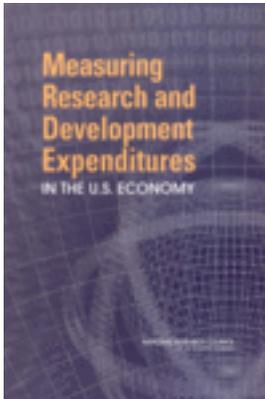


Free Executive Summary



Measuring Research and Development Expenditures in the U.S. Economy

Lawrence D. Brown, Thomas J. Plewes, and Marisa A. Gerstein, Editors, Panel on Research and Development Statistics at the National Science Foundation, National Research Council

ISBN: 978-0-309-09320-0, 212 pages, 6 x 9, paperback (2004)

This free executive summary is provided by the National Academies as part of our mission to educate the world on issues of science, engineering, and health. If you are interested in reading the full book, please visit us online at <http://www.nap.edu/catalog/11111.html>. You may browse and search the full, authoritative version for free; you may also purchase a print or electronic version of the book. If you have questions or just want more information about the books published by the National Academies Press, please contact our customer service department toll-free at 888-624-8373.

This report examines the portfolio of research and development (R&D) expenditure surveys at the National Science Foundation (NSF), identifying gaps and weaknesses and areas of missing coverage. The report takes an in-depth look at the definition of R&D, the needs and potential uses of NSF's R&D data by a variety of users, the goals of an integrated system of surveys and other data collection activities, and the quality of the data collected in the existing Science Resources Statistics surveys.

This executive summary plus thousands more available at www.nap.edu.

Copyright © National Academy of Sciences. All rights reserved. Unless otherwise indicated, all materials in this PDF file are copyrighted by the National Academy of Sciences. Distribution or copying is strictly prohibited without permission of the National Academies Press <http://www.nap.edu/permissions/>. Permission is granted for this material to be posted on a secure password-protected Web site. The content may not be posted on a public Web site.

Executive Summary

The Panel to Review Research and Development Statistics at the National Science Foundation (NSF) was asked to look at the definition of research and development (R&D), the needs and potential uses of NSF's R&D data by a variety of users, the goals of an integrated system of surveys and other data collection activities, and the quality of the data collected in the existing Science Resources Statistics (SRS) surveys. The panel has examined the portfolio of R&D expenditure surveys, identifying gaps and weaknesses, and areas of missing coverage.

The R&D expenditure surveys have owed their growth to a heightened interest in science and technology policy since World War II, as well as the growing federal involvement in R&D policy. Over the years, these data have become the accepted measures of the amounts of R&D spending, and of public and private investment in areas of science and engineering. These data have been called on to serve other purposes as well. They have become a proxy indicator of the direction of technological change. They are consulted to portray the locus of emphasis among the public, private, non-profit, and college and university sectors. Most importantly, they are used by federal agencies, Congress, and the public to frame the national debate over the investment strategy for R&D.

SURVEY CHALLENGES

The NSF research and development expenditure data are often ill-suited for the purposes to which they have been employed. They attempt to quantify three traditional pieces of the R&D enterprise—basic research,

applied research, and development—when much of the engine of innovation stems from the intersection of these components, or in the details of each. Public policy attention to early-stage technology development, the Advanced Technology Program, and process innovation requires data beyond these basic components of R&D. Similarly, the data are sometimes used to measure the *output* of R&D, when, in reality, in measuring expenditures, they reflect only one of the *inputs* to innovation and economic growth. It would be desirable to devise, test and, if possible, implement survey tools that more directly measure the economic output of R&D in terms of short-term and long-term innovation. Finally, the structure of the data collection is tied to models of R&D performance that are increasingly unrepresentative of the whole of the R&D enterprise. The growth of the service sector, the growing recognition of the role of small firms in R&D, the shift in funding from manufacturing R&D to health-related R&D, changes in geographic location, and the globalization of R&D have all served to challenge the current system for depicting the amount and character of R&D in today's economy. New forms of conducting R&D in collaborative environments, using joint ventures or outsourcing arrangements, working through alliances, and outsourcing R&D to foreign affiliates are just a few of the emerging ways of conducting research and development that are not well measured by the traditional R&D surveys.

