368

THE COORDINATION AND INTEGRATION OF GOVERNMENT STATISTICAL PROGRAMS

HEARINGS

BEFORE THE

SUBCOMMITTEE ON ECONOMIC STATISTICS

OF THE

JOINT ECONOMIC COMMITTEE

CONGRESS OF THE UNITED STATES

NINETIETH CONGRESS

FIRST SESSION

MAY 17, 18; JUNE 7, 8, 1967

Printed for the use of the Joint Economic Committee

U.S. GOVERNMENT PRINTING OFFICE

80-826 O WASHINGTON : 1967

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402 - Price 55 cents

CONTENTS

STATEMENTS AND SUBMISSIONS

MAY 17, 1967

MAY 18, 1967

JUNE 7, 1967

JUNE 8, 1967

APPENDIXES

THE COORDINATION AND INTEGRATION OF GOVERNMENT STATISTICAL PROGRAMS

WEDNESDAY, MAY 17, 1967

CONGRESS OF THE UNITED STATES,
SUBCOMMITTEE ON ECONOMIC STATISTICS OF THE
JOINT ECONOMIC COMMITTEE,
Washington, D.C.

The subcommittee met, pursuant to recess, at 10 a.m., in room 6226, New Senate Office Building, Hon. Herman E. Talmadge (chairman of the subcommittee) presiding.

Present: Senators Talmadge and Miller; and Representative Bolling.

Also present: John R. Stark, executive director; James W. Knowles, director of research; and George R. Iden, staff economist.

Chairman TALMADGE. The subcommittee will come to order.

The Subcommittee on Economic Statistics today begins the first of 4 scheduled days of hearings on the coordination and integration of Government statistical programs.

The issues at stake are indeed very significant, for the hearing involves the general quality and adequacy of our statistical programs. General economic knowledge is not as sufficient as it might have been when at least the direction of needed monetary and fiscal policies was clear. Now, when the economy is operating near full employment, accurate quantification of developments is essential. Near full employment, it is also essential to have statstics on the structure of the economy. Such data are needed for price-wage guideposts and for the operation of human resource programs.

The statistical needs of the private sector have also become increasingly demanding. Examples include data for business planning and for the large number of individuals engaged in social science research.

On the side of the supply of economic statistics, there have been technological advances of revolutionary proportion, making possible entirely new dimensions in statistical capabilities. These advances include the application of computer technology and electronic transmission and storage of data.

Developments in the demand for and the supply of economic statistics have not gone unnoticed. During the 80th Congress, almost 20 years ago, the Joint Economic Committee published a staff study entitled "Statistical Gaps" which reviewed some of the pressing statistical measures needed at that time, along with cost estimates. During the 89th Congress the Subcommittee on Economic Statistics, under the chairmanship of Senator Proxmire, conducted a study entitled "Improved Statistics for Economic Growth." The subcom-

mittee invited statements by private economists and statisticians on measures to strengthen our statistical tools. The Government agencies, in turn, were invited to comment on these views.

The current hearings are designed to carry forward this work and to carry out a directive contained in the Joint Economic Committee's Economic Report of 1967—

> * * * 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.

The hearings will focus on the problems and possibilities of our statistical system. It is hoped that the proceedings will furnish information and perspective on the following questions:

(1) To what extent is there coordination and integration of statistical series so that it is possible to relate various series bearing on current problems?

(2) Is the present system efficient in the sense of making full use of available information, and known technology, and of minimizing duplication?

(3) What are the present measures for coordinating statistical series within the Federal Government and between Federal and State and local Governments and between users and compliers of data?

(4) Would a National Statistical Center significantly improve the quality of statistical services rendered to users both within and outside of the Government?

(5) What are the implications of a possible statistical center for the problem of disclosure and of safeguarding the rights of individuals to personal privacy?

The subcommittee will seek to determine the extent to which the demand for economic data is being met and is being met efficiently and to examine alternatives in order to arrive at recommendations for the future improvement of our statistical system.

At this point in the record we will include the press release announcing the hearings and the schedule of same.

MAY 12, 1967.

CONGRESS OF THE UNITED STATES

JOINT ECONOMIC COMMITTEE

SENATOR TALMADGE ANNOUNCES HEARINGS ON GOVERNMENT STATISTICAL PROGRAMS

Senator Herman E. Talmadge (D., Ga.), Chairman of the Subcommittee on Economic Statistics of the Joint Economic Committee, today announced that his subcommittee will hold four days of hearings—May 17 and 18 and June 7 and 8—on the coordination and integration of government statistical programs.

In announcing plans for the hearings, Senator Talmadge said that the hearings can be highly productive through contributing to the adequacy and the efficiency of our government statistical system.

"The demand for government statistics on the part of public and private decision makers is expanding rapidly. We need more, better and more timely data. 'Fine tuning' of monetary and fiscal policies, the efficient operation of our many human resource programs, and State and local data for business planning needs present formidable information and analytical problems. These problems require for their solution relevant statistical series that can be both related and utilized rapidly.

"On the side of the supply of economic statistics, there have been technological advances of revolutionary proportion. These advances include the application of computer technology and the electronic transmission and storage of data. Although scarce, there is also a pool of highly skilled technicians.

"The subcommittee will seek to determine the extent to which the demand for economic data is being met efficiently and to examine alternatives in order to

arrive at recommendations for the future improvement of our statistical system."
A schedule of the hearings is attached.

SCHEDULE OF HEARINGS ON THE COORDINATION AND INTEGRATION OF GOVERNMENT STATISTICAL PROGRAMS, MAY 17, 18; JUNE 7, 8

Wednesday, May 17, 10:00 a.m., room 6226 New Senate Office Building:

A MORE ADEQUATE AND EFFICIENT GOVERNMENT STATISTICAL PROGRAM—THE PROPOSAL FOR A NATIONAL DATA CENTER

EDGAR S. DUNN, Jr., Research Analyst, Resources for the Future, Inc.
RICHARD RUGGLES, Professor of Economics, Yale University

Thursday, May 18, 10:00 a.m., room 1202 New Senate Office Building:

USERS' VIEWS ON THE COORDINATION AND INTEGRATION OF GOVERNMENT STATISTICAL PROGRAMS

JOHN AIKEN, Executive Director, Federal Statistics Users' Conference
FREDERICK STEPHAN, Professor of Social Statistics, Princeton University (Past President, American Statistical Association)

Wednesday, June 7, 10:00 a.m., room 1202 New Senate Office Building:

THE COORDINATION OF GOVERNMENT STATISTICAL PROGRAMS

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

Thursday, June 8, 10:00 a.m., room 6226 New Senate Office Building:

LONG-RUN POSSIBILITIES AND PROBLEMS

ARTHUR M. OKUN, Member, Council of Economic Advisers: *Statistics for Effective Public Policy*
EWAN CLAGUE, Formerly Commissioner of Labor Statistics, Bureau of Labor Statistics: *Goals and Difficulties of the Government Statistical Program*

Chairman TALMADGE. To open the hearings this morning, we are privileged to have as witnesses Dr. Edgar S. Dunn, Jr., and Dr. Richard Ruggles. Both of these gentlemen are eminently qualified to evaluate our statistical system and to discuss the possibilities for its continued improvement.

Dr. Dunn, of Resources for the Future, Inc., completed a report in 1965 entitled "Review of Proposal for a National Data Center," at the request of the Office of Statistical Standards of the Bureau of the Budget. Dr. Dunn's report was especially valuable because he has followed through by writing and speaking on this vital subject.

Dr. Richard Ruggles, professor of economics at Yale University, has often given generously of his time to contribute to the Joint Economic Committee. Professor Ruggles was the chairman of the committee on the preservation and use of economic data which was established by the Social Science Research Council. He also served on the recent "Task Force on the Storage of and Access to Government Statistics," established by the Bureau of the Budget.

Will you please begin, Dr. Dunn?

STATEMENT OF DR. EDGAR S. DUNN, JR., RESEARCH ANALYST, RESOURCES FOR THE FUTURE, INC.

Mr. DUNN. Thank you very much, Mr. Chairman. I have a modest sized statement here that I think is appropriate and is within the framework of the procedure laid down by your staff.

My name is Edgar Dunn. I am a professional economist currently employed as a research associate by Resources for the Future, Inc. I have been directly involved in the problems at issue in the subject matter of this hearing in three capacities: (1) As Deputy Assistant Secretary for Economic Affairs of the Department of Commerce, (2) as a consultant to the Office of Statistical Standards in the Bureau of the Budget in a review of proposals for a national data center, and (3) as a research economist and public servant who makes use of statistical resources in his work.

In my communication from the committee I have been asked to deal with two topics—the integration of Government statistics including the feasibility of a national data center, and the relationship of the problem of personal privacy to these objectives. Both are large and complex topics and I can do no more than highlight the issues in this statement.

First, let's consider the problem of effective statistical services.

Until quite recently the evolution of Government statistical programs have been marked by two principal characteristics. First, attention has been almost entirely limited to measures of particular economic, social or demographic phenomena. Attention was focused largely upon individual series such as the size of the population, the volume of foreign trade and the output of manufacturing. Most of the uses served either public or private management.

The series often found their origin in a particular management need. Second, Federal statistical programs are essentially production and publication programs. Their missions are defined in terms of the collection and tabulation of data with the aim of publication in statistical monographs. The printed publication is the primary device for information retrieval or dissemination. They provide what are hoped to be general-purpose tables for all users. This orientation has created a system that handles all of the problems of producing data in this form with admirable skill and efficiency, but it has also produced one that has little capacity for understanding the problems and the requirements of statistical use and provides no adequate mechanism for the priorities of statistical usage to find expression in program formulation and management. The respondent who supplies information to the system is an object of much greater concern and formal study than the user. It is like an automobile manufacturing industry with the management dominated by engineers.

We have been finding over the years that these management-oriented programs do not serve the information requirements of policy determination or social science research very well. The information needed to formulate and evaluate policy is usually more complex than that needed for its daily implementation or management. Information systems that have grown out of the needs of the latter don't accommodate themselves very gracefully to the service of the former.

We had this forcefully brought home to us in the 1930's when we found that the effort to establish public policy to cope with a stagnant and unstable economy was floundering for want of a comprehensive measure of the economic performance of the Nation—one that would allow us to relate in a meaningful way the components of the national output with the components of the national income. As a consequence we set up a special organization to produce such information—the Office of Business Economics—and over the last three decades we have

developed a system of economic accounts. However, we are increasingly finding the same kind of frustrations that motivated the formation of that program cropping up all across the board. Policy formulation and review in such areas as poverty, health, education, area development, and science policy—in addition to the traditional areas of fiscal and monetary policy—have been severely handicapped by the inability to engage the service of relevant and timely information.

Why is the present system incapable of serving adequately this kind of need?

Because the present system is dominated by the intermediate aggregates that were designed to serve a publication program. Since detailed data cannot be published—because of expense, prohibition of disclosure and, as a practical matter, we would be inundated if it were—the Government statistical programs have developed over the years little boxes with names corresponding to different industrial, demographic, and social characteristics of the individual respondents.

Thus, the data pertaining to the respondent households, individuals, and establishments are sorted out into these boxes—some labeled "male" or "female," and some "apparel" or "textile manufacturing"—and the resultant sums constitute the published statistical record.

There are two problems here. (1) The names on these boxes suffer from the fact that, whatever their origin, they do not provide descriptions of the attributes of the populations they contain that are adequate for many uses appropriate to research or public policy. (2) In the face of published aggregates there is no means for reaching back to the original observations in order to assemble them into different boxes with different names that might be more appropriate for policy or research use.

Now, particularly since the advent of the computer, there is no insurmountable technical limitation upon our ability to rearrange these collection boxes to generate information more appropriate to the requirements of each of the major objectives of policy and research. We should, in principle, be able to "reshuffle the deck" and "deal a new hand" as the situation requires. But to do this requires a Federal statistical program that sees its mission as more than a production and publication task. It requires one that sees its mission as a custodian of the file and a supplier of information services to a broad spectrum of important users. 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."

This means that the Government statistical programs need to develop a new capability for retaining in machine records the recorded attributes of respondents so that they can be retrieved and made subject to retabulation and computation to meet important needs that traditional publication tables cannot serve. Although some agencies have developed a limited capability in this area, by and large it is extremely difficult, if not impossible, to fulfill this kind of requirement.

There are two reasons, both interrelated, why the present system cannot perform in this way—in spite of the almost universal application of computers to the process of statistical collection, tabulation, and publication. First, the simple act of retrieval of respondent attributes in the interest of generating different aggregates or filling dif-

ferent boxes is commonly impossible or extremely difficult, the Federal system has no policy and no systematic organizatonal or funding provision to assure that basic statistical records will be maintained in appropriate archives. It is not uncommon for them to be destroyed after the limited publication objectives have been fulfilled. Where they do exist, they are commonly not documented, maintained or organized so they can serve as an effective basis for retrieval. There is no service to which the user can turn—equivalent to the reference librarian in the documentary library—for assistance in finding access to those elements of a labyrinth record that are appropriate for his use.

Second, even where retrieval can be effectively accomplished, it is often impossible to sort it into a box that may have special meaning for policy. For example, we may want to sort out of the basic record a measure of the number of Negro families in Appalachia with incomes under $3,000, a family size of four or more, with less than a high school education, and who are drawing public assistance. You see, the names on these boxes can become rather lengthy. This is particularly true of information that can serve policy. It doesn't take much imagination to see how vital information of this general type may be for establishing or reviewing public policy related to poverty or education. But, in servicing such information requirements, a special problem usually arises. The attributes of income may come from tax records, the demographic characteristics from census records, the public assistance record from the Social Security Administration, or some other agency.

