369

90th Congress
1st Session

JOINT COMMITTEE PRINT

THE COORDINATION AND INTEGRATION OF GOVERNMENT STATISTICAL PROGRAMS

REPORT

OF THE

SUBCOMMITTEE ON ECONOMIC STATISTICS

OF THE

JOINT ECONOMIC COMMITTEE
CONGRESS OF THE UNITED STATES

Printed for the use of the Joint Economic Committee

U.S. GOVERNMENT PRINTING OFFICE

82-340O WASHINGTON : 1967

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402 Price 10 cents

LETTERS OF TRANSMITTAL

AUGUST 9, 1967.

To the Members of the Joint Economic Committee:

Transmitted herewith for your consideration and use, and for the use of other Members of Congress and other interested parties, is a report on "The Coordination and Integration of Government Statistical Programs" by the Subcommittee on Economic Statistics.

WILLIAM PROXMIRE,
Chairman, Joint Economic Committee.

AUGUST 7, 1967.

Hon. WILLIAM PROXMIRE,
Chairman, Joint Economic Committee,
Congress of the United States,
Washington, D.C.

DEAR MR. CHAIRMAN: Transmitted herewith is a unanimous report of the Subcommittee on Economic Statistics on "The Coordination and Integration of Government Statistical Programs."

This report is based on 4 days of hearings and on the compendium of views and suggestions entitled "Improved Statistics for Economic Growth," which was put together under your chairmanship of this subcommittee. The report contains a summary of our findings, together with our recommendations.

Statements received and testimony taken in the hearings have been very helpful to the subcommittee in formulating its recommendations.

We wish to thank the witnesses for their excellent papers and thoughtful observations. The participating witnesses were:

JOHN AIKEN, *Executive Director,* Federal Statistics Users' Conference.

RAYMOND T. BOWMAN, *Assistant Director for Statistical Standards,* Bureau of the Budget.

EWAN CLAGUE, Formerly *Commissioner of Labor Statistics,* Bureau of Labor Statistics.

EDGAR S. DUNN, Jr., *Research Analyst,* Resources for the Future, Inc.

ARTHUR M. OKUN, *Member,* Council of Economic Advisers.

RICHARD RUGGLES, *Professor of Economics,* Yale University.

FREDERICK STEPHAN, *Professor of Social Statistics,* Princeton University (Past President, American Statistical Association).

Sincerely yours,

HERMAN E. TALMADGE,
Chairman, Subcommittee on Economic Statistics.

CONTENTS

THE COORDINATION AND INTEGRATION OF GOVERNMENT STATISTICAL PROGRAMS

INTRODUCTION

In its 1967 Economic Report, the Joint Economic Committee directed the Subcommittee on Economic Statistics "to look into the possibilities of a truly integrated system providing genuinely comparable statistics consistent with and meshed into an overall system of economic statistics including the Federal, State, and local governments." Two considerations gave rise to this directive: (1) statistical information about the economy is necessary for good management in both the public and the private sectors; and (2) at this time in the evolution of our statistical services, significant improvements lie in the direction of further coordination and integration.[1]

In our modern private enterprise economy and in our system of Federal, State, and local governments, statistics provide much of the information which enables individuals, businesses, and governments to pursue their goals intelligently. A wide distribution of appropriate, accurate, prompt, and comprehensive intelligence is absolutely essential to the efficient functioning of a free society.[2]

Investigations by this subcommittee and by others have indicated that further significant improvement in our statistical services depends upon a higher degree of integration and coordination of our statistical programs. Indeed, there are strong indications that this is the aspect of the statistical system where progress is needed most.

The central theme of respondents in this subcommittee's compendium of views and suggestions by experts on improving our economic statistics was a recognition of the need to improve the organization of statistical data so that they can be more effectively used in analysis and policy. This general feeling was succinctly expressed by Prof. Wassily Leontief, of Harvard:

> The postwar revolution in economic programing and management techniques reinforced by unprecedented advances in data processing technology has made the traditional approach to collection, organization, and dissemination of facts and figures describing the operation of the American economy and the social conditions of the public and private life of the 180 million of American citizens completely obsolete.[3]

The staff of the Joint Economic Committee in a recent study of *Productivity, Prices, and Incomes* pointed out:

[1] By coordination and integration is meant a system of interrelated, accessible statistics, which conform to uniform definitions, classifications, time periods and quality standards.

