics. Data for this component of MEPS are collected by the Census Bureau.

National Agricultural Workers Survey

The National Agricultural Workers Survey (NAWS; http://www.doleta.gov/naws/) is an activity of the U.S. Department of Labor's Employment and Training Administration. It provides data on wage and migration history, type of crops worked, unemployment benefits, housing, health care, use of public programs, and other characteristics of the U.S. crop labor force. The information, which is used by numerous federal agencies for occupational injury and health surveillance, estimating the need for services for workers, allocating program dollars to areas of greatest need, and program design and evaluation, is obtained directly from farm workers through personal interviews.

Since 1988, when the survey began, nearly 53,000 workers have been interviewed. The survey samples crop workers in three cycles each year to reflect the seasonality of agricultural production and employment. Workers are located at their farm job sites. During the initial contact, arrangements are made to interview the respondent at home or at another location convenient to the respondent. Depending on the information needs and resources of the various federal agencies that use NAWS data, between 1,500 and 4,000 workers are interviewed each year.

National Automotive Sampling System

The National Automotive Sampling System (https://www.nhtsa.gov/research-data/national-automotive-sampling-system-nass) is an administrative-records–based data collection system of the National Center for Statistics and Analysis (NCSA) of the National Highway Traffic Safety Administration in the U.S. Department of Transportation. The system was created in 1979 as part of a nationwide effort to reduce motor vehicle crashes, injuries, and deaths on U.S. highways. It samples accident reports of police agencies in randomly selected areas of the country.

The National Automotive Sampling System has two components, one on crashworthiness and one on general estimates. For the crashworthiness component, NCSA field researchers collect detailed information from police accident reports for selected crashes on a wide range of factors, including exterior and interior vehicle damage, occupant injury, and environmental conditions at the time of the crash. For the general estimates component, which covers a larger sample of crashes, only basic information is recorded from the police accident reports. The National Automotive Sampling System infrastructure is also used for special studies and surveys, such as the National Motor Vehicle Crash Causation Survey, conducted in 2005–2007, which sampled police accident reports in real time and obtained on-scene information in addition to the information reported by the police.

National Center for Veterans Analysis and Statistics

The National Center for Veterans Analysis and Statistics (NCVAS; https://www1.va.gov/vetdata/) in the U.S. Department of Veterans Affairs (VA) develops descriptive, diagnostic and predictive analytics on a broad range of topics about veterans and VA programs. NCVAS collaborates with other federal agencies to survey and analyze the veteran population and represents the VA in interagency statistical, data science, and data management forums.

NCVAS sponsors the National Survey of Veterans (NSV), of which six rounds have been conducted to date, the latest in 2010. The 2010 NSV included veterans, active duty service members, demobilized National Guard and Reserve members, family members, and surviving spouses. This was the first NSV iteration to include groups other than veterans.

National Resources Inventory

The National Resources Inventory (NRI; https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/technical/nra/nri/) is a statistical program of the Natural Resources Conservation Service (NRCS) in USDA. The current NRI is a longitudinal survey of soil, water, and related environmental resources designed to assess conditions and trends on nonfederal U.S. land parcels. NRCS has conducted the NRI in cooperation with the Iowa State University Center for Survey Statistics and Methodology since 1977.

The NRI was conducted on a 5-year cycle from 1982 to 1997; beginning in 2000 it is now conducted annually (with major data releases occurring at 5-year intervals). Before 2000, NRI data were collected every 5 years for 800,000 sample sites; annual NRI data collection covers slightly less than 200,000 sample sites. Year-by-year data on conditions for the same sites enable analysis of the effects of resource conservation programs and other applications.