Even where there is no problem of retrieving these data, problems arise because the different agencies and different programs define the basic respondent unit in different ways that preclude the association of their characteristics. For example, some business data are collected on an enterprise basis and some on an establishment basis. Matching attributes of these records becomes exceedingly difficult. Again, different agencies often define the characteristics differently. The electrical appliance industry in Bureau of Labor Statistics may be defined to contain a different array of establishments than the electrical appliance industry in Census. The respondent units may carry similar tags which, in fact, mean different things.

In short, serving policy or social research often involves bringing together data which are *separately generated* in the collection process but which pertain to *inherently connected* relationships in economic and social behavior.

This means that, if the potential for serving public policy and social research in a computer age is to be realized, government statistical programs will have to take as a fundamental part of their mission providing the services that are essential. This, in turn, implies changes in program orientation that reach all the way back to the standards and procedures for collecting, classifying and tabulating data, as well as providing the hardware and software capability for file retrieval, tape translation, file rearrangements, record matching and standard statistical routines.

What is the stake or payoff involved in reforms that would provide the capability of serving public policy and social research in this way?

There is unfortunately no way of giving a precise answer. We are talking about an altogether new kind of capability. The demand upon

the system cannot be estimated with any accuracy in advance of the existence of the capability. We can say from experience that there is a tendency to grossly underestimate the value of new systems in the information field. Let me give you two examples. Back in 1950 when the Eckert-Mauchly group were putting out the Univac I, which was the first large-scale commercial computer, IBM undertook a careful market study to determine whether they should try to get into this market. They concluded that there was a market for something like five or six of these machines in the entire United States and decided to stay out of the field. With 5 years 1,275 machines had been sold and the industry was turning to the design of a whole new generation of computers. IBM was late getting into the market for this reason. Another example, I was told recently by an official of the National Academy of Sciences that before they first acquired a Xerox machine they made a careful survey of the staff to estimate its use and decide upon appropriate equipment. Within a period of less than 2 years they had exceeded their estimate by something like a factor of 10 and had gone through two changes of equipment.

The great value of an innovation in statistical services of the kind envisaged cannot be measured by trying to find out who is going to use it, for what purpose, and then estimate their benefits. Indeed, if we could identify and measure the uses to which it would be applied we could be sure that we would have, in fact, a more limited capability than we seek. The value lies in the great flexibility that is inherent in its design. It can more easly respond to the needs of different requirements and adapt to the needs of new requirements. No techniques in technology or social organization is more prized in a complex and changing world than one which is flexible to changing requirements—that incorporates a basic or generic capability. This is the prize we seek. And, while recognizing that the governmental statistical programs are by no means the only or always the most important sources of information for policy, it is still fair to say that the real stake is the degree of our success in public policy and public management in a complex and changing world. Good information is the root of success in that endeavor.

How, then, are we to achieve such reforms in the servicing capabilities of government statistical programs?

There are those who know the problems well who maintain that the best solution (some say the only real solution) is the development of a centralized National Statistical Bureau that would integrate in the same agency all of the general-purpose, public-serving statistical programs of the Federal Government. (This should not be taken to mean that the statistical programs specific to the other missions of operating agencies should not continue to be decentralized.) Indeed, the United States is one of the few, if not the only country that has such a decentralized statistical program—and it has been becoming progressively more decentralized. However, if one is sensitive to the political and bureaucratic milieu in which such a radical transformation would have to take place, one cannot help but feel that the changes of such a reform are very slim. Furthermore, one can at least make a case that the way for such a centralized bureau to come into being is through an emergent or evolutionary process tied to the solution of problems.

At the other extreme there are those who feel that each statistical agency should be supported separately in the attempt to extend these servicing capabilities along the lines internally.

The difficulty with the latter is that this is a problem that does not appear to lend itself to solution by agency incrementalism. There are thresholds that stand in the way and these thresholds are identified precisely with the necessity for interagency coordination in terms of both records and procedures.

As a consequence, my report on this subject (and in a similar spirit the more recent report of the Kaysen committee and the earlier report of Mr. Ruggles' committee) suggested a middle ground. We proposed a Federal statistical servicing center that would begin to arrange for the provision of more flexible user services. Such a center could make the investments in hardware and software system development that could serve the record retrieval and record association process in all of the agencies. It could request funds and allocate them according to systemwide standards and priorities to assure that archives are maintained and that appropriate file documentation and maintenance are provided for the key records. It could provide reference services for users with respect to all the governmental statistical programs. It could undertake file rearrangement, tape translation, record matching and cross-tabulation and standard statistical routines in a facility designed to provide a service capability for the users. Since much of the work that would be undertaken in individual agencies and programs, it should, above all, serve to evaluate and reflect the requirements of users in establishing priorities for the resources going to individual agencies in serving the needs. And it should guide the Bureau of the Budget and the agencies in the difficult task of establishing the new standards and procedures that will have to be incorporated into the basic collecting and tabulating activities. The object would be to provide an explicit instrument whereby the need for a flexible user servicing capability can be developed as a functional superstructure to traditional production oriented agency programs.

Now, let's consider the problem of personal privacy.

The statistical service center proposed has come generally to be referred to as a national data center. In retrospect I feel that this is an unfortunate appelation because it calls to people's minds the naive notion that the solution to this problem is simply bringing all the machine readable records from the Federal statistical agencies into a common repository. This, in turn, has raised in many people's minds the specter of a private dossier easily accessible to people who want to pry into their private lives. The whole matter has been caught up in considerable public controversy over the alleged threat that such a national data center would pose to personal privacy.

It has become associated in the public mind with such anxiety generating concepts as lie detection machines and electronic wiretapping. There is an emotional response that appears to be fed by a vague disquiet associated with the popular myths that have grown up concerning the omnipotence of modern electronic technology.

I, for one, do not feel that a concern with the issue of personal privacy is inappropriate. It is one of the important human values we must maintain and protect as best we can in a complex social environment. However, I am convinced that much of the concern that has attached itself to the proposal for a statistical service center is mis-

placed. This concern grossly overestimates the temptation to perversion that a genuine statistical system of the kind under discussion poses, and grossly underestimates the formidable protection that has and can be erected.

Much of the confusion about this issue rests upon a failure to differentiate between two different kinds of information systems—intelligence systems and statistical information systems. There is a basic difference in the purposes and, therefore, the organizations and functions that characterize them.

Intelligence systems generate data about individuals as individuals. They have their purpose "finding out" about the individual. In special purpose form they include such things as medical and educational records essential to the performance of many private and public functions. There is a great threat to privacy inherent in their assembly into personal dossiers to serve a general purpose intelligence function. This is certainly appropriately a matter of concern.

A statistical information system produces information that does not relate to the individual. It only identifies characteristics that relate to groups of individuals or so-called populations. It is concerned with generating aggregates and computing indexes, averages, percentages, et cetera, that describe the characteristics of and the relationships between groups of individuals. No information about the individual is generated as output and no information about the individual needs to be available to anyone outside the system under any circumstances for the statistical information system to perform its function.

You will notice that I have made the distinction on the basis of purpose, organization, and function. There is also characteristically some difference in record content that favors us, but this is not the primary distinction. Some people have attempted to seize upon the false hope that an effective statistical system can be built upon a file of population aggregates and, therefore, can be contrasted with an intelligence system by claiming that it will contain no individual records. Indeed, a part of a useful statistical system is made up of such aggregates. These are all that are essential for many important uses particularly in the management realm. However, if I have been successful in the first part of this presentation in communicating an understanding of the problem of statistical usage especially critical to policy and research, it should be plain that individual records are basic to the development of a flexible statistical system.

In view of this fact, we must directly face this question: Can a statistical information system be developed and administered in a way that asssures that it cannot be perverted for use as an intelligence system yielding outputs to inquiries concerning an individual? As a practical matter I think the answer is undeniably "Yes." This assertion rests upon two grounds.

First, there is the fact that major protections have and can be erected to protect the files from this kind of use. It has long been a traditional matter of social practice that, when we have something valuable, we put it in a safe place. One builds a trustworthy repository and then places his trust in it. We learned from the bank failures in the 1930's that there is no other option. A major part of our statistical system, notably the census, has operated for many years under strict legal and procedural strictures against the release of data that would disclose information about any individual respondent—be it a

person or an economic establishment. They have built a truly admirable record of success in meeting this obligation and the documentation of that record is worth serious consideration.

Both legal and technical protections can be strengthened and extended considerably beyond present practice. Indeed, employing the computer in the development of integrated file systems can make available for the first time certain protective devices and policies that are impossible outside the kind of reforms that we are advocating. For example, such reform will permit the use of internal control codes that can be made the object of special surveillance and it can make posssible a whole range of machine monitoring devices that can extend the range of human surveillance. One can argue that the capital costs and design problems associated with extending protections to personal privacy almost require an increase in system integration.

Second, my assertion rests upon the fact that there will be much less temptation to pervert a statistical system for intelligence purposes than is commonly supposed.

(*a*) The core of the record content of a general purpose statistical file consists of information concerning the *public face* of an individual—for example, obvious demographic characteristics—rather than his *private face*—criminal records, and so forth. Only a part of such a statistical record might represent information not available almost literally to public inspection.

(*b*) Statistical files are often loaded with information irrelevant for intelligence purposes—concerning deceased respondents, no longer existing enterprises, and so forth—that are important for statistical purposes but increase the inefficiency of the files for intelligence purposes.

(*c*) A general purpose intelligence system, if it is going to be efficient has to be as nearly complete as possible. Ideally it should constitute a census so that every possible individual search request could be fulfilled. Statistical systems do not contain "all of the data on everyone," to borrow a purple phrase, but only some data on some of the people. Even the census is largely a sampling instrument. It is not very tempting to run all the expense and risk of trying to violate a well-protected system to find out something about an individual if there is only one chance in four his record is even contained in the system.

(*d*) Most important of all, those who seek dossier type information have too many easy options available to them. We can rely on the market. No one is going to pay $500 for a suit if he can get the same suit at a lot less trouble for $50. Anyone who has any knowledge of informal intelligence sources and the formidable intelligence industry in this country can readily see that the temptation to pervert a properly protected statistical system would be practically nonexistent.

In conclusion, Mr. Chairman, we stand on the threshold of an era where relevant and timely information of special value to public policy and social research can be made available in ways undreamed of before. Given the rapid increases in the complexity of social life and the issues of public policy, this opportunity may be coming none too soon. We badly need to avail ourselves of these tools. In undertaking to provide them we must take seriously the necessity of protecting personal privacy. I am convinced that the provisions of this new capability can be made consistent with adequate protections for the privacy of the individual.

Chairman TALMADGE. Thank you for your excellent statement, Dr. Dunn.

My first question pertains to the archival function of the statistical system. You have said:

> The federal system has no policy and no systematic organizational or funding provisions to assure that basic statistical records will be maintained in appropriate archives.

My question is, Why does not the Bureau of the Budget order this to be done?

Mr. DUNN. I don't have any authority to speak for the Bureau of the Budget on that matter.

Chairman TALMADGE. Do they have the authority to accomplish it if they desire to do so?

Mr. DUNN. It is my impression that they do have the authority. In fairness to the Bureau we should bear in mind that it requires something more than an order from the Bureau of Statistical Standards—it would have to be implemented in accordance with budget requests of the various agencies and a set of standards for maintaining such archives that would be more effective if established upon a broader base of agency participation.

Chairman TALMADGE. Would you elaborate on the difficulties which stem from a lack of uniformity in defining the respondent unit?

Mr. DUNN. One of the common ones is the fact that a good deal of the data—for example, for business firms—is collected in the form of establishments. Some of it is collected from enterprises. Some of the data may pertain to an enterprise that is a corporation or a business partnership as a total enterprise and that partnership may be made up of several establishments located in various locations and engaged in different processes.

Now, if the data is collected on an establishment basis for one purpose in one agency, and other data is collected from the same enterprise but on an enterprise basis in another agency, the problem of trying to associate the enterprise data with the records for the establishment for the purposes of analysis is extremely difficult to accomplish. In many cases it is impossible.

Chairman TALMADGE. What types of analysis does this hinder?

Mr. DUNN. I would like to attempt to answer that, but, fortunately, we have here Dr. Ruggles who has engaged in some considerable effort in attempting to match these enterprise statistics with established statistics.

Chairman TALMADGE. Would you care to comment on it, Dr. Ruggles?

Mr. RUGGLES. Well, it is true that our aggregated statistics do not necessarily reflect the behavior of individual establishments and without being able to match the individual establishments we cannot create adequate statistical information on behavior.

For example, suppose we are interested in the behavior of wages in the economy. Our present statistics, as I point out in my own testimony, are built up by collecting man-hour figures from one set of establishments, and then collecting wage payments also, on an aggregated basis, and dividing the one into the other to obtain average hourly earnings. It is quite possible that the changes in average hourly earnings can come about by employing relatively more people in high-wage establishments than were previously employed, and thus raising

the average. In the aggregate statistics it will look as if the average is moving up. But, if you questioned each establishment individually, asking how much have the earnings of the employees in that establishment increased, you would have gotten quite a different answer. And this is true of productivity changes, and practically every change that takes place in the economy. We just don't have enough information on how these changes come about. If you are trying to achieve price stability, encourage economic growth, and so on, and you are trying to design legislation that will affect the behavior of the economic system, you have to understand how this behavior comes about. We don't know.