At the same time that the foundation of R&D statistics is coming under increasing pressure, the league of uses and users continues to expand. The National Science Board continues to make sophisticated use of these data in producing the comprehensive volume, *Science and Engineering Indicators*, every 2 years, which places additional stress on the data in terms of quality and timeliness. The data are used by the administration, particularly the U.S. Office of Management and Budget (OMB) and the Office of Science and Technology Policy, to paint a complete picture of federal and nonfederal investment in R&D. Congress not only relies on the NSF data but also has directed collection of data necessary for evaluating the need for public investment in R&D. New uses of the data for purposes for which they were not originally intended are springing up. The inclusion of R&D investment in national income and product accounts, as well as in estimates of multi-factor productivity, are two examples of the emerging uses that refocus attention on these data sources.

Finally, as the data have come under increasing use, they have come under increasing scrutiny. Some users are deeply troubled by the apparent discrepancy between reports of federal spending on R&D and the amounts that academia and industry report that they have received from the federal government. This large discrepancy casts doubt on the reliability of some of the data sources.

SURVEY REVIEW

Against this backdrop, the panel undertook an in-depth review of five of the recurring statistical collections by the Science Resources Statistics Division of the National Science Foundation:

- The Survey of Industrial Research and Development,
- The Survey of Federal Funds for Research and Development,
- The Survey of Federal Science and Engineering Support to Universities, Colleges, and Nonprofit Institutions,
- The Survey of Research and Development Expenditures at Universities and Colleges, and
- The Survey of Scientific and Engineering Research Facilities.

In addition, the panel considered the provisional Survey of Innovation, which has been conducted twice, in somewhat different forms.

Industry R&D Survey

The panel devoted much of its attention to the critically important survey of industrial R&D. This survey is conducted for NSF by the U.S. Census Bureau. It was last redesigned in 1991, to expand the sample into the service sector of the economy and make other changes. The panel recommends that NSF address the problems associated with this survey first, and lists its recommendations below in order of priority.

The improvements to the industry survey are extensive and expensive and call for a reconsideration of the basis for administering this survey. For several reasons, including the need to increase the professionalism of the staff of the Science Resources Statistics Division, the panel urges SRS to take the lead in the work on the industrial survey. While leaving the exact form of this more active role up to the designs of NSF and the Census Bureau, the panel suggests using the tools of the interagency agreement, the oversight of a high-quality methodological staff, and the input of highly qualified outside experts. This lead role should be undertaken while working collaboratively with the Census Bureau (**Recommendation 8.1**).

The panel strongly recommends that the National Science Foundation and the Census Bureau resume a program of field observation staff visits to a sampling of reporters to examine record-keeping practices and conduct research on how respondents fill out the forms (**Recommendation 3.11**). The first step in this process will be to make contact with respondents. The panel supports the initiative to identify individual respondents in companies as a first and necessary step toward developing an educational interac-

tion with respondents so as to improve response rates and the quality of responses (**Recommendation 8.3**).

Although the survey has been modified and adapted over the past decade, it has largely failed to keep up with the fast-changing environment for the conduct and organization of research in the private business sector, or with advances in data collection and analysis techniques. Results from field observations should inform this redesign, and NSF should also conduct research into record-keeping practices of reporting establishments by industry and size of company to determine if they can report by more specific categories that further elaborate applied research and development, such as the categories utilized by the Department of Defense (DoD) (**Recommendation 3.1**).

NSF and the Census Bureau should test the ability to collect some disaggregated data by the newer, more detailed North American Industrial Classification System (NAICS) codes used in the industry survey today. The record-keeping practice surveys should be used to assess the feasibility and burden of providing this additional detail on industrial reporters. With this information in hand, NSF and its advisory committee should decide whether the collection of reliable R&D line-of-business data is feasible, and, if so, whether for all or a subset of reporters, and at which frequency (**Recommendation 3.5**).

The panel notes that a special emphasis panel of R&D officials in large companies, which had provided advice and spending projections during the 1980s, had been disbanded in 1990 for reasons of funding shortfalls and concern over whether the body was sufficiently representative of industrial R&D. Today, NSF has no standing advisory body to which it can turn for advice on measurement issues in the industry survey. We recommend that NSF again develop a panel of R&D experts, broadly representative of the R&D performing and R&D data-using communities, to serve as a feedback mechanism to provide advice on trends and issues of importance to maintaining the relevance of the R&D data (**Recommendation 3.8**).