[2] The importance of statistics for a free society was stressed by Raymond T. Bowman, Assistant Director for Statistical Standards, Bureau of the Budget, in his address "Achieving an Integrated Federal Statistical Program," delivered to the Annual Meeting of the Federal Statistics Users Conference, Oct. 2, 1957, and reprinted in *The National Economic Accounts of the United States*, hearings before the Subcommittee on Economic Statistics, 1957, p. 70.

[3] "Improved Statistics for Economic Growth—A Compendium of Views and Suggestions from Individuals, Organizations, and Statistics Users," materials submitted to the Subcommittee on Economic Statistics of the Joint Economic Committee, July 1965, p. 90.

> * * * there are still major difficulties in regard to the integration of and availability of statistical data. * * * On the other hand, the advances in statistical methodology and in data processing equipment have made practical enormous advances in the quality, prompt availability, and integration of economic data.[4]

A recent task force, commissioned by the Bureau of the Budget, made the following evaluation in their report:

> As it is presently operated, the statistical system is both inadequate—in the sense of failing to do things that should and could be done and inefficient—in the sense of not doing what it does at minimum cost, or getting less for what it spends than might be possible.[5]

The coordination and integration of Government statistical programs involves far more than the organization of Government bureaucracy, although the potential increase in efficiency and resource saving is important in itself. Nothing less than the quality of our public and private economic policies is at stake. In 1957, the National Accounts Review Committee, set up by the National Bureau of Economic Research at the request of the Bureau of the Budget, made the following statement:

> Integration of the national economic accounts is desirable from three points of view. First, many economic problems require the use of several different kinds of information, and it is often necessary to move from the information provided by one kind of economic accounts to that provided by another. Second, from a statistical point of view, integrating the various kinds of economic accounts makes best use of the available data, with less duplication and with improvement in statistical accuracy. Finally, for the user of the national economic accounts, a single integrated system is easier to understand and use correctly than a number of different apparently unrelated or overlapping systems.

The potential improvements in efficiency through increased integration are large—not only in terms of the quality and quantity of statistical output, but also in terms of cost reduction for a given output. The estimated budget obligations for fiscal 1968 are $122 million for current programs and $163 million for all programs including periodic programs, which do not include the costs and inconvenience imposed on private respondents. It is therefore clear that every effort should be made to obtain maximum use of these data which have been secured at significant cost to the Government and to private firms and individuals.

In view of the Committee's directive, and of such expert opinions as cited, and in recognition of the primary importance of an integrated statistical system, the subcommittee held 4 days of hearings—May 17, 18 and June 7 and 8—on the "Coordination and Integration of Government Statistical Programs." The focus of the hearings was on the following questions:

(1) To what extent is there coordination and integration of our statistical programs?

(2) What are the implications on the efficiency of our statistical system?

---

[4] James W. Knowles, Executive Director, Letter of transmittal, Dec. 21, 1966, *Productivity, Prices, and Incomes*, materials prepared by the Committee staff for the Joint Economic Committee.

[5] "Report of the Task Force on the Storage of and Access to Government Statistics," Bureau of the Budget, Oct. 1966, p. 5.

(3) What proposals are there for a more integrated statistical system?
(4) Would further integration and particularly a national data center present a threat to personal privacy?
(5) Is the present administrative machinery adequate for the tasks at hand?

To What Extent Is There Coordination and Integration?

The statistical programs of the Federal and State Governments, although good by historical or relative standards, are not sufficiently integrated or coordinated. Indeed, we do not have "a national statistical system" but rather several subsystems and some rather unrelated pieces. The present situation has evolved over many years, whereby new statistical series were initiated in response to the changing forces in our economy and the changing concerns of the time. The various bodies of data were collected with different objectives in mind and by different agencies of Government. Different objectives were often unavoidable and even desirable, but they should not be allowed to lead to duplication or the neglect of general high priority statistical needs, as has too often been true.