Chairman TALMADGE. Dr. Dunn, my next question, Are industrial classification systems similar throughout the Government?

Mr. DUNN. Yes, sir; they are.

I would like to interject here a qualifying statement. One of the difficulties in preparing a short statement is that, in order to communicate effectively in a short period of time, you have to leave out qualifying phrases.

I don't want to give the impression that the Federal statistical system is falling apart and isn't a useful system. It is a useful system. They have done many things over the years to improve the quality of statistics, and they have an admirable reputation in many respects.

We have such things as a Standard Industrial Classified Code or codes which are in effect and utilized by different agencies. But even here you can get into many kinds of problems, some of which we haven't really addressed the energy and resources to coping with. The descriptive phrases which are used to label these collection boxes are supposed to be standard for the different agencies, but there are some practical problems associated with the actual process of sorting the individual respondent's data out into these boxes. It is possible for one agency or one statistical program to wind up with a box with the same standard classification label as another, but each containing a different collection of respondents. Whether a respondent unit is sorted into a box labeled the electrical appliance industry will depend upon such criteria as whether or not at least 50 percent of its product output is of this type. Agencies may occasionally establish different cutoff points that determine which box will actually receive a given respondent unit. A second problem can arise from the fact that over time, the same agency may find that the product mix of an establishment may be changing in such a way that the same set of rules may place it in one box in 1960 and another box in 1970. The problems of going back and rearranging the contents of these boxes and ways in which they are fully comparable has never been coped with. It is only recently that we have come up against the problem of matching these records in ways that can serve more effectively policy and research and therefore come face to face with these issues.

Only a short while ago these kinds of exercises were so unfeasible as a practical matter that the set of standards which were appropriate for a statistical system of that era of technology just simply aren't appropriate in this era of technology. It is a lag process.

Chairman TALMADGE. Will you please elaborate on your statement "respondent units may carry similar tags which in fact mean different things?"

Mr. DUNN. Well, our discussion just completed really is an illustration of that—the fact that something called electrical appliance in BLS may carry the same tag as the electrical appliance industry in Census. But because of certain anamolies and the way in which, as a matter of fact information of respondent units are sorted out in these boxes you may wind up with a collection of respondents in the BLS box, even though it has the same tag on it, it is different in the collection of respondents in the census box with the same tag on it.

Chairman TALMADGE. To what extent do State and local data systems mesh into the Federal systems?

Mr. DUNN. They don't.

Chairman TALMADGE. Are they completely different?

Mr. DUNN. There are times when you can get data that is in a condition which has sufficient quality that it will allow merging in ways that are useful, but these are rather rare instances. Usually, the data that is useful come out of programs that are associated with Federal programs.

The Division of Regional Economic Analysis in the Office of Business Economics makes some economic estimates for individual counties for intercensal years. They make use of tabulations which they gain from the States which are administered in the unemployment insurance programs in the States. These are State agencies, and the data comes from the State agencies, but their programs are cooperative programs under Federal supervision.

There are many problems associated with using these data in the estimating process because of their source and the way in which they are generated, but they have been on occasion used successfully. This was under the most favorable circumstances where what was generated was under a program tied in with a federally-sponsored program, so a cross-State consistency could be established. More commonly, State generated data cannot be effectively compared for interstate analysis. My understanding is that State Governors, controllers, and policymakers have had great difficulty in analyzing such things as the State expenditures for State services and make justifiable comparisons between what they were doing in their States and what people were doing in other States because of the great difficulty in developing statistics that have any comparability.

Chairman TALMADGE. We have a rule in the committee where each member is allowed 10 minutes. My 10 minutes has expired. I have some further questions, but I will ask Mr. Knowles to give them to you and you can supply the answers for the record.

(Additional responses supplied for the record of testimony of Edgar S. Dunn, Jr., appear below:)

I.1 What would be the cost of a Federal statistical servicing center which would significantly improve our system?

Answer: This is an extremely difficult question to answer satisfactorily at this stage simply because the staff work essential to the specification and costing of program options has not been done. When asked a similar question by the Bureau of the Budget I provided the judgment that the range of services and program adjustments required would call for an expenditure of between 1 and 2 million dollars annually during the initial years and rising to the neighborhood of 10 million annually over a period of 5 to 10 years. I think that this is as good a judgment as one can give until more staff resources are brought to the planning task.

II.1 At the present time, does the Government have uniform policies on matters of (a) disclosure and (b) personal privacy?

Answer: No, sir. Not for all of the programs that might be characterized as part of the Federal Statistical System. Some programs and agencies operate under very stringent legal and administrative controls forbidding the disclosure of information and the violation of personal privacy. This is particularly true of an agency like the Census that collects information under a legal mandate that requires the respondent to comply with the request for information. These controls have been rigorously enforced and, in the process, a considerable body of techniques and experience has been accumulated concerning the protection of personal privacy. Other bodies of data, especially those that come into being as the byproduct of administrative activities and programs are subject only to administrative protections and these often vary considerably in stringency.

II.2 Is it likely that a statistical service center might reduce the risks to personal privacy and to firm disclosure, compared with present arrangements?

Answer: There is no doubt in my mind but that it would, if we do a proper job of building in all of the protections that are available to such a system. There is no question that both legal and technical protections can be strengthened considerably beyond present practice. As a matter of fact, it seems likely that the necessity for more universal legal controls combined with the capital costs and system design problems associated with extending protections to personal privacy require an increase in system integration.

II.3 Would you distinguish between the short run and the long run problems of safeguarding personal privacy in the event that the government establishes a statistical service center?

Answer: Yes, I would, for the following reasons. A flexible user-servicing capability of the kind that I have been talking about cannot possibly come into existence except over a period of some years. As I tried to indicate, there is more involved than simply setting up a centralized computer and bringing together indiscriminately all the tapes you can get your hand on. To perform the statistical services that are important to policy and social science research one needs to reach back into the primary agencies to modify collection and tabulation practices through the establishment of new standards and new procedures. We not only need the capability to reshuffle the deck, we need to reshape the deck over time to a form that can be shuffled in this way. This is a big job. We will have to make a beginning with particularly valuable subsets of the existing records and work to enlarge and generalize our capability on an incremental basis over time. I cannot see that there is any way of escaping this.

This has two implications from the point of view of the personal privacy issue that makes a distinction between long run and short run problems useful. First, as long as we are pursuing goals of developing a statistical user service, it will be some time before we will have restructured a large enough portion of the Federal Statistical System so that there would be emerging anything like a dossier record with a set of attributes of wide enough scope to be of special interest as an intelligence source. We have a substantial period of grace before that aspect of the personal privacy issue would assume the threatening proportions that the scare reaction presumes. Second, some of the most effective devices that we can develop to protect personal privacy are going to require some experimentation and system design work that can be undertaken in connection with the implementation of the first phases of generalizing the capabilities of the Federal Statistical System. This suggests to me the conclusion that the notion of the threat to personal privacy is especially exaggerated with respect to the short run and that our beginning efforts in the short run will provide a valuable testing ground for developing the protections essential to an efficient, flexible, general-purpose system.

Chairman TALMADGE. Congressman Bolling?

Representative BOLLING. I would like to pursue a little further this proposal that apparently has been made by virtually all who have studied this matter with a background of real knowledge, if I may use a purple phrase.

You have all proposed a Federal statistical center, a data center, or something along that line. I notice later on in the statement that you indicated that it should guide the Bureau of the Budget and

the agencies in the difficult task of establishing the new standards and procedures and so on and so on. The object would be to provide an explicit instrument whereby the need for a flexible user servicing capability can be developed as a functional superstructure to traditional production-oriented agency programs.

I will have to confess that if I read those reports when they were issued I haven't them in my mind and this may have been covered in the reports. But where are you going to put this agency? Where in the structure of government do we put it so that it avoids the dilemma I remember confronting when as an original member of this subcommittee a good number of years ago we were trying to break through the morass of agency individualism, agency provincialism, agency parochialism, while we were trying to develop a more centralized, but not too centralized, operation. I am curious as to where you put a Federal statistical systems center. Outside the President's door?

Mr. DUNN. You mean as a matter of governmental organization?

Representative BOLLING. Yes, sir.

Mr. DUNN. I wouldn't presume to make any kind of judgment or recommendation on this subject, because I am not an expert on government organization and I am not sensitive to all the problems of political and bureaucratic accommodation that would be essential in setting up the program of this kind. There are many different kinds of organizational options that are available.

Representative BOLLING. Suggest some.

Mr. DUNN. One can set up a statistical servicing center as an independent agency, one could put it as one of the functions of an existing agency such as the Bureau of the Census, one could take the census out of Commerce and add it to the census function, and it was even suggested at one time—it is not an idea that appeals to me—that the Office of Emergency Planning be considered on the grounds that it has a large computer and a good deal of information addressed to the problem of evaluating post-attack impacts and security and has had a fair amount of experience in assembling large amounts of data into a computer organized system. There are, I am sure, many options.

I wouldn't presume to offer a judgment about what is appropriate in this case. I think that is a political and administrative decision here which isn't a part of my function.

Representative BOLLING. Dr. Ruggles, would you care to comment on that?

Mr. RUGGLES. Well, I think this is obviously an interagency function. There are examples of a number of other interagency functions which now exist—the National Archives carries out such a function dealing with a large number of agencies; the Government Printing Office is another one; General Services Administration is another one. I think we have to realize that there are interagency functions to be performed. Whether you set these functions up in an independent agency, or whether you give some existing agency additional authority, I think is a matter of bureaucratic convenience and so on. But I do feel that an interagency authority of some sort is needed here, similar to that of Archives or the Government Printing Office.

Representative BOLLING. Do you think that any of the existing interagency operating agencies would be the appropriate place for this additional function?

Mr. RUGGLES. I would feel that the major difficulty is one of technical know-how. If you are going to engage in an operation of this sort you want to bring to bear on it the best talent that is already available within the Government. Merely to use an existing Government organization is not as important as it is to get the best talent available.

Therefore, I would feel that you have to use some part of the existing statistical system to essentially support the new effort. This would be my inclination.

Representative BOLLING. If I understand correctly, the pool of adequate talent is relatively limited, is that correct, both within and without the Government? Is this not a field in which there is a shortage of real experts?

Mr. RUGGLES. That is true, but it is quite possible that by creating this organization you would economize on the amount of talent needed.

In other words, as it now is constructed, the Federal statistical system is large, sprawling, and highly decentralized. To the extent that a reorganization would introduce more efficiency, the existing resources could be better utilized.

Representative BOLLING. I do not want to pursue this point too far, but I am very much interested in a result, not only of the hearing, but I happen to feel that this kind of thing is perhaps a little overdue. What would be wrong with putting whatever you are going to call it, the servicing center, under the Bureau of the Budget which already has an interagency coordinating authority in this area?

Mr. DUNN. Are you addressing it to me?

Representative BOLLING. To each of you—despite your denials of any competence in the field of organization. I think this is a problem that was one of those difficult ones where it is political and bureaucratic, and it is organizational, but unless it is made very wisely in terms of the techniques involved it could be a disastrous political and organizational decision. I apologize for pursuing it.

Mr. DUNN. My impression is that the Bureau of the Budget would present some special difficulties because in effect it would be giving the Bureau a very substantial operating function when it is not really organized to form operating functions and doesn't really conceptualize its mission in this way. As a matter of fact, I think that some of the difficulties which have beset the Office of Statistics and Standards over the years in performing its functions is because it has had a fairly substantial operating function in a nonoperating agency.

Mr. RUGGLES. I think I would concur in this, and I might add one comment. I think the reasoning behind this is quite sound in that it would be useful to have such an interagency in the Executive Office of the President, and the Bureau of the Budget is in the Executive Office. However, I think that to attach the National Data Center to the Bureau of the Budget would so distort the Bureau of the Budget in terms of operational activity that it would be inappropriate. But there are other agencies in the Executive Office. It might be useful to take the Office of Statistical Standards out of the Bureau of the Budget, move it into a National Data Center, together with some parts of the present Federal statistical system that have both the computer capability and the operational data handling capability. The merger of those elements might be very useful.

Representative BOLLING. Thank you, gentlemen. My time is up.

Chairman TALMADGE. Senator Miller?

Senator MILLER. I take it from the conversation that while this would be a National Data Center, that it would be for the primary use of the executive branch of Government, is that right?

Mr. DUNN. No; I would not think it would be fair to presume it would be limited to Federal agency service any more than the Federal statistical system is currently limited to Federal agency service. This is one of the largest and dominant parts of the user community, but many of these files—particularly those that were established on the basis of legislative authority to provide a general public service function like the census—provide statistical outputs for everyone. As a matter of fact, their principle source of general output is the publication of monographs which are circulated throughout public repositories of all kinds.

Senator MILLER. Do we wish this primarily for the assistance in decisionmaking by the lawmakers or primarily for decisionmaking on the part of the administrators, or primarily for decisionmaking on the part of the private sector.

Mr. DUNN. I think all three. I think you are going to have representatives of the Federal Statistics Users' Conference in a day or two and they will testify as to the value of the Federal statistics programs and what they have meant over the years to private decisionmaking. There are certain kinds of information that private decisionmakers would have had absolutely no means of accessing were it not for certain standard Federal statistical programs.