Among the issues facing the managers of the industrial R&D survey is the wastefulness of surveying large numbers of establishments to find a relatively rare activity: R&D was reported only for about 3,500 of the 25,000 firms in the sample. The panel recommends the use of supplemental lists of R&D performers in drawing the sample. There are a number of practical problems to be solved in using one or more supplemental lists. Lists may overlap, and duplicates must be handled in some way. The units on the lists may not all be the same—establishments may be mixed in with companies, for example—and some editing will be needed in advance of sampling. However, the payoff in efficiency could be substantial, and the panel thinks that this approach is worth investigating (**Recommendation 3.2**). Low response rates in the survey are a concern of every user, as they

may signal a problem with the quality of the estimates. The panel recommends increased reliance on mandatory reporting between economic censuses, and additional research on the topic of voluntary versus mandatory reporting (**Recommendation 8.5**).

The panel concludes that appropriate assignment of industrial classification to industrial R&D activity requires additional breakdowns of data at the business unit level. We urge NSF and the Census Bureau to evaluate the results of the initial collection of R&D data in the Company Organization Survey to determine the long-term feasibility of collecting these data.

The panel is concerned about the possibility that the editing process, replete with analyst judgment, could introduce unmeasured and undocumented errors into the publicly released data. The panel recommends that the industrial R&D editing system be redesigned so that the current problems of undocumented analyst judgment and other sources of potential error can be better understood and addressed (**Recommendation 3.12**). As NSF turns to modernizing the industrial R&D survey, the panel urges it to sponsor research into the effect of imprinting prior-period data on the industrial R&D survey in conjunction with testing the introduction of web-based data collection (**Recommendation 8.2**).

The panel took note of a recent pioneering effort to improve understanding of the impact of foreign investment in R&D in the United States by linking Census Bureau R&D data to the foreign direct investment data of the Bureau of Economic Analysis. The panel commends the three agencies for this initiative and encourages this and other opportunities to extend the usefulness of the R&D data collected by enhancing them through matching with like datasets. We urge that the data files that result from these ongoing matching operations be made available, under the protections to assure the confidentiality of individual responses that are guaranteed by the Census Bureau's Center for Economic Studies, for the conduct of individual research and analytical studies (**Recommendation 3.9**).

Innovation Survey

The panel considered the several attempts to collect data on innovation here and abroad, as well as the need for such data to illuminate the amount and outcomes of innovation activity in the economy. The panel concludes that innovation, linked activities, and outcomes can be measured and the results used to inform public debate or to support public policy development.

Furthermore, the panel recommends that resources be provided to SRS to build an internal capacity to resolve the methodological issues related to collecting innovation-related data. The panel recommends that this collection be integrated with or supplemental to the Survey of Industrial Research and Development. We also encourage SRS to work with experts in univer-

sities and public institutions who have expertise in a broad spectrum of related issues. In some cases, it may be judicious to commission case studies. In all instances, SRS is strongly encouraged to support the analysis and publication of the findings (**Recommendation 4.1**).

An additional recommendation is that SRS, within a reasonable amount of time after receiving the resources, should initiate a regular and comprehensive program of measurement and research related to innovation (**Recommendation 4.2**).

Surveys of Federal R&D Spending

In reviewing the accounting framework basis for the federal funds survey, the panel considered the growing, important uses of the federal science and technology (FS&T) budget. The panel recommends that NSF continue to collect those additional data items that are readily available in the defense agencies and expand collection of expenditures for those activities in the civilian agencies that would permit users to construct data series on FS&T expenditures in the same manner as the FS&T presentation in the president's budget documentation (**Recommendation 5.1**).

The panel reviewed the basis for collection of the data from federal agencies and compared the NSF procedures with the collection methodology employed in the RAND Research and Development in the United States (RaDiUS) database, which uses data from primary contract, grant, and cooperative agreement files as the data sources. Currently, the RaDiUS database is not adequate for obtaining estimates of federal government spending by science field. The panel urges NSF, under the auspices of the E-Government Act of 2002, to begin to work with OMB to develop guidance for standardizing the development and dissemination of R&D project data as part of an upgraded administrative records-based data system (**Recommendation 5.2**).

Similarly, the panel recommends that NSF devote attention to further researching the issues involved with converting the federal support survey into a system that aggregates microdata records taken from standardized, automated reporting systems in the key federal agencies that provide federal support to academic and nonprofit institutions (**Recommendation 5.3**).

Academic R&D Surveys

Noting that it has been some three decades since the field-of-science classification system has been updated, and that the current classification structure no longer adequately reflects the state of science and engineering fields, the panel recommends that it is now time for OMB to initiate a review of the *Classification of Fields of Science and Engineering*, last pub-

lished as Directive 16 in 1978. The panel suggests that OMB appoint the Science Resources Statistics Division of NSF to serve as the lead agency for an effort that must be conducted on a government-wide basis, since the field classifications impinge on the programs of many government agencies. The fields of science should be revised after this review in a process that is mindful of the need to maintain continuity of key data series to the extent possible (**Recommendation 6.1**).