Government statistical programs focus attention on measurements of particular economic, social, or demographic phenomena. The programs themselves are oriented to the production and publication of statistical series, and the primary device for information retrieval and dissemination is the printed publication.

In general, the bodies of data do not mesh according to any overall system and there is much inflexibility which often prevents fitting the micro data to behavioral models. Reasons for the incompatibility include the following:

(1) differing definitions, classifications, and timing of respondent reports when uniformity is needed;
(2) difficulties of access to original or source data;
(3) differing qualities of data and inconsistent documentation.

In regard to the coordination among producers and users of data, the present system appears to be pragmatic and informal. In many cases, the statistical programs are connected with the administration of particular programs, and decisions bearing on statistical matters (for example, questions concerning the publication of series and changes in definition), tend to be highly influenced by the groups most closely affected by the operating programs. In some cases, the agencies have advisory groups and in other instances, the agency contacts with users tend to be more informal. To the extent that statistical decisions are influenced by particular groups, there may be a tendency for general needs to go unmet.

The decentralized organization of the statistical system, like the programs themselves, has evolved without a detailed, overall plan. This decentralization has imparted a dynamic character to the system and has enabled it to meet many emerging needs. At the same time, the coordinating efforts of the Bureau of the Budget have prevented many of the major adverse effects of decentralized efforts. However, under present circumstances, much more needs to be done and can be done to further integration—without sacrificing the benefits stemming from the basic decentralization of the data producing processes. As

one of the witnesses, Professor Richard Ruggles, stated with reference to the present system:

> * * * The term "decentralized," while applicable, may be somewhat misleading * * * it has been the result of a jungle-like growth of statistical activities by different government agencies having widely differing purposes. The result has been extensive duplication and lack of coordination. * * * Although the Office of Statistical Standards has led to considerable improvement and rationalization, the Federal statistical system continues to operate on a highly decentralized and uncoordinated basis.

What Are the Implications on the Efficiency of Our Statistical System?

On the basis of testimony, the subcommittee concludes that the present methods of operating our statistical programs have become obsolete in the sense that the need for integration and data availability have outstripped progress. At the present time, statistical programs require radical improvement to take advantage of modern technology and to meet the new and rapidly expanding needs of policy and analysis.

Although the situation described in the previous section was not ideal before the advent of the computer, it was then infeasible to relate large bodies of data either at one period of time or over long periods of time. However, computer technology has drastically increased the speed and efficiency of handling data and has made possible entirely new kinds of analysis. It has also greatly simplified the problem of data storage and retrieval.

The statistical series produced by Government agencies serve reasonably well the traditional tasks of management. But they are totally inadequate to meet the changing policy needs of our times—to aid in the planning of modern corporations and in the design and operation of new Government programs and the improvement of existing ones. Our rather isolated pieces of information do not fit into a general system. But this is a minor part of the difficulty since every general system has its limitations and inflexibilities. The real difficulty with the present system is its lack of flexibility, which is needed to interrelate different bodies of data. One witness, Dr. Edgar Dunn, expressed the present situtation as follows:

> * * * We are still living in identical statistical "rowhouses," so to speak, when we have the technology and means to adopt the architecture to the size and interests of the family.

The information contained in the statistical publications cannot be easily manipulated or related. Moreover, the micro data that were used to produce the aggregated series have frequently not been maintained in usable form, nor are the various bodies of basic data compatible due to a lack of coordination in the original production processes. The detailed data, then, which are being collected at great cost and inconvenience are not being fully utilized to produce the wide variety of aggregate statistics needed or to study micro relationships. While some efforts are now going forward to put new additions to the stock of data into machine readable form for improved accessibility, much more effort on a Government-wide basis is needed.

In the area of national economic policy, improved integration would provide valuable information on productivity, prices, and incomes. For example, meaningful discussions of price-wage guideposts require quantitative information on productivity, prices, costs, and inputs for the total economy and by industry. Yet, at the present time, the output, input, price, and income data are estimated from different samples and by different statistical bureaus. The result is that we often cannot be sure whether a small computed change in unit-labor costs in a particular industry is due to sampling variability, differing sampling units, or differing classification procedures.