I would also like to emphasize that I think the Nation has a very great stake in improvements in the availability of information of this sort, because of the contribution it can make over the years for the success of social science research. I think many of our most serious public problems these days are getting to a realm of policy having to do with health, education, poverty, and so forth, where the domain of the problem is not very well known and many of the aspects of the problem are not very well understood. We need to research these things. A lot of this is going to have to be undertaken by the social science research community.

Senator MILLER. I am sure they use some of these facilities at the Library of Congress now, do they not?

Mr. DUNN. Yes; they do.

Senator MILLER. Would there be any objection to having this Center under the cognizance of the Library of Congress?

Mr. DUNN. This is one of the organizational options which has some possibilities.

Senator MILLER. Following Congressman Bolling's question, if the Bureau of the Budget would not appear to be the most suitable agency for this, what about the Council of Economic Advisers?

Mr. DUNN. Here again, if you added this function to the Council of Economic Advisers you would completely transform the character of the function. They are quite different functions. The Council is a very, very small organization made up of professional people who spend all of their time providing advice to the President on current problems. They make extensive use of information and certainly their efficiency would be substantially improved if they had more effective access to data and had more flexibility in arranging and combining data. But this kind of statistical servicing has been in no way any

part of their previous mission. If you were to associate these two you would find they wouldn't have very much functionally or operationally in common. I would think the size and scope and operational problems of the latter would either swamp the functions of the Council of Economic Advisers, or else they would just continue to operate independently.

Senator MILLER. Somebody is going to be swamped. They are going to take on something that they have not had before. The Council of Economic Advisers is not exclusively advising the President. Here is a publication—the Economic Indicators, issued monthly. This is not prepared for the President. It is prepared for the Congress.

Mr. DUNN. True.

Senator MILLER. It seems to me one of the vital functions of the Data Center would be to enable Congress to have data readily available which would take less time to legislate in a more timely fashion and in a more responsive fashion than we are doing now. So, I am not troubled by the swamping results. Somebody is going to get that. But somebody is going to have to make a decision where this is going to be located and the Council of Economic Advisers might be a possible source for this.

One thing that I am wondering is whether or not we are troubled more by the lack of timeliness of the data or more by the lack of responsiveness of the data we are getting. What I am getting at, for example, is this: Are you familiar with the report of this Subcommittee on Economic Statistics of the Joint Economic Committee?

Mr. DUNN. I have not read it with care. I have seen it.

Senator MILLER. I think a reading of that report indicates that we were not troubled by the lack of rapidity as by the lack of responsiveness of the data. We were not getting figures on underemployment and we were not getting figures which really reveal how much certain unemployment was and in what sectors of society it was. We were not troubled about getting the data. We were not getting the right data. I am wondering if we are putting the emphasis on the wrong place here, because this Data Center is not going to be worth anything if we do not have the kind of input needed, which is our major problem here.

Mr. DUNN. The emphasis on these reports has been clearly on the problem of the appropriateness of the data used rather than timeliness and the ability to have flexibility and record-matching association, so you can bring records together in ways that would illuminate the relationships between economic and social phenomenon.

I think that, for certain problems of public policy, timeliness is not as important an attribute. It happens to be the one we have become most exercised about earliest in our history. We have placed great emphasis on timeliness because of our concern with business-cycle phenomena, national income, and so forth. We have been concerned about what the economy is doing currently during this quarter and this month. This has placed great emphasis on timeliness and a substantial amount of resources, interest, and attention has gone into making the data just as current and timely as it can be for these purposes. But for the vast bulk of the other kinds of policy problems for which information is critical, the issue of timeliness is nowhere near as important a problem as the ability to associate records and the ability to relate records over historical timespan in the way that will establish trends and give it some notion as to the way changes are taking place. It is these uses and requirements that have been most poorly satisfied and it

is in the interest of providing more appropriately for these requirements that such recommendations as were made earlier have been made.

Senator MILLER. I could not help but think when you were talking about this data center in answer to questions of the chairman, that it seems to me there are many types of statistics that are peculiar only to one agency. For example, the Department of Agriculture, the Post Office Department, Defense Department—statistics that are vital to those agencies but are really of no interest to anybody else. We are going to have to make a pretty careful survey to determine what types of data would be handled by the data center and what types would be left exclusively under the control of the agency, would we not?

Mr. DUNN. Yes, sir. One thing should be clear at the outset. As the various committees have considered this problem and as the Bureau of the Budget has pondered these issues initially, there has never been any thought that a centralized statistical system of either limited or general type would undertake to centralize all of the statistical functions of the Federal Government. Many important statistical programs are agency-specific—that is, they are central to the operating missions of the agency. Certainly these programs belong in those agencies under agency control. The idea here is that those statistical programs and those statistical resources from the agencies that could be made to serve a general-purpose, public-servicing function would be the ones that would be subject to this kind of reorganization and reform.

Senator MILLER. My time is up. I have just one more question, Mr. Chairman, if I may?

Chairman TALMADGE. Proceed.

Senator MILLER. I am wondering if, before we established a national data center, with the ramifications which appear possible here, it might not be a good idea to first have a pilot project. What I have in mind is this: The Air Force for a number of years has had computer programs. About 3 or 4 years ago they decided they ought to have a sort of service center where experimental types of use agencies would be the function. I believe they set this up in Dayton, Ohio. I think it has had good success. It has not served to centralize the data; however, it has served to run experimental types of programs with a view to improving the system—new uses, extrapolations, for example.

I wonder if we might not start out, let us say, in a experimental type project to see what its potential would be before we get too deeply involved in something like this. Do you not think this may be a good approach?

Mr. DUNN. Well, I think it has to be a good approach because I think it is really, in the end, the only approach. Let me explain what I mean by that.

I think there has developed a false impression that somehow or other what is proposed, and, indeed, what is possible is to establish overnight something that would represent a complete, de facto, operational national data center. This simply cannot be done. In my testimony I tried to indicate that in order to bring about the kinds of reforms that would really make possible record matching, for analytical purposes of the kind that we are talking about, the changes would have to take place would reach all the way back into the statistical collecting and tabulating procedures of the agencies—establishment of standards and things of this kind. These things don't happen over-

night and no matter how it started organizationally, with what kinds of authorities, with what kinds of ultimate objectives, it is going to take a considerable period of years for a capability of this kind to come into being and we are going to have to make a beginning with certain segments of the file which we feel are most critical to public policy requirements and which will have the biggest payoff. We will need to generate experience with those records and expand the system as we go along. There won't be any option to proceeding this way. I would take exception to the pilot program notion only in this respect. It seems to me, to make an effective beginning on this thing, we've got to do something more than set off some kind of disassociated pilot program in the corner and say, "Let's run it for 5 years and come back and take a look at it." There is a certain minimum threshold of getting started and that minimum threshold probably represents setting up an organization which has some kind of interagency authority and funding and staffing.

We will need this to even make this kind of beginning, to get into the problem.

Senator MILLER. Thank you, Mr. Chairman.

Chairman TALMADGE. Thank you, Senator Miller.

Our next witness is Dr. Richard Ruggles, of Yale University.

You may proceed as you see fit.

STATEMENT OF DR. RICHARD RUGGLES, PROFESSOR OF ECONOMICS, YALE UNIVERSITY

Mr. RUGGLES. During the past 2 years a heated controversy has developed over the future of the Federal statistical system. A proposal for a national data center which would pool data from many government agencies has been labeled as an "invasion of privacy" and even a "threat to democratic institutions." To put this controversy into its proper perspective, there are certain directly related questions which should be considered. First, was the data base in the past adequate for the needs which it was intended to serve? Second, what has been the impact of the computer on the processing and use of data, and how has it affected the existing situation? Third, how is the present structure of the Federal statistical system related to its functioning and to the adequacy of the data base? On the basis of the answers to these questions, it can then be asked what steps, if any, should be taken now to assure that the data system will be capable of meeting future needs fully and efficiently, without invading personal privacy.

THE ADEQUACY OF THE DATA BASE

The Federal statistical system as it now exists evolved in response to the changing forces in our economy. The first requirement, specified in the Constitution, was for a head count, which had as its major purpose the determination of representation in the Congress. The demographic census has now become one of the basic sources of economic and social data, describing not only the geographic distribution of people, but summarizing their economic and social characteristics. Most nations today consider the demographic and industrial censuses as the backbone of their statistical systems. As the economy developed, other kinds of data were created as a byproduct. Customs data yielded

information on international trade. The establishment of the banking system brought with it information on money and credit. The attempt to solve agricultural problems led to a large mass of data on prices and production of farm products. In the field of labor, price indexes of the cost of living and information on wages and hours were collected and published. The depression of the 1930's led to measurement of unemployment. The regulatory activities of the Government also produced a flood of data. Railroads, public utilities, security markets, and interstate commerce all were required to provide information for the public record, and to produce data for the purpose of regulation.

The establishment of the corporate and personal income taxes and the social security system put in the hands of the Government detailed records on almost every individual and business in the Nation, and although these were not made public they have been used for administrative purposes and for the development of statistical data. In the fields of health, education and welfare, large quantities of data have also been amassed. Information about births and deaths, diseases, education, and crime have been collected from a wide variety of State and local institutions.

The ever-growing body of data has found a great many uses, but it has also brought with it a great many headaches. Masses of different tabulations of interrelated but not necessarily consistent data have been published by a large number of different agencies. In trying to summarize the quantities of detail, economists first went in the direction of developing indicators which would show the level of activity in the economy. Such diverse things as steel production, freight car loadings, the wholesale price index, and stock market prices were used as general measures of what was taking place. After the great depression of the 1930's and especially in response to the needs of World War II, an attempt was made to develop a more systematic and integrated set of information for all parts of the economy. The national income accounts, initiated in the late 1930's, have now developed so that they draw on a large number of statistical sources and provide a set of internally consistent estimates of all sectors of economic activity.

However, national economic accounting is still highly aggregative. It manages to reduce the picture of the total economy to a manageable set of tables only by omitting the underlying detail and interrelationships. It is not generally possible, furthermore, to move from the national economic accounts to the more detailed tabulations on particular aspects of the economy provided by the different Government agencies. Besides the intellectual problems of comparability and consistency, the sheer mass and wide diversity of the underlying data restrict its usefulness for general statistical purposes, although, of course, to those concerned with one part of the system individual tabulations by themselves are very useful.

The first major revolution in the processing of data was the development of punchcard machines. Punchcards were useful not only because they afforded a mechanical means of handling and tabulating data, but because they forced upon the system a degree of rigor in the development of codes and classification systems. The use of punchcards on any large scale, however, introduces an unavoidable element of inflexibility. For large operations, the coding, keypunching, and tabulation have to be rigorously supervised, and scheduled in advance from the first recording of the data to final tabulation. Generally speaking, feed-

back from the initial phases of such an operation cannot be used to alter and improve later phases. The final tables are essentially cross-tabulations at a highly aggregated level. If the statistical work is repeated periodically, a time series results.

Attempts were seldom made to match a reporting unit in one period to the same reporting unit in later periods. Comparability over time was seriously impaired, since the tabulations reflected differences in classification procedures, definitions, and coverage.

Matching of data from one set of information to another was really not feasible. The Census Bureau points out that even with the population census arranged in the best possible form, the cost of matching information for a specific individual, to be used as proof of citizenship, was $4 to $5 per match, and that the match was unsuccessful on average 15 percent of the time. Matching of different sets of data either over time or between Government agencies was seldom even attempted. When a Government agency wished to use information which was contained in existing sets of data processed for another purpose, it was generally cheaper to completely redo the work rather than try to use the existing data.

From the point of view of the social scientist, the increasing flood of information has been both welcome and disheartening. Data is as important to the social scientist as laboratories are to the scientist or libraries to the humanist. The social scientist has become increasingly ware that progress in his discipline is closely tied to his ability to analyze, explain, and understand the empirical information on the behavior of the economic and social system. But the information available has tended to swamp empirical research workers. Until recently the cost of data handling and processing on any significant scale has been prohibitive.

The lack of consistency between sets of data and the lack of comparability of classification systems further complicated the work. Even where suitably disaggregated data existed, the individual scholar rarely could gain access to it. As a result, the economist has generally taken refuge in macroeconomic data such as the national accounts which are manageable and are presumably comparable and consistent. Unfortunately, the use of macroeconomic models has methodological disadvantages and limitations. But until very recently the social scientist wishing to engage in empirical research had little alternative.

In view of these considerations, it does not seem relevant to criticize the adequancy of the data base. The basic limitation in the past has been the cost of processing and handling of data, which restricted users to a partial and fragmentary basis. Even the imperfect and incomplete statistical system yielded greater amounts of data than could be efficiently used. The problem was not a lack of data, but rather the inability to use efficiently all the pieces which did exist.

THE IMPACT OF THE COMPUTER

With the introduction of the computer, a new set of forces was set in motion. Starting in the 1950's, the Bureau of the Census pioneered in the use of electronic equipment for data processing. UNIVAC I, now in the Smithsonian Institution, was a monumental step forward, although it was only a modest beginning of what turned out to be a completely new technology.

Each succeeding generation of computers has incorporated improvements in the size of memory, the speed of computation, and the density of data storage such that the capacity and speed of operation have been increased many times over. By now the technological revolution has become so significant that a reexamination of the situation is urgently needed.