The panel recommends that NSF engage in a program of outreach to the disciplines to begin to develop a standard concept of interdisciplinary and multidisciplinary research and, on an experimental basis, initiate a program to collect this information from a subset of academic and research institutions (**Recommendation 6.2**).

We are concerned that the apparently growing collaborative environment for the conduct of R&D is not adequately reflected in the academic spending survey. The panel recommends that NSF consider the addition of periodic collection of information on industry-government-university collaborations as a supplemental inquiry to the survey of college and university R&D spending. A decision on the viability of this collection should be preceded by a program of research and testing of the collection of these data (**Recommendation 6.3**).

With regard to the academic expenditure survey, the panel observes that the exact procedure used by NSF for imputation is not well documented, but it appears that imputation is used for unit nonresponse—a practice that is highly unusual in surveys. In most surveys, unit nonresponse is handled by weighting, as it was in this survey in 1999. At a minimum, NSF is urged to compare the results of imputation and weighting procedures (**Recommendation 6.9**).

We balance our concern over the burdensome nature of the survey of academic scientific and engineering research facilities with evidence that the data have important uses, including to those who provide that data. Sensitive to these concerns, the NSF staff has recently introduced several innovations in the questionnaire and in process automation. The panel recommends that the experience in the fielding of the revised questionnaire in 2003 be carefully evaluated by outside cognitive survey design experts, and that the results of those cognitive evaluations serve as the foundation for subsequent improvements to this mandated survey (**Recommendation 6.7**). This recommendation supplements our recommendation that NSF continue to conduct a response analysis survey to determine the base quality of these new and difficult items on computer technology and cyber infrastructure, study nonresponse patterns, and make a commitment to a sustained program of research and development on these conceptual matters (**Recommendation 6.8**).

Nonprofit Sector Survey

In reviewing the attempts by NSF to collect data on the nonprofit sector, the panel noted that there were evident problems that were well documented in the methodology report on this survey. Nonetheless, the panel recommends that another attempt should be made to make a survey-based, independent estimate of the amount of R&D performed in the nonprofit sector (**Recommendation 3.10**). The panel also recommends that NSF evaluate the possibility of collecting for nonprofit institutions the same science and engineering variables that pertain to academia (**Recommendation 5.3**).

DISCREPANCY BETWEEN SURVEYS

In evaluating the potential sources of the apparent discrepancy between the federal reports of spending on R&D and the reports of performers of R&D, the panel concludes that much of the discrepancy is caused by the use of improper metrics. The panel recommends that future comparisons of federal funding and performer expenditures be based on outlays versus expenditures, not obligations versus expenditures (**Recommendation 7.1**). However, the discrepancy can be an early and important sign of problems in one or more of the surveys. The panel's recommendation is that a reconciliation of the estimates of federal outlays for R&D and performer expenditures be conducted by NSF on an annual basis (**Recommendation 7.2**).

SURVEY ADMINISTRATION AND MANAGEMENT

The panel makes several recommendations concerning the administrative and management functions of NSF with regard to the surveys. Noting that the SRS division is considered a full-fledged federal statistical agency but that it is somewhat buried in NSF, the panel nonetheless could find no compelling reason to suggest that SRS be relocated organizationally within NSF. However, we have the sense that an elevation of the visibility of the resource base for SRS would be a positive step and would serve to direct attention to the needs of the programs for sustainment and improvement.

There are several tools that NSF has in its toolbox that will help the agency gain more control over aspects of survey operations. As a start, the panel recommends that NSF, in consultation with its contractors, revise the Statistical Guidelines for Surveys and Publications to set standards for treatment of unit nonresponse and to require the computation of response rates for each item, prior to sample weighting (**Recommendation 8.4**).

The panel would like to note that significant progress has been made

by the Science Resources Statistics Division in fostering an environment for the improvement of data quality. We continue to be hopeful that these recent initiatives, buttressed by additional resources and supplemented by further initiatives such as those outlined in this report, will lay a basis for further improvements in the future.