A second example where a more integrated system of data might significantly improve economic policy is in the area of fiscal policy. What determines the level of investment and thus full employment? Is it capacity utilization, corporate profits, previous investment, the cost of borrowing, or some other variable? Answers to questions of this type often involve bringing together financial data collected for enterprise units with nonfinancial data collected for establishment units. Significant progress in answering these questions might be made if researchers could relate relevant financial and nonfinancial data over an extended period of time or compare the situations of many diverse firms at one point in time.

In the field of human resource programs, further integration would help to improve the design and administration of government programs. For example, in the study of the causes of poverty, it would be advantageous to merge Census data with data from the operation of the Social Security Administration, and with income data from the Internal Revenue Service. The ability to combine data sets and to compare data over time would also help to refine the dimensions of poverty. Efforts to relate some of these bodies of data have been and are being made sporadically but these efforts must not only be increased but made a regular and organized feature of our statistical system.

The integration of State and local data with Federal data would aid in the design and operation of government programs which provide for special educational and training resources, or special aids to businesses. It would also be valuable to State and local government planning and to businesses in their planning of investment and sales. Here, too, efforts have begun, but they require far greater scope and support.

In addition to the loss of analytical capabilities stemming from the present disorganized system, there is unnecessary duplication of effort by government agencies and duplicative reporting by respondents. In part, these duplications arise because of the absence of a master directory of business respondents available to all statistical agencies.

In addition to duplication, the present dispersal of activities means that economies of scale in the use of automatic data processing and technical experts are not being fully realized. This, also implies that the information is being produced more slowly than would be required by modern processes.

What Proposals Are There for a More Integrated System?

The subcommittee heard proposals designed to improve the operation and particularly the integration of our statistical programs.

The most complex and potentially far reaching of the proposals was that of establishing a national statistical data center (sometimes referred to as a national statistical servicing center). Little agreement was expressed on specific aspects of such a center; however, its proposed general functions would include the following:

(1) To establish and maintain an index of all sets of data available within the Federal system and usable for statistical purposes.

(2) To assemble the data from selected sources, integrate them to the maximum feasible extent, and preserve the data in usable and highly accessible form by use of large scale automatic data processing equipment. (With modern technology, it may be unnecessary to physically locate the data in the center, in view of computer hookups and computer consoles.)

(3) To function under uniform disclosure standards, which should also apply to other Federal statistical agencies, in order to insure maximum protection of personal privacy.

(4) In cooperation with users in and out of government and with the collecting agencies, to formulate standards for further collection efforts in order to promote maximum integration and minimum duplication without undue sacrifice in terms of specific functions of the agencies.

(5) To define the accessibility and cost of various sets of data to potential users inside and outside of government. Special tabulations could also be performed by the center, with due consideration for disclosure restrictions and with costs assessed against the users.

The potential benefits from such a center are difficult to estimate since it would create an entirely new dimension in statistical servicing capability. In general terms, the center would provide for a more efficient use of collected information by providing greater flexibility.

First, the center would provide a staff and other resources that would be ready, willing, and able to service special requests for data and for information about data from users both inside and outside of Government. This would greatly increase the efficiency of data users since it would save hours of searching for the data, negotiating to obtain the data, and readying the data for use.

Second, the center's operation would gradually improve the quality and increase the quantity of data available, while either reducing the burdens on respondents or at least not increasing them. In part, this would arise because the basic micro data from the separate programs and agencies would be permanently stored, coded and made usable under standards of the national statistical center. Thus these micro data could be related to each other and/or recombined into new aggregates. At the same time these efforts would reinforce the other major function of the center—namely the improved coordination of the processes of collection, processing, storage, and retrieval of data by all the Government agencies. Since the operations could be carried out inside the center, manipulation of the data (for example, matching operations) could be carried out without violating disclosure principles or threatening personal privacy.

The financial cost of a Federal statistical servicing center cannot be satisfactorily estimated because the staff work essential to the specification and costing of program options has not been done. Realistic cost estimates can only be developed gradually as experience is gained. Mr. Raymond Bowman estimated that an initial budget of about $2 million per year for the first 2 years would be required to start a

statistical data center. He went on to estimate that in the following 3 years, the costs would be about $5 million per year, exclusive of computer hardware costs.