Office of Environmental Information

The Office of Environmental Information (OEI; https://www.epa.gov/aboutepa/about-office-environmental-information-oei) in the Environmental Protection Agency (EPA) is headed by the chief information officer, who reports to the EPA administrator. OEI manages the life cycle of information to support EPA's mission of protecting human health and the environment. It works to ensure the quality of EPA's information and the efficiency and reliability of EPA's technology, data collection and exchange efforts, and access services. It provides technology services and manages EPA's IT investments. OEI operates the Environmental Dataset Gateway (EDG),

which is a web-based metadata (data about data) portal that supports the discovery of and access to EPA's environmental dataset resources. The EDG contains metadata records contributed by EPA offices and links to geospatial and non-geospatial resources (e.g., data, services, or applications) described by those metadata records. Unrestricted information that is contributed to the EDG is shared with interagency data sharing portals, including data.gov and geo.data.gov. Examples of datasets in the EDG are state listings of impaired water sources and fish consumption advisories.

Panel Study of Income Dynamics

The Panel Study of Income Dynamics (PSID; http://simba.isr.umich.edu/data/data.aspx) is a longitudinal survey that has followed several thousand families since 1968. It is conducted by the Survey Research Center of the Institute for Social Research of the University of Michigan with funding from a consortium of federal agencies.

The PSID emphasizes the dynamic aspects of economic and demographic behavior, but its content is broad, including sociological and psychological measures. From 1968 to 1996, the PSID interviewed individuals in the original sample of about 4,800 families every year, whether or not they were living in the same dwelling or with the same people. In 1997 interviewing was changed to every other year, the original sample was reduced, and a sample of Hispanic families that had been added in 1990 was replaced by a sample of post-1968 immigrant families of all ethnic groups and their adult children. The current sample of families, including those formed by children leaving their parental homes, is about 8,700.

Since 1968, more than 3,000 journal articles, books and book chapters, government reports, working papers, and dissertations have been based on the PSID. The PSID was founded to study poverty and the effects of programs to combat poverty. As the survey has added content and extended its period of observation, the data have also contributed importantly to studies of intergenerational patterns of work, welfare receipt, and other behaviors; international comparisons with panel data from other countries; neighborhood effects on family well-being (using data files augmented with census-based characteristics of sample members' communities); and long-term trends in marital and fertility histories and living arrangements.

REFERENCES FOR APPENDIX B

All URL addresses valid as of April 2017. *Please note that, as of February 2017, U.S. Office of Management and Budget documents previously on the main whitehouse.gov site were relocated to an archived site. Wherever possible, a citation is provided to a stable site such as federalregister.gov.*

Norwood, J. L. (1995). *Organizing to Count: Change in the Federal Statistical System*. Washington, DC: The Urban Institute Press.

U.S. Office of Management and Budget (2001). *Statistical Programs of the United States Government: Fiscal Year 2002*. Statistical and Science Policy Office, Office of Information and Regulatory Affairs. Washington, DC: U.S. Government Printing Office. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/inforeg/02statprog.pdf.

U.S. Office of Management and Budget (2007). Implementation guidance for Title V of the E-Government Act, Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA). 72 *Federal Register* 33362 (June 15, 2007). Available: https://www.federalregister.gov/d/E7-11542.

U.S. Office of Management and Budget (2014). Statistical Policy Directive No. 1: Fundamental responsibilities of federal statistical agencies and recognized statistical units. 79 *Federal Register* 71609 (December 2, 2014). Available: https://www.federalregister.gov/d/2014-28326.

U.S. Office of Management and Budget (2016). *Analytical Perspectives, Budget of the U.S. Government, Fiscal Year 2017*. Washington, DC: U.S. Government Printing Office. Available: https://www.govinfo.gov/content/pkg/BUDGET-2017-PER/pdf/BUDGET-2017-PER.pdf.

U.S. Office of Management and Budget (2017). *Statistical Programs of the United States Government: Fiscal Year 2017*. Statistical and Science Policy Office, Office of Information and Regulatory Affairs. Washington, DC: U.S. Government Printing Office. Available: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/information_and_regulatory_affairs/statistical-programs-2017.pdf.

COMMITTEE ON NATIONAL STATISTICS

The Committee on National Statistics was established in 1972 at the National Academies of Sciences, Engineering, and Medicine 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, a National Agricultural Statistics Service cooperative agreement, and several individual contracts.