Measuring Research and Development Expenditures

IN THE U.S. ECONOMY

Panel on Research and Development Statistics at the
National Science Foundation

Lawrence D. Brown, Thomas J. Plewes, and Marisa A. Gerstein,
Editors

Committee on National Statistics

Division of Behavioral and Social Sciences and Education

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

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

THE NATIONAL ACADEMIES PRESS • 500 FIFTH STREET, N.W. • Washington, DC 20001

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

This study was supported by Contract No. SRS-0244598 between the National Academy of Sciences and the National Science Foundation. Support of the work of the Committee on National Statistics is provided by a consortium of federal agencies through a grant from the National Science Foundation (Number SBR-0112521). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

Library of Congress Cataloging-in-Publication Data

Measuring research and development expenditures in the U.S. economy / editors, Lawrence D. Brown, Thomas J. Plewes, Marisa A. Gerstein.
p. cm.

Panel to Review Research and Development Statistics at the National Science Foundation. The panel was convened in January 2002 by the Committee on National Statistics of the National Academies at the request of the National Science Foundation to conduct an in-depth and broad-based study of the Research and Development Statistics program of the NSF Science Resources Statistics (SRS) Division. The study was later mandated by Congress in the National Science Foundation Authorization Act of 2002.

Includes bibliographical references.

ISBN 0-309-09320-1 (pbk.) — ISBN 0-309-54595-1 (pdf)

1. Scientific surveys—United States. 2. Science and industry—United States. 3. Research, Industrial—United States. 4. Research—Economic aspects—United States. 5. Research, Industrial—Economic aspects—United States. 6. Science and industry—Government policy—United States. 7. Research, Industrial—Government policy—United States. I. Brown, Lawrence D. II. Plewes, Thomas J. III. Gerstein, Marisa A. IV. National Research Council (U.S.). Panel to Review Research and Development Statistics at the National Science Foundation.

Q149.U5M32 2004

658.57'0973—dc22

2004022518

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, <http://www.nap.edu>.

Printed in the United States of America.

Copyright 2005 by the National Academy of Sciences. All rights reserved.

Suggested citation: National Research Council. (2005). *Measuring Research and Development Expenditures in the U.S. Economy*. Panel on Research and Development Statistics at the National Science Foundation, Lawrence D. Brown, Thomas J. Plewes, and Marisa A. Gerstein, Editors. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

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

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

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

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

www.national-academies.org

PANEL ON RESEARCH AND DEVELOPMENT STATISTICS AT THE
NATIONAL SCIENCE FOUNDATION

LAWRENCE D. BROWN (*Chair*), Department of Statistics, The
Wharton School, University of Pennsylvania
JOHN L. ADAMS, RAND Corporation, Santa Monica, California
WESLEY M. COHEN, Fuqua School of Business, Duke University
FRED GAULT, Science, Innovation, and Electronic Information Division,
Statistics Canada, Ottawa
JAY HAKES, Jimmy Carter Library and Museum, Atlanta
BRONWYN H. HALL, Department of Economics, University of
California, Berkeley
CHRISTOPHER T. HILL, Public Policy and Technology, George Mason
University, Fairfax, VA
STEVEN KLEPPER, Department of Economics and Social Science,
Carnegie Mellon University
JOSHUA LERNER, Department of Finance and Department of
Entrepreneurial Management, Harvard Business School
BARUCH LEV, Stern School of Business, New York University
GARY McDONALD, Department of Mathematics and Statistics,
Oakland University, Rochester, MI
NORA CATE SCHAEFFER, Department of Sociology, University of
Wisconsin, Madison
RICHARD VALLIANT, Joint Program in Statistical Methodology,
University of Michigan

THOMAS J. PLEWES, *Study Director*
MARISA A. GERSTEIN, *Research Assistant*
TANYA M. LEE, *Program Assistant*
BARBARA A. BAILAR, *Consultant*
MICHAEL McGEARY, *Consultant*
WILLIAM F. LONG, *Consultant*

COMMITTEE ON NATIONAL STATISTICS
2003-2004

JOHN E. ROLPH (*Chair*), Marshall School of Business, University of Southern California
JOSEPH G. ALTONJI, Institute for Policy Research and Department of Economics, Northwestern University
ROBERT BELL, AT&T Research Laboratories, Florham Park, NJ
LAWRENCE D. BROWN, Department of Statistics, The Wharton School, University of Pennsylvania
WILLIAM EDDY, Department of Statistics, Carnegie Mellon University
ROBERT M. GROVES, Survey Research Center, University of Michigan
JOHN HALTIWANGER, Department of Economics, University of Maryland
PAUL HOLLAND, Educational Testing Service, Princeton, NJ
JOEL L. HOROWITZ, Department of Economics, Northwestern University
WILLIAM KALSBECK, Survey Research Unit, Department of Biostatistics, University of North Carolina
ARLEEN LEIBOWITZ, School of Public Policy and Social Research, University of California, Los Angeles
VIJAYAN NAIR, Department of Statistics, Department of Industrial and Operations Engineering, University of Michigan
DARYL PREGIBON, Google, New York, NY
KENNETH PREWITT, Russell Sage Foundation, New York, NY
NORA CATE SCHAEFFER, Department of Sociology, University of Wisconsin, Madison