Although the center would incorporate several functions designed to further integration, some of these possibilities were mentioned as separate possibilities.

A master list of firms and establishments was proposed to aid in the design and operation of sampling and in the classification of data, for example by area and industry. Under the present situation, the Standard Industrial Classification System does not guarantee that a particular firm or establishment will be classified in the same industry by each of the different agencies, particularly in cases where the firm has several establishments, one or more of which produce products in different industries.

Several witnesses also testified as to the need for an index describing the statistical data collected by the Federal Government. The index would indicate such characteristics as where the data are stored, their accessibility, and sampling characteristics.[6]

The subcommittee probed for suggestions on ways that the Federal Government could encourage cooperation among statistical programs at the Federal, State, and local levels of government. Numerous witnesses stressed the need for Federal aids to promote an integrated network of statistics for the United States. Suggestion was made that operating programs involving two or more levels of government should have provisions to insure comparable and meaningful statistics.

Would Further Integration and Particularly a National Data Center Present a Threat to Personal Privacy?

Further integration can and should be attained without sacrificing principles of personal privacy. Modern technological capabilities for data association, storage, and retrieval, like most technological changes, present new possibilities for use or misuse. The rational approach to these new capabilities is not fear to act, but rather action to control technology for man's use.[7]

From the standpoint of the national data center as it relates to the issue of personal privacy, relevant considerations include (1) the determination of what information might go into the center, and (2) the design of adequate safeguards to prevent misuse of information.

In regard to the first point, it should be pointed out that the proposed national data center differs fundamentally in design and purpose from central files designed to hold information about individual people or individual firms. The purpose of the data center under discussion is to furnish information primarily about the "public face" of groups of people, households, firms, etc., and their interrelationships—never about the individual units themselves. This would involve

[6] The *Supplement to Economic Indicators*, prepared for the Subcommittee on Economic Statistics by the committee staff and the Office of Statistical Standards, Bureau of the Budget, describes some of the most commonly used series in the economic field. Major Federal statistical agencies publish catalogs, price lists, and other types of guides to their statistical publications. Agency catalogs and other guides to statistical publications are listed in and complemented by the bibliography of "Principal Statistical Publications of Federal Agencies" published by the Bureau of the Budget in its booklet, *Statistical Services of the United States Government*.

[7] The issues concerning personal privacy have been the subject of investigation not only in the recent hearings of this subcommittee, but also hearings of the Special Subcommittee on Invasion of Privacy of the Committee on Government Operations, U.S. House of Representatives; and of the Subcommittee on Administrative Practice and Procedure of the Committee on the Judiciary, U.S. Senate.

statistical samples and only in rare cases would the center contain a 100-percent sample.

In regard to the second consideration, it should be pointed out that, although the purpose of the center necessitates that the identity of the statistical observations be maintained for matching operations, the identities of individuals and firms could be carefully safeguarded by strict legal and technical measures. For example, internal codes could be utilized to protect the identity of particular individuals or firms. It should be noted that some of our statistical agencies have already had long and successful experience with programs which have safeguarded personal privacy and prevented the disclosure of information about specific businesses. The chief drawback of the present operation of safeguards is that they are unevenly applied and where they are strictly applied, the use of the data may be more restricted than necessary.

The problem of safeguarding privacy should be treated seriously. The problems of legal and technical design and of day-to-day operation are formidable. Work on a national data center, if instituted, should go forward at a pace that is slow enough to allow reasonable assurance that the disclosure problems can be and are being controlled. Indeed, the national data center might improve on the present system safeguarding the rights of individuals to personal privacy and of confidentiality of business information. The data center would force a more explicit consideration of these pressing issues of our electronic age. Moreover, the center might cause us to move from the present ad hoc system to one of uniform and far-reaching principles.

Is the Present Administrative Machinery Adequate for the Task at Hand?

Although the subcommittee did not deal with legal or administrative aspects of our Federal statistical system in detail, it believes that the present legal and administrative machinery is generally adequate for the problems of coordination and integration.