From the outset, the computer, like other forms of automation, reduced the amount of labor required in the processing of data. This fact in itself is not remarkable, but it is the magnitude of this reduction that is revolutionary. Data processing which would have required hundreds of thousands of man-hours to carry out 15 years ago can now be carried out with much greater efficiency in a matter of a few minutes. Data processing tasks which would never even have been considered because of the magnitude of the cost involved now can be done cheaply and easily. Equally important is the fact that the timelag between the input of information and the final output has been substantially reduced. Operations which took 8 months to a year to complete with a small army of clerks can now be reduced to a matter of days or weeks with relatively few people. This shortening of time has not only meant an increase in efficiency in terms of overhead and other fixed elements, but it has also resulted in making important information available more promptly.

The computer has also made possible new kinds of analysis. Editing instructions to test the reasonableness of the basic information can be built into the processing programs. It has become possible to examine and edit data much more carefully. Computers can "wash" the information, and find inconsistencies which would have gone unnoticed in hand editing. For some Federal agencies the ability of the computer to make consistency tests is very important. Thus, the Internal Revenue Service currently uses computers to check the internal consistency of items contained in each individual tax form. Such an operation is basic to the major administrative function of this agency, but before the introduction of the computer it was too expensive and time consuming to be feasible except for a very small number of cases. In such uses the computer is adding a new dimension to the work and increasing the overall efficiency of the agency.

The problem of data storage and access has also been greatly simplified. Before the development of the computer the storage of original data was difficult. The Internal Revenue Service had warehouses full of bales of tax forms tied in packages. The Census Bureau had every available corridor lined with cases of punchcards, and many punchcards had to be destroyed merely because there was no physical room to keep them. Computer tapes have radically changed this situation by dramatically reducing the storage space required. This reduction in required storage space has at the same time resulted in greater accessibility of the data. As long as data were in bales or even in punchcard form, the cost of access was so great that for most purposes it was not feasible to use the data. Now, however, it costs relatively little to put a tape on the computer and to use it as desired.

The existence of large sets of relatively accessible data on computer tape makes matching operations more feasible. The usefulness of matching operations has always been apparent, but the prohibitive cost made them impractical. This last March the Census Bureau started publishing in its "Data Access Descriptions" a report (MS–1) on

"Matching Studies." This report describes methods by which census records can be matched for selected groups of persons. The confidentiality of replies is preserved in the matching operation, since the organization requesting a matching study for a list of individuals receives from the Census Bureau only statistical tables summarizing the characteristics of persons on the list. For example, three major matching studies relating to mortality, mental illness, and juvenile delinquency were conducted after the 1960 census. Samples were drawn respectively from death records, hospital admission records, and court records. The geographic basis for the matching work was in one case the entire Nation, in another case two States, and in another a populous county. In each study the investigator had a listing of the events in which he was interested; that is, deaths, hospital admissions, and cases of juvenile delinquency which reached the courts, which occurred in a time period appropriate for census matching.

The mortality study was concerned with variations in the proportions of deaths by a broad range of social and economic conditions within specific age and color categories. The mental illness and juvenile delinquency studies were concerned primarily with the family characteristics of the subject populations compared with the characteristics of the population at large. These studies produced results considered useful by their planners, but they were very expensive—although not as expensive as new fieldwork—and only 70 to 85 percent of the cases were matched.

In the case of a sample of deceased persons, a matching study may offer the only hope of obtaining additional data needed. In the case of mentally ill persons or delinquent juveniles, fieldwork is possible, but a matching study may be a more reliable way of getting data. In the 1970 census, an attempt will be made to build the file of data in a manner such that matching studies can be carried out more economically and more accurately. This will be done by matching the address and general characteristics of the household, such as age, sex, color, and marital status.

The possibility of matching data from individuals, establishments, and other reporting units makes it possible to interrelate different bodies of data in a meaningful and useful manner. Studies of individual behavior over time can be made. For example, it has been customary to measure average hourly earnings for an industry by adding up total man-hours reported by all establishments and the total wage bill reported by all establishments, and dividing these two aggregate figures. Changes in average hourly earnings over time were measured by comparing the average hourly earnings in one period with that of other periods. The observed change measured in this way, however, could come about in two ways. Employment of workers in high-wage establishments may change relative to that in low-wage establishments, thus altering the average of the industry. Alternatively, average hourly earnings may change in each establishment. Without a matching study, the behavior of individual establishments could not be ascertained. But an understanding of the factors affecting wage behavior is essential if we are to hope to achieve the goal of reasonable price stability.

In similar manner, there are questions as to how productivity change occurs in the economic system, how consumers save, how businesses invest. Such questions as these require statistical examination of indi-

vidual behavior, and for this purpose it is necessary to observe the same individual reporting unit over a period of time.

The increased accessibility of data via computer means also that existing data which was originally developed for one purpose, or which is a byproduct of administrative procedures, can be used for a wide variety of other purposes by reclassifying it or reprocessing it. Many public programs dealing with urban renewal, poverty, and medical care could use already existing data for the design and implementation of policy. This raises the problem of disclosure of confidential information. Given the computer, however, it is now possible to process the data in such a form that useful results are produced without disclosure of any individual information.

Social scientists engaged in empirical research in academic institutions also have much to gain from the resources of the computer and the availability of microdata. The technique of using highly disaggregated data for economic analysis is rapidly developing. For example, Joseph Pechman, of Brookings, used a sample of 100,000 individual tax returns to analyze the impact of alternative changes in the tax law on individual taxpayers and on total tax revenue.

For this type of analysis it was merely necessary to program the computer to recompute each tax return according to proposed changes in the tax law, and to compare such a computation with alternative proposals. The Bureau of the Census has put in the hands of social scientists another extremely valuable set of data. On the basis of the 1960 Demographic Census they constructed a 1-in-1,000 sample which provided information on the economic and social characteristics of households. This set of data has been used in a variety of ways, including, for example, a simulation of a life process model to determine the characteristics of the future retired population and their probable income level. The ability of social scientists to obtain highly disaggregated data permits them to use techniques of analysis which are inherently much more powerful and can separate out the structural changes of the system from the changes in behavior of individual units.

Thus, the revolution which the computer has caused in the processing and analysis of data has completely altered the point of view of both Government agencies and research scholars with respect to the nature and adequacy of the data base, and the pertinent question which remains is whether the present organization of statistical activities is consistent with the changed conditions.

THE ORGANIZATION OF THE FEDERAL STATISTICAL SYSTEM

From this brief discussion of the evolution of the Federal statisticial system it is obvious that the term "decentralized," while applicable, may be somewhat misleading. Decentralization could come about through a conscious splitting up of responsibility to provide a division of labor.

The decentralization of the Federal statistical system, however, does not represent such rationalization; rather it has been the result of a jungle-like growth of statistical activities by different Government agencies having widely different purposes. The result has been extensive duplication and lack of coordination. In view of this, the Hoover Commission in 1949 recommended the establishment of the

Office of Statistical Standards in the Bureau of the Budget, to produce a higher degree of integration and centralization. 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.

In 1959, the Social Science Research Council appointed a committee to study the problems of the preservation of data and access to data in the different agencies of the Federal Government. After considerable study of the statistical activities of 20 Federal agencies producing some 600 bodies of data represented by 100 million punchcards and 30,000 computer tapes, the committee came to the conclusion that the present organization of statistical information in the Federal Government resulted in loss of important data, inaccessibility of information, and excessive duplication. Despite the use of computers and punchcards, Government agencies were still primarily geared to producing statistical tabulations.

Because most agencies had operating functions, they did not consider the preservation of data to be an important function, and they were unable to service the legitimate requests for information by other Government agencies or by outside users. Although disclosure sometimes represented a proper obstacle to the interchange of information, it was often used as a cloak for inefficiency and as a device for avoiding the performing of service functions for others. Even internally, many Government agencies were careless in the preparation and preservations of their own statistical materials. The basic computer tapes were often left in an unedited and uncorrected state, so that they could not be used over again, and documentation concerning the tapes was almost always inadequate. Because each agency was concerned with the development of its own data without respect to the activities of other Government agencies, there was excessive duplication in the procurement of information from respondents. The same reporting unit was required to provide approximately the same information to a number of different Government agencies. In brief, although the decentralization of the Federal statistical system was acceptable in the precomputer period in a situation where the effective limits on the use of data were imposed by difficulties of processing and handling of data, this same decentralization poses serious problems of coordination and integration given the data processing potential of the new computer technology.

PROPOSALS FOR THE COORDINATION AND INTEGRATION OF GOVERNMENT STATISTICAL PROGRAMS

In view of these findings, the Social Science Research Council's Committee of the Preservation and Use of Economic Data recommended the establishment of a Federal data center which would have the authority to obtain copies of computer tapes and other machine readable data produced by all Federal agencies. Such a Federal data center would have the functions of providing data and service facilities, so that within the proper safeguards concerning the disclosure of information both Federal agencies and users outside the Government would have access to basic data. The Bureau of the Budget, in response to these recommendations, undertook its own evaluation of the desirability of such a facility. A report prepared by Edgar S. Dunn, of

Resources for the Future, consultant to the Bureau of the Budget, concluded that it would be desirable to establish a national data service center which would direct the file storage and management of significant archival records in machine readable form for all participating agencies, and provide a central reference source and explicit facilitating services for all users of Federal statistics.

After the Dunn report was issued, a task force under the chairmanship of Carl Kaysen, of the Institute for Advanced Studies, was appointed to study the storage of and access to Government statistics. The task force studied the question over a period of a year, and then proposed the creation of a national data center which would be given the responsibility for assembling in a single facility all large-scale systematic bodies of demographic, economic, and social data generated by the present data-collection or administrative processes of the Federal Government.

It would be the function of the national data center to integrate the data in such a way as to preserve as much as possible of the original information content of the whole body of records, providing ready access to the information within the laws governing disclosure to all users in the Government and where appropriate to qualified users outside the Government on suitably compensatory terms. The task force emphasized the necessity of developing safeguards to preserve the right of the individual to privacy in relation to information he discloses to the Government either voluntarily or under legal compulsion.

During the period when the desirability of a national data center was under discussion, hearings were held by a special Subcommittee on the Invasion of Privacy of the House Committee on Government Operations. A number of witnesses testified that it was their belief that a national data center would be a serious threat to individual privacy. Much of the discussion centered about the question of building up dossiers containing improper, irrelevant, but harmful information which could be used to the disadvantage of individuals. The view was expressed that the establishment of a national data center would encourage the collection and preservation of such information.

It cannot be denied that in some respects there has been a significant erosion of privacy over the past 50 years. The introduction of the income tax perhaps represents one of the largest intrusions, since a person must reveal all of his sources of income to the Internal Revenue Service.

The social security system, also, has made the employment history of an individual a part of administrative records. Such records such as motor vehicle licenses, driver's licenses, medical care, and public assistance records have further reduced privacy. In the past the operations of certain congressional investigations and the knowledge of the existence of security files kept by various Government agencies have led to further uneasiness. It is therefore not remarkable that the average person views with some alarm an apparent plan to centralize all records.

This uneasiness reflects, however, more a fear of possible misuse of information than an objection to its existence. For example, few individuals would mind that the Social Security Administration has the administrative information it needs if they were confident that no one else would have access to it. The Census Bureau has long realized that they could not expect to get accurate information if those provid-

ing the information feel that other Government agencies might have access to such information and use it against them. For this reason the Census Bureau has had written into the law confidentiality restrictions which prevent the disclosure of census information to any other parts of the Government or to private users. Thus, for example, the Internal Revenue Service cannot get even the names and addresses of those included in a census. Unfortunately, not all Government agencies are so circumspect concerning the information they obtain from individuals, and State and local governments provide even less protection of information.

Recently, for instance, the New York Times cited the case of the New York State government selling the names and addresses of 6,400,000 motor vehicle owners to marketing services for a sum of $86,000. There are many examples of other abuses by all levels of government, but perhaps what concerns individuals most is the existence or use of secret files about them within government, containing unverified and often erroneous information.

Although the emphasis in the privacy hearings was mainly on the possible danger of centralizing records, they also brought out that in some instances the centralization of files can result in increasing the protection of individual privacy in situations where there have been flagrant abuses. For example, New York State is currently setting up a central identification and intelligence system. Before the establishment of the centralized system, there were some 70 million files in the various agencies of criminal justice in New York. These related to police departments, prosecutors, criminal courts, and probation, correction, and parole agencies, all of whom dealt with individuals who came within the jurisdiction of the law. They include of course local agencies as well as those of the State government. In all, some 3,600 agencies were involved, of which over 600 were police department. Under the decentralized system of duplicate files, the cost of maintaining the files was very great, and there was for the most part no agreement as to the kind of evidence which it was proper to maintain in the files.

In many cases useful information could not be brought to bear upon a pressing problem. The files often were barren of material they should contain, and instead were a collection of newspaper clippings loose notes, unverified and irrelevant information. Violation of files was frequent. Police reporters looking for a good story were given free access to files on suspects, and as a result were able to publish in the newspaper some interesting but in many cases misleading, irrelevant, and damaging pieces of information. Those police chiefs who tried to protect the confidentiality of their files received poor press treatment, so that they would be encouraged to cooperate with the press more fully in the future.

With the establishment of the statewide identification and intelligence system, one of the first steps was to define what material should be contained in the basic system. Unreliable and inadmissible evidence was excluded. Each agency contributing information was given the right to specify what other agencies should be allowed access to that information. Each administrative unit in the system has access only to that kind of information in the central file which it has been agreed in advance is proper. The intelligence system, furthermore, keeps a record

of all information provided to each individual user so that violations of disclosure rules will be apparent.