CONSTANCE F. CITRO, *Director*

Contents

Preface	ix
Executive Summary	1
1 Introduction	10
2 Uses and Users	33
3 Measuring R&D in Business and Industry	48
4 Measuring Innovation in Business and Industry	91
5 Measuring R&D Spending in the Federal Government	102
6 Measuring R&D Activity in Academic Institutions	119
7 Analysis of the R&D Data Funding Discrepancies	139
8 Survey Management and Administrative Issues	152
References	177
Appendix A Workshop Agenda	187
Appendix B Biographical Sketches of Panel Members and Staff	191

Preface

The nation's commitment to research and development (R&D) has a history that goes back to the earliest days of the United States; some date that commitment to the patent clause in the Constitution, which set the stage for an interplay between the government and the scientific community that continues to this day. Scientific discoveries, technical applications, and, more recently, federal R&D programs have been integral to the nation's development. The role of R&D in the U.S. economy has been characterized as essential to national prosperity. The statement "Research and development widely recognized as being key to economic growth" begins the discussion of trends in R&D expenditures in the National Science Board's report, *Science and Engineering Indicators—2002* (p. 4). The report goes on to suggest that R&D decision making is critical to the future of the U.S. economy and national well-being.

Recognizing the importance of research and development to national prosperity, the United States and other countries have developed systems that conceptualize, define, and measure research and development activities so as to understand and to quantify the influence of R&D on the economy. It is claimed that the level of expenditures and the composition of those expenditures in R&D may be considered a proxy measure of national and sectoral commitment to economic growth and development, and that expenditures are an indication of the "perceived economic importance of R&D relative to all other economic activities" (National Science Board, 2002:4).

Despite this high level of attention to the role of R&D in the economy, the U.S. government devotes little of its treasury to the measurement of

R&D. If the relative spending of the federal government on the measurement of R&D expenditures can be viewed as a measure of its commitment to R&D, that commitment is problematic. In fiscal year 2003, the federal government spent approximately \$4.2 million on the National Science Foundation programs to measure R&D expenditures, a minuscule part of the nearly \$4.7 billion national investment in statistical programs.

The fast-changing environment for research and development severely impacts the R&D Statistics Program managed by the Science Resources Statistics Division (SRS) at the National Science Foundation (NSF). The NSF mission includes being a source of information for policy formation by other agencies of the federal government and providing data and analysis on a broad policy area for public- and private-sector constituents. SRS products also inform congressional policy makers and NSF itself.

As a major organization in the National Science Foundation, SRS has a requirement for a 5-year review of its programs by an NSF Committee of Visitors.

The review requirement has been partially met in the recent past by such activities as a U.S. Census Bureau advisory committee that recently reviewed the statistical foundation of the R&D industry survey; an academic R&D advisory panel that contributed guidance on collection of information on R&D activities in colleges and universities; and a series of workshops, such as a federal agency R&D administrator workshop in 1998, that have been periodically convened to examine certain aspects of program operations. The R&D surveys have also been reviewed in several recent studies by the National Research Council (2000, 2001c) and also by both the U.S. General Accounting Office (2001) and the Congressional Research Service (1999).

These reviews identified problems with specific R&D surveys, with inconsistencies in data among surveys, and with the inability of the current portfolio to address important policy questions and research issues generated by changes in how R&D is carried out in the United States today. Examples of these problems include high nonresponse rates to items in the industrial R&D survey; discrepancies of several billion dollars between the amount of money that federal agencies reported as R&D support and the amount that the performers of the R&D work reported spending; inconsistent classification of R&D by field across surveys; and the failure of existing surveys to capture interdisciplinary research and collaborations among firms, between research institutions, and across sectors.

In view of the special role of SRS, NSF leadership decided that a review of the R&D statistics programs by a panel of the National Research Council's Committee on National Statistics (CNSTAT) would meet the internal requirement for a periodic program review. A review of the entire portfolio of R&D data collection activities of SRS is timely to identify its

strengths and weaknesses and to recommend ways to improve data quality and relevance. In addition to fulfilling the requirement for a periodic outside evaluation, the NSF leadership sponsored this review as a means of developing a comprehensive, integrated review of concepts, definitions, survey methodology, and dissemination.