The major responsibility for the coordinating function is with the Bureau of the Budget, through its Office of Statistical Standards. Legislation provides the Bureau with strong backing for its task of coordination. The Budget and Accounting Procedures Act of 1950 in Title I, Part I, Section 103, states that "The President, through the Director of the Bureau of the Budget, is authorized and directed to develop programs and to issue regulations and orders for the improved gathering, compiling, analyzing, publishing, and disseminating of statistical information. * * *" This provision of law is carried out under Executive Order 10253.

Specific authority is also provided by the Federal Reports Act of 1942 for the Director of the Bureau of the Budget (a) to transfer the responsibilities for the collection of statistical information from one agency to another and, with certain safeguards, to transfer information among agencies to avoid duplication and promote efficiency; and (b) to review, and approve or disapprove, reporting proposals by Federal executive agencies for obtaining information from the public.

Conclusions and Recommendations

Major and far-reaching reforms are needed to make possible full utilization of our statistical data, affecting every major function of the statistical system, including concepts and definitions, sampling design, collecting and tabulating procedures, and storage, retrieval, and dissemination of information. The changes cannot be accomplished overnight, but efforts should be greatly accelerated at once.

The subcommittee has the following recommendations:

I. Work should proceed toward the establishment of a national statistical servicing center. A beginning should be made with certain data judged to be most critical to public policy requirements; and as experience is gained, the operations of the center should be expanded.

As an early aspect of the servicing center, serious consideration should be given to the need, expressed by witnesses, for a national statistical index and library to serve users of Government statistics.

The design and operation of the center, even in its developmental stages, should not be relegated to a role of secondary importance, since by its very nature it requires cooperation at the highest levels of Government. The subcommittee believes that in this case the practice of the Office of Statistical Standards of assigning primary responsibility to a department is a sound one. Although subject to the review of the Office of Statistical Standards for administrative considerations, the subcommittee assumes that the Department of Commerce would be a logical place for the statistical data center.

II. Improvements in the efficiency, flexibility, and servicing capability of our statistical system should not and need not wait for the full establishment of a national statistical servicing center, since the Bureau of the Budget, through its Office of Statistical Standards, already has significant authority to establish uniform standards within the Federal Government. Steps to further integration should be taken now by the Office of Statistical Standards since the improvements are justified on their own merits and, in most cases, are necessary prerequisites to the implementation of a statistical servicing center. The steps should include but not be limited to the following:

(*a*) Develop and promulgate general operating principles for safeguarding rights of personal privacy to be applied uniformly throughout Government.

Difficulties in insuring personal privacy should be worked out simultaneously with steps taken toward the establishment of a statistical servicing center.

(*b*) Set down standards to insure that, in general, the micro data in the various agencies are stored in usable form—preferably machine-readable form—edited, and with full explanatory notes.

(*c*) Outline in detail the necessary steps to insure more uniform classifications throughout the Federal Government. For example, a master list of firms and establishments could aid in achieving consistent industrial classification and could result in cost reductions by reducing duplication in reporting and duplication of efforts in classifying respondents.

III. The integration of area data presents very difficult and complex problems. Although the subcommittee was encouraged by the reports of progress, it believes that study and innovation are particu-

larly needed to promote integration of statistics among the State and local governments, between the Federal and State Governments, and among nations and international organizations. In view of the complexity of the problems involved, the subcommittee recommends that the Bureau of the Budget prepare a detailed report of the problems involved, of specific ways in which the Federal Government can be of significant assistance, and of the costs of various alternatives.

IV. Additional resources for the Federal statistical system will be needed to carry out the recommendations outlined above. These additional costs, although they cannot be precisely estimated, would be relatively small in comparison to the overall statistical programs. The greater efficiency achieved, thereby, would well justify the investment.

At this point, it is important to stress that it is good economics and good business to spend the money for an effective information system. Although current statistical programs have been expanded and are more costly than in earlier years, they are still very small when compared to the impact of the public and private policies which they influence. The most significant increases in recent years have been for labor and demographic statistics—and within these categories the emphasis has been on social statistics—health, welfare, education, and poverty. This is appropriate, but at the same time, major advances in economic statistics must not be neglected. Too often it is difficult to engender support for general statistical programs, since they do not appear to have a specific and immediate impact on particular individuals or groups. But economic data influence the efficiency of our near-trillion-dollar economy and thereby directly affect the material well-being of every individual.

○