One of the strongest supporters of the new identification and intelligence system has been the Civil Liberties Union, which has argued that the new system regularizes the kind of information available, increases its accuracy, and protects the rights of the individuals involved. The identification and intelligence system itself has no operational or administrative responsibilities. It has been set up as an independent and impartial organization designed to meet legitimate requests for information and to protect against the misuse of information. It is interesting to note in this connection that the new agency does have responsibilities to provide data within proper disclosure rules for legitimate research work in the fields of crime, juvenile delinquency, mental health, and other concerns of social research.

The key to the problem of protecting privacy is not to depend blindly upon the inefficiency which may accompany decentralization—in many instances the decentralization may result in flagrant abuses which are difficult to uncover simply because of the extent of the decentralization. Furthermore, the problem of controlling disclosures cannot be solved, as has been suggested by some, by special devices such as eliminating the names of people in the central files, or collecting only a sample of the total data. As has been pointed out above, the matching of individual records is essential for the proper construction of statistical data, and impairment of the possibility of matching by omission of major identifying characteristics or by restriction of the data to samples is highly undesirable.

Instead, the problem should be attacked frontally in a manner similar to that currently pursued by the Bureau of the Census. The establishment of a national data center which will pool information obtained from all cooperating agencies would provide an opportunity to develop the required safeguards against improper disclosure of information both within government and outside of it. This will do more to insure individual privacy than permitting the present uncontrolled system of data access to continue. As has been pointed out by all of the previous reports on the topic, the pooling of data will at the same time reduce the redundancy and the cost of obtaining and maintaining statistical data, vastly improve the statistical base, and remove from operating government agencies the burden of servicing purely informational needs of those outside the originating agency.

Finally, by creating a national data center which supplements the present statistical system and improves the coordination and integration of data, we will still be able to retain the present advantages of decentralization which now provide a maximum amount of freedom for each government agency to determine its own statistical program and thus encourage innovation and experimentation.

Chairman TALMADGE. Thank you, Dr. Ruggles, for a very fine statement.

You testified we have too much duplication of effort in the statistical program. Will you give us some examples.

Mr. Ruggles. Yes, I think I can.

A wide variety of Government agencies collect information on establishments. If this information could be collected on a joint basis it would eliminate the need for the same establishment to file reports saying practically the same thing to different agencies.

As another example, the Bureau of the Census has been trying to use some of the income tax material to develop lists of names to help in the self-enumeration for the 1970 census. The Census Bureau can obtain from administrative records in the Government such as social security, internal revenue, et cetera, lists which they can use to send out census forms. A self-enumeration census improves confidentiality since the respondent will fill out the census return and send it in himself, rather than having a local census taker coming around and asking questions. Such a self-enumeration census will reduce costs and duplication, and will also protect privacy.

Chairman TALMADGE. Do present operating procedures take full advantage of modern techniques, professional specialists, and large-scale machines?

Mr. RUGGLES. We are all learning, and the new computer technology is rather recent. For instance, the IRS is very rapidly increasing its computer use on things like the matching of withholding statements from firms with the tax returns of the individuals in question, and so on. This is obviously not a static situation. I would expect large productivity gains in almost every operational program over the next 3, 4, 5 years.

Chairman TALMADGE. Do you think that much can be done to reduce the lag between the receipt of the information and its availability in usable form?

Mr. RUGGLES. Yes; very much so. I think this is one place where our progress is going to come very fast. The timelag generally resulted from the inability to get the information into the system more rapidly when it had to be done by coding and keypunching, all hand operations.

Today with reading and scanning devices the input of information can be accomplished very much more quickly and much more accurately, so that the computer can produce the results very fast indeed.

Chairman TALMADGE. What are the principal issues concerning centralized versus decentralized statistical systems?

Mr. RUGGLES. I do not consider that a national data center would be a centralized system. It would merely bring together and pool the output of the decentralized system. It would be able to perform many of the functions which a centralized system would have, because it could operate with the raw data in its original form. A decentralized system has the great advantage that the person desiring to use information for operational purposes can construct his data in the way that he wishes. I would hesitate very much before recommending the creation of a centralized system which would specify for all users what information they should have and what information they should not have.

Chairman TALMADGE. Is our system more centralized than those of other modern powers?

Mr. RUGGLES. It is considerably less centralized than most foreign powers.

Chairman TALMADGE. What were the main conclusions of the "Report of the Task Force on the Storage of and Access to Government Statistics"?

Mr. RUGGLES. The task force report recommends essentially that there be an interagency facility which would take the records of the contributing and cooperating agencies and within the proper safe-

guards of disclosure would perform statistical services both for them and for outside users, thus taking the burden of performing such services off the operating agencies.

Chairman TALMADGE. Without objection, I would like to insert in the record as an appendix to these hearings the "Report of the Task Force on the Storage of and Access to Government Statistics." (See page 196.) I believe that the report indicated that if the task force were to design a Federal statistical system de novo, it would clearly recommend the creation of a single central statistical agency. Why did the task force then go only part way in recommending a national data center?

Mr. RUGGLES. I think because any system is the result of evolution. Perhaps if you were to build New York City brand new today it would not be the same New York City that you happen to have at this moment. However, it is not therefore rational to destroy New York City and build a new city in the same location. I think the same thing applies to statistical systems.

One of the reasons why the U.S. system is more decentralized than those of other countries is that we started earlier, and we have more statistical capital involved in our present decentralized system than many countries who came later to the statistical field and designed a system at that time. You take advantage of the resources you have and you adapt them accordingly.

Chairman TALMADGE. What would be the principal functions of such a center?

Mr. RUGGLES. Of a national data center?

Chairman TALMADGE. Yes.

Mr. RUGGLES. I would conceive as its principal function would be supplying the kinds of information required for the design and implementation of policy.

Chairman TALMADGE. How much would it cost?

Mr. RUGGLES. This is a difficult question to answer because it is highly related to the cost of our present decentralized statistical system. In many cases a national data center would reduce costs, mainly because functions that are now carried out very expensively in different statistical agencies would be to a central facility.

On the other hand, I think anyone who has looked at the rate of increase in the budgets of the statistical agencies and seen how fast these are rising could not reasonably expect an actual reduction in total cost in the near future due to a creation of a national data center.

I do not think we can foresee; this will depend largely on the needs that are placed upon it.

Initially, the Kaysen committee, I think, felt that a starting budget of some $3 to $5 million would be required.

Chairman TALMADGE. How long would it take to become operational?

Mr. RUGGLES. If it undertook at first the bringing together of some of the most important data I would expect that it could become operational very quickly, indeed.

Chairman TALMADGE. What do you think would be the proper administrative set up for it?

Mr. RUGGLES. We have already discussed this question, and I would agree that I am somewhat perplexed as to just how it should be lodged within the Government. I feel it is essential for such an organization

to have interagency authority to obtain computer tapes, and the ability to service different agencies and groups outside the Government.

Chairman TALMADGE. Where is the bulk of the statistics now, in the Department of Commerce, Bureau of Labor Statistics, Department of Agriculture, HEW, or where?

Mr. RUGGLES. That depends whether you consider tax returns, for example, as statistics. Tax returns and the social security information probably are the largest bodies of data. However, both of these agencies are not considered to be primarily statistical agencies.

In terms of actual data, I believe the Census Bureau is the largest. In terms of budget, the Department of Agriculture may be the largest. However, the Bureau of Labor Statistics is also very important.

Chairman TALMADGE. Thank you very much, Professor Ruggles and Dr. Dunn.

Before adjourning, I would like to place in the record at the close of today's proceedings, a recent address by Dr. Dunn on "The Idea of a National Data Center and the Issue of Personal Privacy." The subcommittee will stand in recess until 10 a.m., tomorrow, when we will meet in room 1202 of the New Senate Office Building. The witnesses will be John Aiken, executive director, Federal Statistics User's Conference, and Frederick Stephan, professor of social statistics, Princeton University, past president, American Statistical Association.

Thank you very much.

(Whereupon, at 11:56 a.m., the subcommittee recessed to reconvene on Thursday, May 18, 1967, at 10 a.m., in room 1202, New Senate Office Building.)

(The address by Dr. Dunn, referred to above, follows:)

THE IDEA OF A NATIONAL DATA CENTER AND THE ISSUE OF PERSONAL PRIVACY[1]

Edgar S. Dunn, Jr., Resources for the Future, Inc.[2]

In late 1965 a report was submitted to the Office of Statistical Standards of the Bureau of the Budget entitled "A Review of Proposals for a National Data Center."[3] That report analyzed some of the anomalies that prevent the most effective use of the resources of the Federal Statistical System in the establishment of public policy, the management of public affairs, and the conduct of research. It recommended changes in the mission of the Federal Statistical System that could transform it into a more effective source of information services for today's needs.

During the time that this report was under review by the Administration it became "caught up" in a substantial public controversy over the alleged threat to personal privacy embodied in its recommendations. The report and the Administration's intentions were made the object of hearings before the subcommittees of Senator Long of Missouri in the Senate and Congressman Gallagher of New Jersey in the House. Through extensive comment in the public press, the report acquired the image of a design to establish a gargantuan centralized national data center calculated to bring Orwell's "1984" at least as close as 1970. Is the theme of this paper that the image embodied in the "purple phrases" that characterized the public reports do not reflect either the realities of the proposals or the balance that Congressman Gallagher and Senator Long attempted to bring to this issue in the hearings. The author wishes to take this means of correcting certain obvious misinterpretations and set forth more explicitly some views on the very important issue of personal privacy.

The topic will be presented in two progressions: from the particular to the general and from the short run to the long-run. We must start with the particular:

---

[1] Adapted from an address presented before the MENSA Society, New York, Oct. 21, 1966.
[2] Reprinted from *The American Statistician*, February 1967.
[3] Published by the Office of Statistical Standards as "Statistical Evaluation Report #6."

the author's report to the Bureau of the Budget and then move to a more general perspective of the issue. Because of the overriding importance of realistic time dimensions in the evaluation of this problem we also need to make a distinction between the short-run and the long-run and we shall progress, in our treatment, along this time path. (It will help to bear in mind that the author's concept of the short-run implies something like ten to fifteen years.) The treatment of this issue in the press and in public hearings has confused the particular and the general and the short-run and the long-run. We will begin by reviewing briefly what the "Dunn report" does and does not say. First, contrary to reports, it does not constitute a formal plan in any sense. It is primarily an informal review of certain problems and prospects associated with the management and organization of *statistical information* generated by *public, general-purpose statistical programs* within the federal government. It was a preliminary review carried out with limited time and resources. The conclusions of this review were circulated within various administrative agencies as a basis for discussion and taken for evaluation by the administration.

The report did state that certain obstacles to effective use of federal statistical files might require some centralization of function. It did not at any point recommend what changes should be made or what agencies or files should be involved. It contented itself with a generalized treatment of the problem and some indication of the general direction in which the solutions might lie. It was not presented as a final program design.

Turn now from what the report did not say that people think it said to what it did say that few people have noticed. First, it pointed out that the new technology has made possible a new kind of capability in servicing the requirements of public policy and public management for statistical information, and that this capability is an order-of-magnitude different from any capability we have had in the past. It identifies a major "pay-off" in improving the public and private decision process, and in supporting research aimed at improving our understanding of the social process. This emerging potential of statistical programs received short shrift in congressional hearings and was totally ignored by the press.

Consider the benefit side of the equation for the moment. This benefit is not to be identified in terms of the direct savings and costs for some given number of traditional requests or inquiries made upon statistical files. It derives from the fact that there is available a new dimension of information service for the decision process. The source of this benefit stems from the fact that the central problem of using statistical records for purposes of policy, management, or research has always been one of associating statistical records. No number conveys any information by itself. It acquires meaning and provides useful insight only when compared with other numbers. The policy maker wants to know the answer to questions such as this: "What proportion of the residents of Appalachia have incomes under $3000 a year and how do their age, race, sex and educational characteristics differ from those with higher incomes both within and outside Appalachia." In this one typical question we can identify dozens of separate statistical attributes of interest. Categories of income, age, race, sex, education, geographical designation, etc. all have to be specified and *related in some way that is relevant for the problem at hand.* This often requires that attributes of different and separate statistical records be brought together to reveal the particular collection of characteristics of use in analysis. In the not too distant past this kind of record association could be done only if the attributes of interest existed in traditional publications and conventional tabulations in a form from which a statistical clerk could, without exorbitant expense, extract a new tabulation that met the program requirements. In the more recent past there has been a growth in the possibility of doing special tabulations from machine records where the attributes of interest happen to be found in compatible records—usually within the same agency or program.

Commonly, however except for a limited portion of the satistical uses of interest to policy, many policy decisions have to be made on the basis of very sketchy information because there has never existed the kind of capability that could fulfil these needs.

Until recently the principal reason for this has been the lack of technical capacity to manipulate large quantities of numbers fast enough and with sufficient economy to make it feasible to meet the information requirements of policy in a flexible way. That technical capacity now exists and is growing. This has tempted many people to the conclusion that we should computerize all the data we have in the back room so that we can match any number instaneously with

any other number at the push of a button! A careful reading of the report will reveal that a major part of its effort was directed at denying this hypothesis. It rejected the "naive data bank concept"—or what some call "data dumps"—and indicated that its widespread acceptance among some advocates of extended statistical systems was a matter of concern.