As result of this interest, the Committee on National Statistics convened the Panel on Research and Development at the National Science Foundation in early 2003 to conduct an in-depth and broad-based study to look at the Research and Development Statistics Program of the Science Resources Statistics Division. The goal of the panel has been to look at how R&D surveys are currently conducted and how they should be conducted to capture the country's R&D activities over the coming decade.

The CNSTAT committee has conducted its work in cooperation with a separately appointed panel of the Board on Science, Technology, and Economic Policy (STEP). The STEP panel focused on the issues of composition, structure, sourcing, and location, particularly in the context of the industrial R&D and federal funds surveys, covering the majority of U.S. funding and performance. The STEP panel's workshop, held in April 2003, informed the deliberations of this panel.

Our panel represents expertise in the fields of survey methodology, data analysis, statistics, economics, research analysis, and the organization and conduct of R&D in the public and private sectors. In selecting the members of the panel, the National Research Council sought to include representation from data users in the fields of science and technology policy and the analysis of technological change, as well as data providers from industry and academia, the two major sources of raw data. Because many of the practices and procedures for providing data depend on tax law, panel expertise in the valuation of intangible assets and the treatment of R&D in the tax code was also considered essential.

In conducting this review, the panel examined existing R&D data collection systems and relevant literature, commissioned appropriate papers, identified gaps in current methodology, and held a workshop on R&D measurement methodology. The panel held five meetings beginning in January 2003, as well as a workshop on July 24-25, 2003. The papers the panel commissioned address specific issues. In focusing on issues of statistical accuracy and reliability, the panel had the benefit of advice and analysis from two meetings with staff of NSF's Division of Science Resources Statistics and the public and private organizations directly responsible for data collection and compilation. In addition, subject-matter experts from the panel met to explore issues of statistical methodology and cognitive aspects of data collection.

The panel issued an interim report in March 2004. This interim report presented the panel's findings and conclusions regarding the present array

of surveys on matters of statistical accuracy and reliability, as well as interim recommendations on near-term improvements that should be considered and could be implemented by NSF in developing plans and making resource decisions for the next several years. The interim report's major findings and recommendations have been fully incorporated into this report.

The interim report found that significant progress has been made by the Science Resources Statistics Division in fostering an environment for improvement of data quality. The panel expressed hope that these recent initiatives, buttressed by additional resources and supplemented by further initiatives such as those outlined in this report, will lay a basis for further improvements in the future.

The interim report focused on four basic methodological issues: web-based collection, the practice of providing prior-year data to survey respondents, the designation of respondents, and nonresponse adjustment and imputation.

The very newness of web-based collection in the academic and federal government surveys suggests both opportunities and challenges. Web-based collection can afford efficiencies and economies and promises to improve such functions as editing and imputation. However, in reviewing how web-based collection is now implemented, the panel raised several cautions, suggesting that additional research is needed on such issues as self-imputation that is forced by demanding a data entry in each cell.

The panel considered the practice of providing data collected in prior-year responses to respondents. Although we have some confidence that the practice does not generate significant errors, the panel urged NSF to sponsor research on the effect of imprinting prior-period data on the industrial R&D survey in conjunction with testing the introduction of web-based data collection.

The interim report found that the industrial R&D survey gives limited attention to interaction with survey respondents, beyond the largest reporters, and then only when questions or issues of nonresponse are encountered. The panel supported the initiative to identify individual respondents in companies as a first and necessary step toward developing an educational interaction with respondents, so as to improve response rates and the quality of responses. The panel also strongly recommended that NSF and the U.S. Census Bureau resume a program of field observation staff visits to a sampling of reporters to examine record-keeping practices and conduct research on how respondents fill out the forms.

The R&D surveys have very different approaches to the treatment of nonresponse and the imputation of missing values. The panel recommended that NSF revise its Statistical Guidelines for Surveys and Publications to set standards for treatment of unit nonresponse and to direct the computation of response rates for each item, prior to sample weighting.

The interim report also looked at and made recommendations in regard to each of the surveys. The major findings of the interim report are reflected in the discussions of each of these surveys in this final report.

The Panel on Research and Development Statistics at the National Science Foundation gratefully acknowledges the contributions of many individuals and organizations which supported our activities over the 2-year study period. Not all of those who gave so freely of their time, energy, and knowledge can be adequately acknowledged. As we conducted our work, it was obvious that the topic of measurement of research and development is of intense interest to a wide-ranging community of government, academic, private, and private nonprofit individuals and organizations. Without their assistance and encouragement, we could not have completed this work.

Our appreciation begins with the sponsor of the study, the Science Resources Statistics Division of the National Science Foundation. The director of this division, Lynda Carlson, was instrumental in identifying the need for the study and assisting the panel in framing the scope and approach to the study. Her insightful comments on the occasion of the first meeting of the panel presaged a very successful undertaking. She ensured that appropriate members of her staff were available to assist the panel as we approached our task, and invariably recognized the independence of the panel.

Special recognition should be given to John Jankowski, program director, Research and Development Statistics Program, who was patient, supportive, and invariably helpful in fielding the many requests for information about the programs from the panel members and staff. His key survey managers and support staff, including Rich Bennof, Leslie Christovich, Mage Machen, Ron Meeks, Francisco Moris, Brandon Schackleford, and Ray Wolfe, provided support and assistance beyond the expected. Other members of the staff of the SRS division likewise contributed their time and talents to this product. The division's assistant director, Mary Frase, as well as Fran Featherstone, Ron Fesco, Rolf Lemming, and Jeri Mulrow, assisted in laying out issues of statistical methodology and data quality. The constant interest and support of Norman Bradburn, the head of NSF's Directorate for Social, Behavioral, and Economic Sciences, was absolutely essential to the success of the enterprise. Although we recognize the crucial assistance of all the members of the NSF staff, we want to emphasize that the deliberations and recommendations of the panel are the panel's own.

Almost uniquely among federal statistical agencies, SRS relies on the services of the U.S. Census Bureau, as well as private contractors, especially QRC Macro Corporation, to perform many key functions in the surveys, including data collection. The panel, in turn, relied on the knowledgeable staff of these organizations to assist in understanding the survey operations. Staff of these organizations gave presentations at both of the panel's open

meetings and participated fully in two “mini-meetings” involving members of the panel who focused on specific aspects of survey operations. The Census Bureau staff who were invariably helpful included William Bostic, Stacey Cole, Paul Hsen, Kimberly Moore, John Slanta, and Julius Smith. Adding strength and perspective to the discussion of cognitive issues in survey design at the mini-meeting on cognitive issues was Don Dillman of Washington State University.

The panel thanks QRC Macro staff members Susan Akin, Dan McMaster, Mark Morgan, and Michael Rossi, and Jim Smith of WESTAT, who freely assisted us in understanding the strengths and limitations of their operations.

The panel is also indebted to many others who made presentations on data and methodological issues, assisting us to focus on all aspects of the surveys, from concepts to uses. In addition to Lynda Carlson and John Jankowski, who participated in the initial meeting of the panel, the panel benefited from the work of Michael Bordt, Statistics Canada; Donna Fossum, RAND Corporation; Barbara Fraumeni, Bureau of Economic Analysis; Michael Gallaher, Research Triangle Institute; August Goetzfried, EUROSTAT; David Goldston, House Science Committee; Tomorhio Ijichi, University of Tokyo; Anita K. Jones, University of Virginia; Key Koizumi, American Association for the Advancement of Science; Anna Larsson, EUROSTAT; Gregory Tassej, National Institute of Science and Technology; and John Walsh, University of Tokyo.

The study director conducted many interviews with several other prominent users, providers, and producers of data on research and development. In addition to those who gave presentations at the open meetings of the panel, the panel recognizes the contributions of David Appler, Federal Laboratory Consortium for Technology Transfer; William Bonvillian, Office of Senator Lieberman; Rick Cheston, Government Accountability Office; Michael Davey, Congressional Research Service; Floyd DesChamps, Senate Committee on Commerce, Science, and Technology; Mark Herbst, Office of the Secretary of Defense; Ken LaSala, Senate Committee on Commerce, Science, and Technology; Chan Lieu, Senate Committee on Commerce, Science, and Technology; Charles Ludlam, Office of Senator Lieberman; Bob Palmer, House Committee on Science; David Radzanowski, U.S. Office of Management and Budget; Maurice Swinton, Small Business Administration; Jean Tol-Eisen, Senate Committee on Commerce, Science, and Technology; and David Trinkle, U.S. Office of Management and Budget.

The panel wishes to thank the senior staff of the General Motors Research and Development Center, which hosted members of the panel at its facilities in October 2003 for a wide-ranging discussion of the company’s organization for research and development and its use of federal R&D