For statistical purposes the main consideration in developing a more serviceable information system is not the assembly of a large data bank of files from a wide variety of sources. The effective use of data in statistical analysis requires data that have classification attributes *relevant* to the decision problem and that have *reliable* qualitative characteristics. Without these conditions being fulfilled you can push numbers around in machines ad infinitum without effectively serving the decision process—in fact you would more likely confuse and mislead it. Those who know the problem of effective computer use and statistical analysis have a favorite acronym—GIGO. This stands for "Garbage In—Garbage Out."

The biggest problem in developing a more effective statistical system has little to do with the development of computer hardware systems to do the job; it has to do with needed changes in production processes that generate and make available the statistical records and synthetic series that form the basic analytical capability.

Most of the report dealt with the institutional and procedural anomalies that result in statistical files that are unresponsive to the needs of many vital policy programs. Briefly, they are such constraints as the following:

1) Important historical records are sometimes lost because of the absence of a consistent policy and procedure for establishing and maintaining archives.

2) The absence of appropriate standards and procedures for file maintenance and documentation lead to low quality files that contain many technical limitations in statistical usage.

3) Many useful records are produced as a by-product of administrative or regulatory procedures by agencies that are not equipped to perform a general purpose statistical service function.

4) No adequate reference service exists that would allow users to determine easily whether or not records have the characteristics of quality and compatibility that are appropriate to their analytical requirements.

5) Procedures for collecting, coding and tabulating data that were appropriate when developed now lead to some incompatibilities in record association and usage required by current policy problems and made possible by computer techniques.

6) There are serious gaps in existing data records that stand in the way of bringing together records of greatest relevance for today's problems.

7) The need to by-pass problems of record incompatibility in developing statistics appropriate for policy analysis, places severe strains upon regulations restricting the disclosure of information about individuals. Technical possibilities for using the computer to satisfy these statistical requirements without in any way violating personal privacy have not generally been developed and made available by the agencies.

Changing the practices of the Federal statistical agencies to bring about more relevant and effective statistical resources for modern public policy and management will be a time-consuming and resource-consuming job. We could not possibly spare the intellectual and financial resources to make *all* statistical files meet the necessary standards. *And, if they do not meet the necessary standards for effective statistical use, there is no need for them to be incorporated into a computerized statistical service system!*

This means that we must move first to identify the most vital problem areas affecting public policy and management; second, to determine the statistical requirements that will meet these needs; third, to establish the standards and practices essential to the generation of *relevant and reliable* statistical records to fulfill these requirements; fourth, to provide the institutional forms and mission concepts to provide effective statistical services; and fifth, to support these records with computer systems able to provide the essential flexible servicing capability. It is the fifth and easiest step that many naively feel is sufficient to solve the problem of statistical services. Anyone familiar with the operation of statistical systems and the history of the federal general-purpose statistical programs will recognize this as a formidable problem that can only be solved in stages over a number of years.

We have examined what the report did not say that people think it did and what the report did say that was unfortunately ignored. Now let's turn to what the report did not say which it should have said.

First of all, it said very little about the issue of personal privacy. This now stands revealed as a gigantic oversight for which the author takes full responsibility. For anyone who views this report and this issue from a perspective outside of the Federal statistical system, this oversight must seem incomprehensible. The reason is very simple, if inexcusable. In preparing this report it was addressed, in terms of its subject matter, to improvements in the servicing capability of the Federal statistical system and, in terms of its audience, to the members of the administrative family engaged in the activities of that system and thoroughly knowledgeable about its characteristics. Consequently, it was assumed that the protection of personal privacy was a given condition that was understood by everyone concerned. This was thought justified in dealing with such an audience because legal and procedural protections against revealing information about individuals have been a very basic part of the operation of the Federal statistical programs for many many years! Furthermore, these protections have been phenomenally successful! The protection of personal privacy has long been an obsession with the directors of federal statistical programs because the success of these programs have always depended upon cooperation of the respondents who supply information. No successful statistical program could exist without full confidence that personal privacy was secured!

The second big omission, stemming from the nature of the report and its audience, was the failure to distinguish clearly between statistical information systems on the one hand and intelligence systems on the other. This distinction was introduced in testimony before the Gallagher Committee.

The distinction is basic. *Intelligence systems* generate data about individuals *as* individuals. They have as their purpose "finding out" about the individual. They are widespread and common and essential in our private and public business. They include such things as the medical records a doctor keeps to trace the changes in the well-being of his patient and the educational records the teacher keeps to trace the progress of a student. They include requirements essential to public administration, such as the results of tests by driver licensing authorities concerning vision or tax information needed by the tax authorities.

A *statistical* information system produces information that does not relate to the individual. It only identifies characteristics that relate to groups of individuals or so-called "populations". It has as its purpose answering such questions as these. "In what way does the mix of economic activities in New York City differ from that of Chicago?" "What proportion of the registered voters turned out in a recent primary and how were they divided between Republicans and Democrats, urban and rural, white and nonwhite?" The range of the questions is infinite.

The important point to emphasize is that a statistical system is concerned with generating aggregates, averages, percentages, etc. that describe relationships characteristic of groups or populations of individuals. No information about the individual is generated as output and no information about the individual needs to be available to anyone outside the system under any circumstances for the statistical information system to perform its function.

This distinction divides the issue of personal privacy into two parts. The first part of the issue is reflected in this question: Can a statistical information system be developed and administered in a way that assures that it cannot be used as an intelligence system? The author is sure that the answer is yes.

Here the distinction between the short-run and the long-run comes particularly into play. We have seen that the coordinating requirements in the statistical system often will require reformulation of programs in the production of data. Thus, over a period of some years a modification of the system will have to proceed with only those limited subsets of all conceivable existing files that are relevant to the most urgent policy requirements. Consequently they will deal mostly with traditional statistical records that have contained information dealing largely with the public face of the individual (such things as the demographic characteristics like age, race, sex, etc.) in contrast to the private face of the individual (such things as criminal records, medical records, psychological tests, etc.). There is nothing in sight in the short-run future that would change the scope or content of such general-purpose data or their organization in any way contrary to the existing tradition of protecting personal privacy. In the future as the system evolved in scope and effectiveness it would be possible to extend the legal and procedural protections against the misuse of a statistical system for intelligence purposes. Computer technology cuts two ways. It provides us with new and powerful techniques for controlling and protecting the misuse of the record.

The Gallagher Committee and at least one of the witnesses seemed unwilling to accept the distinction between the short-run and the long-run on the one hand, and the distinction between the statistical information systems and intelligence information systems on the other.

This skepticism grows out of (1) the fact that a statistical system must contain information about individual respondents, thereby rendering it potentially useful for intelligence purposes; (2) the fact that no system designed exists for providing *foolproof* protection against file misuse in this way; and (3) the argument that the pace of technology is proceeding so fast that there will be no technical limitations on accumulating "all of the data about everyone." They would claim, therefore, that the long run is already upon us.

The author would still maintain that these distinctions are valid and vital.

It is true that with respect to *intelligence* systems, there is no point in making any effective distinction between the short-run and the long-run. The detailed and careful standards with respect to collecting, coding, and tabulating data essential to a meaningful matching of attributes for statistical analysis is not essential for single-record searches for intelligence purposes. Therefore, the development of intelligence systems is not constrained by a prior requirement for production reforms.

Furthermore, the system design, from both the hardware and software point of view, essential to perform a satisfactory intelligence retrieval function is already available for some quite large files and capable of rapid development. As Mr. Baran (a computer expert from RAND Corporation) pointed out the Gallagher Committee, we are already building the bits and pieces of an intelligence data bank that can be quite easily merged. He pointed out that no one planned a national railway system. It started out as short routes connecting local population centers. It gradually merged into a larger system. In effect, an intelligence system may be 50 per cent in being already. In substance, he is saying "instead of distinguishing between a short-run and long-run it is essential to realize that it may be later than you think." Mr. Baran is right with respect to intelligence systems. He is wrong, we think, with respect to the threat of misuse of a statistical system in the short-run. Consider briefly the reasons:

(1) As we have seen, a statistical information system of greater utility for policy cannot be developed without making substantial changes in production practices. This yields the necessity for changes to come slowly. Subsets of traditional files will need to be modified and integrated for matching purposes on an incremental basis with priorities established by important requirements.

Mr. Baran's railway or telegraph merge concept does not apply so simply here. We might transform the analogy to provide a more legitimate comparision if we visualize the development of the separate railway systems having taken place in such a way that each segment was technically incompatible with each other segment so that the system of lines might require something approximating a 100% replacement capital through obsolescence if the system were to be integrated. The present problem the world faces in integrating incompatible TV systems or incompatible systems of measurement might form a closer analogy. In short, no one is going to plan a complete integration of all statistical records over any short-run period of time. It would cost a great deal, and regardless of cost it would take a considerable period of time to put into effect.

(2) An *intelligence system*, if it is going to be efficient, has to be as nearly complete as possible. Ideally it should constitute a census so that every possible individual search request could be fulfilled. There has been an implicit assumption in committee hearings that this is also true of a statistical system, but it emphatically is not. We have found increasingly that the efficient statistical system (since it generates related information about groups of people and never about individuals) doesn't want "all the data on everyone." It only wants some of the data on some of the people—enough to be relevant for the important problems of analysis by private business, government, and researchers and enough to support reliable inferences. To build a complete file is inordinately expensive, and we have found, for most statistical purposes less reliable. Indeed, the national "census," since it is conductd primarily for statistical purposes rather than intelligence purposes, is a complete census for only a very few attributes of the population. The bulk of information is collected on a sample basis only.

(3) The existing statistical systems have had considerable experience and an admirable record in protecting personal privacy through legal regulations supplemented by operational procedures. Initial moves to improve the matching characteristics of federal statistical records *for statistical purposes* could be carried out under an extension of well-established protection procedures.

(4) Often data that is irrelevant for intelligence purposes (concerning deceased respondents, no longer existing enterprises, etc.) is a prized content of a statistical file because of its utility in permitting the analysis of statistical trends and other indicators of social change.

In short, the changes in the Federal Statistical System currently needed would not generate files sufficiently comprehensive is either scope (that is, the numbers of individuals) or content (data would be primarily restricted to the public face of the individual) to turn it into a comprehensive intelligence system. It is, furthermore, already protected by well established procedures that can be extended and improved. A statistical file would have so many gaps in the kind of information important for intelligence use and contain so much information irrelevant to intelligence use that it would be grossly inefficeint instrument as a source of personal intelligence. The incentives to pervert such a statistical system for intelligence purposes are missing because less costly and less risky intelligence sources are already available and are more complete.

At this point let me insert an example to illustrate both the new dimension in statistical analysis that is emerging and the dramatic payoffs that are available. Through most of the evolution of the Federal statistical system attention has been almost entirely limited to measures of particular economic and social phenomenon. Attention was mostly focused upon individual series such as the size of the population, the volume of foreign trade and the output of manufacturing. Most of the uses served either public or private management. The series usually found their origin in some particular management need.

We have been finding over the years that these management-oriented series do not serve the information requirements of policy determination very well. We had this forcefully brought home to us in the 1930's when we found that the effort to establish public policy to cope with a stagnant and unstable economy was foundering for want of any comprehensive measure of the economic performance of the nation. As a consequence, over the last three decades we developed a system of economic accounts—the national income and product accounts produced by the Office of Business Economics. The interesting thing about this statistical program is that it produces synthetic statistics *based upon the matching and coordination of great masses of data* taken from individual management-type series. Because of the basic incompatibility of many of these records it has taken several decades to bring them to their present stage of refinement and the task is not yet complete. The problem is that these kinds of analytical problems in policy formation require that data be brought together that are *separately generated* in the collection process but which pertain to essentially *connected* forces in economic and social behavior. In analyzing these connected forces in the establishment of policy we must examine the relationship of many data series.

Our experience with the national income and product accounts is a monument to the difficulty of serving policy with this kind of information and a monument to the payoffs that are available when we do. The return on the money spent on this program must surely be several thousand percent. Without it the successful role of public policy in the monetary, fiscal and budgetary process in promoting the stabilization and growth of the American economy would almost certainly not have been realized.

Most of the effort and expenditure and most of the development time essential to the production of the national accounts was consumed in reconciling incompatible characteristics of the individual statistical series that had to be associated in analysis.

The growing information requirements of public policy are forcing us more and more into these forms of data integration in a situation where the production process that characterizes the statistical system are generating management-oriented series that are often technically incompatible for policy use. We can't go on building new Offices of Business Economics to force the recalcitrant series into synthetic forms appropriate for policy, or go on providing huge sums of money to individual agencies (as we have in the case of OEO) to carry out special statistical missions that are completely beyond their capacity to perform. The cost in resources and in time is becoming formidable. Yet policy must be served!

There is now emerging via the computer the possibility of integration of data at the primary level of data production, provided that the standards and procedures are modified to realize this potential. To do so will not only allow existing comprehensive series (such as the national economic accounts) to be improved and produced more cheaply, they will open the way to a much more flexible response to fulfilling the information requirements in other policy areas. The

payoffs should be every bit as spectacular as they have been from the development of the national accounts.

It is easy to see that the principal element in the development of this new statistical capability is the change in mission concepts and the coordination of practice in the Federal statistical system. The computer makes it feasible to work towards a system more serviceable to policy, but the development of large banks of data of *indiscriminate type is* definitely not an attribute of this program. Unless the data can meet the rigorous production standards for effective policy use it would be a handicap to have it in the system.

Can a "foolproof" system be developed to prevent the misuse of a statistical system for intelligence purposes? The answer is no. But this is not the relevant form of the question. Mankind has never been able to develop a foolproof method of safeguarding any human value. It is unreasonable to expect anything different here. If one asks whether the cost of improper use can be made prohibitive, the answer is an unqualified yes and we already know a lot about the techniques for accomplishing this and have a good record to build on as far as statistical systems are concerned. Suffice it to say that, with respect to statistical systems, the long-run is far enough away to allow ample progress in technology and law to protect personal privacy adequately as statistical systems develop in flexibility.

In the heat of public controversy on this issue we have heard impassioned demands that we halt any improvement in the statistical system until complete protection of personal privacy can be designed. Such an argument condemns us to failure before the start. The only doorway to the long-run is through the short-run. Systems like this are not designed in the abstract. They emerge out of practice and experience in meeting requirements. One reason we are in such a strong position to proceed with adequate protections against the invasion of personal privacy is the fact that, through practice, we already have developed useful techniques and experience that can be extended in new system design.

Turn now to a more general perspective.

What has been happening in the recent public controversy seems to be understandable and probably necessary. Every major social and technological innovation in the advance of human society serves as a powerful stimulus both to man's legitimate dreams and to his legitimate fears.

Robert Bugoslow has recently written a little book with the catchy title, *The New Utopians*. In it he argues that "there is a new breed of Utopians afoot threatening to rush down all of the exciting pathways and blind alleys frequented by Utopians since the days of Plato." The Utopians are identified with the computer gang and systems designers who see in the new mechanized systems a panacea for many of our ills. Bugoslow is right. There are such creatures about.

It is not surprising that, in the face of untempered dreams, we should encounter untempered fears. The specter of the negative utopia of Orwell seems an obvious instrument to counterbalance the scale. The public conscience seems to become obsessed, in turn, with the benefits and then the costs. The public education about complex issues often seems to require such a dialectic exchange. Our task for the future is to guide the dialogue to a joint evaluation of the cost-benefit ratio. We need to develop the capacity of dream and fear at the same time and reconcile the two in the day-to-day task of getting the job done.

The necessity of the latter is an inescapable aspect of the human condition. Every personal and public advance is won at some cost because each individual and each society has competing interests. They have to be reconciled in the day-to-day process of personal development and social evolution.

Many of these conflicts of interest take on a public vs. a private context. We know many sources of these competing interests and work with them every day. There is the interest of national security versus the free dissemination of knowledge; proprietary interests versus the freedom of scientific knowledge; private property versus the public domain; and, of course, personal privacy versus the needs for public information. We labor daily in a free society to reconcile these interests where possible and, where not, to choose between them.

The competing interest between personal privacy and public information is one of this ubiquitous class of conflicts. It has many facets of much wider significance than has been recognized in the current record of investigation and public controversy. Just to identify a few important subsidiary problems, there is the issue of personal privacy versus effective government; personal privacy versus behavioral research; personal privacy versus law enforcement; and personal privacy versus free dissemination of the news.

But expressing the conflicts of interest in this way has the effect of making the conflict appear to be a contest between governments and individuals. How

often our thinking becomes trapped in this oversimplification. The conflict, in the end, is between conflicting aspects of our own individual personal interests. Law enforcement, behavioral research, freedom of the press and effective government enlarge some of our personal freedoms through the instrument of at least partially restricting others. The problem, of course, is that we have never devised a way, in a free society, to allow every individual an unrestricted choice in his market basket of freedoms and still maintain the viability of the society upon which all freedoms depend.

It is not proposed that the problem of establishing a socially acceptable balance in this area of information and personal privacy is necessarily of the same order of importance as all other conflicts of interest. The instinct of the Congressional committees and of the public is correct. The question of how to develop information and how to use it is without a doubt one of the most vital of all public issues. Information is power. But both information and power are morally neutral—each has the ability to enslave and to release.

Here on the most general and philosophic plane we have the issue restated. In the final analysis what are the benefits to be derived from improved information systems? The stake is our success as a nation in our complex world. The future of mankind is bound up in the accumulation and effective use of pertinent information. In our case, if there is any critical deficit, it is in the realm of information that will serve society's need to establish policy and manage its public affairs.

But if the stake is large, so is the risk. The risk is the dehumanization of man; or perhaps, put more accurately, that we shall fail in our long-term effort to become fully humanized.

This brings us to a critical point. In the end, the important thing is not that we must strike an operative balance in solving these problems, as we will and must. The important thing is what standards serve as our guide as we attempt to strike the balances and restrike them every day and year. What is it that motivates our purpose? Which will be the dominant ethic—if you will? The author would wish every decision in the generation and application of new knowledge to be governed by the contribution it can make to provide every man with the best chance for the realization of his human potential. This same point of view is very likely motivation behind Congressional interest in this issue. Their legitimate concern is to find ways in which this motive can be supported.

In short, the issue of personal privacy is really only part of the larger and more fundamental issue: How can information, which is really the codification of all human knowledge, be made to serve the goal of national development and human enrichment. In this context the long range evolution of statistical systems is seen to be essential to the achievement of these goals.

In closing, the thesis presented here has these essential elements.

(1) We are engaged in discussing a public issue that is of the greatest importance to the future of our society. The emerging prospects of better and more useful information systems hold great promise for human welfare. They also contain the possibility of misuse if not guided by an appropriate social ethic and safeguarded from improper violation of constitutional rights. Neither proposition can be honestly denied.

(2) The legitimacy of these concerns is current and immediate. Intelligence systems already exist and are already compiling records of misuse. This potential abuse can grow rapidly even in the short-run. We obviously need to turn as quickly as we can to establish policy and protections in this area.

(3) It is unnecessary and unfortunate if we identify the prospects for more rational utilization of existing statistical resources with this personal privacy issue in the short-run. The short-run choices are not painful or excessively dangerous. There is at least a decade or more in which we can work to effect an internal rationalization of these resources and serve policy and management better without running the risk of creating an instrument with great potential for the invasion of privacy. We need to get on with the task because of its importance for public welfare and because in the exercise of this effort we have the best chance for developing the extended protections we must develop as an adjunct of all future information systems.

(4) We must somehow commit resources immediately to the task of developing the techniques for safeguarding human rights as we exploit the great advantages of our new technology and as we take up the task of thinking through a national information policy. Not only do we need to engage the effort of leaders in the scientific community we must provide for the initiation of a new kind of technical staff work.

This paper has sought to achieve a balanced approach to these problems. It has not been made clear to the public that the leaders of the congressional investigations have also sought a balanced approach. In his opening statement Congressman Gallagher said, "What we are looking for is a sense of balance. We do not want to deprive ourselves of the rewards of science; we simply want to make sure that human dignity and civil liberties remain intact." This same concern was reflected in Mr. Gallagher's summation on the floor of the House, when he made the constructive proposal that a "symposium" of the scientific community be established to help resolve this issue and to provide the information to guide the formation of public policy. These congressional leaders are to be commended for their concern with this issue. One may hope that the public image will soon be restored to the sense of balance that was sought by the legislative branch and that a professional dialogue supported by adequate staff work will bring agreement upon a path by means of which our goals may be achieved.

THE COORDINATION AND INTEGRATION OF GOVERNMENT STATISTICAL PROGRAMS

THURSDAY, MAY 18, 1967

CONGRESS OF THE UNITED STATES,
SUBCOMMITTEE ON ECONOMIC STATISTICS OF THE
JOINT ECONOMIC COMMITTEE,
Washington, D.C.

The subcommittee met, pursuant to recess, at 10 a.m., in room 1202, New Senate Office Building, Hon. Herman E. Talmadge (chairman of the subcommittee) presiding.

Present: Senator Talmadge and Representative Bolling.

Also present: James W. Knowles, director of research, and George R. Iden, staff economist.

Chairman TALMADGE. This morning, the subcommittee continues its hearings on the coordination and integration of Government statistical programs by hearing users' views on ways to improve our statistical programs and ways to improve our statistical system. The participants are Mr. John Aiken, executive director of the Federal Statistics Users' Conference; and Mr. Frederick Stephan, professor of social statistics at Princeton University.

Gentleman, we appreciate your coming to discuss with us how the Government statistical programs can better serve the public needs.

Mr. Aiken, will you please proceed as you wish.

STATEMENT OF JOHN H. AIKEN, EXECUTIVE DIRECTOR, FEDERAL STATISTICS USERS' CONFERENCE; ACCOMPANIED BY DR. ROY E. MOOR, VICE PRESIDENT AND ECONOMIST, THE FIDELITY BANK; MARVIN FRIEDMAN, ECONOMIST, AFL-CIO; AND DR. JOSEPH E. MORTON, THE W. E. UPJOHN INSTITUTE FOR EMPLOYMENT RESEARCH

Mr. AIKEN. My name is John H. Aiken. I am executive director of the Federal Statistics Users' Conference. I am accompanied today by representatives of the three major groups which make up our membership. From the business group is Dr. Roy E. Moor, vice president and economist with the Fidelity Bank; from the labor group is Mr. Marvin Friedman, economist with the AFL-CIO; and from the nonprofit research group is Dr. Joseph E. Morton with the W. E. Upjohn Institute for Employment Research. They are here to participate in the discussion after my oral statement.

FSUC is an association comprising approximately 160 organizations generally classified as business firms, labor unions, and nonprofit research groups, who have a common interest in obtaining adequate, timely, and reliable information from Federal statistical programs.

Our membership is highly diversified. More specifically, it includes representation from trade associations and industries engaged in advertising, banking and finance, insurance, manufacturing, retail trade, printing and publishing, economic and market research, and so forth. Almost every segment of the economy that uses Federal statistics is represented in the conference.

We are pleased to have this opportunity to express our views on the coordination and integration of governmental statistical programs.

General Comments and Observations

Because of the unique nature of our organization and functions, we obviously have a real interest in the problem of coordination and integration of governmental statistical programs. Generally, however, our attention is concentrated on appraising users' requirements and assessing the degree to which existing and proposed programs meet such requirements. We regularly and consistently evaluate individual statistical programs and make suggestions for improvement. We also strive to keep our members informed and up to date on the status of current and proposed Federal statistical programs.

Thus, while the focus of our interests and activities has been primarily aimed at specific programs and responsibilities of individual agencies concerned with statistical activities, rather than at the system as a whole, many of our suggestions and recommendations for improvement require a certain degree of coordination and integration within the system in order to effectuate them. In testimony before congressional bodies we have repeatedly pointed out the dangers and possibilities of effort which means that coordination and integration is necessary if duplication is to be avoided or eliminated.

The decentralized nature of our statistical system and the division of responsibility requires effective coordination and integration of the various programs if we are to achieve maximum effectiveness, to reduce duplication, and to find areas for improvement. As we all know, the responsibility for carrying out this task has rested with the Office of Statistical Standards of the Bureau of the Budget for some time. In recent years, there have been many significant developments relating to our system that surely must complicate and make more difficult the task of OSS in carrying out its mission in this regard.

In view of the significant changes occurring in our statisitcal information system, the question is whether or not our present instruments for coordination and integration are adequate for doing the job required and where could we strengthen them so that a better job can be done. In our testimony before the Joint Economic Committee in February 1964, we said:

> It is about time to have another careful review of the Government's statistical program as a whole.

Implicit in this recommendation is a review of our means and resources for coordination and integration. These hearings of this subcommittee are a welcome step.

For the sake of time, I can only highlight some of the developments that compound the difficulties of developing a fully coordinated system of information.

GROWTH AND COMPLEXITY OF THE STATISTICAL SYSTEM

The amount of money allocated for Federal statistical programs has tripled, from $40 million in 1957 to about $125 million in fiscal 1967. The number of employees engaged in statistical activities has increased from 11,226 in fiscal 1959 to 14,619 in fiscal 1965, or 30 percent. However, these figures do not indicate the growth or wide range of statistical programs now being undertaken by the various agencies. But even here, if numbers were available, one would have to take into account the size and scope of individual statistical projects. In this connection, however, the House Committee on Post Office and Civil Service points out that in fiscal 1965, there were 1,185 Federal contracts for statistical services costing about $57 million.[1] The breakdown on these contracts is as follows:

Government: Federal, State, and local	426
Private	450
Universities and colleges	309

In our own analysis of the budgets for statistical programs for fiscal 1968, contained in special analysis K of the Federal Budget, we were able to identify roughly 56 new or expanded current statistical programs. With regard to one single agency, we note that the Office of Research and Statistics of the Social Security Administration, in its "work plan" for fiscal years 1967–68, is undertaking roughly some 79 research projects in addition to its recurring activities.

The point we are making here is that in terms of numbers, if nothing else, the vast proliferation of research and statistical programs and projects now being undertaken is so overwhelming that we wonder how any agency or organization can keep abreast of these many programs in order to discover probable areas of duplication and to determine where coordination might be required.

We, as an organization, and our individual members, consider ourselves fairly sophisticated and knowledgeable about the sources and availability of Federal statistical data. But, it is not easy to keep fully informed except in specific subject areas of major interest. Think what this task must be for the less sophisticated user or for the occasional or potential user. What is lacking and is sorely needed is a comprehensive and uniform inventory of research and statistical data and an indexing system.

ACCELERATED IMPROVEMENT AND USE OF THE COMPUTER

The use of computers in the Federal Government has grown from two in 1950 to over 2,600 in 1966. The number of agencies with ADP equipment was 45 in fiscal 1966, 39 of which have electronic digital computers and six of which use punched-card equipment only.

We know that steps have been taken to provide a greater measure of central policy direction, coordination, and guidance to the Federal agencies in the development of computer-based systems and the acqui-

[1] 1965 Report of Statistical Activities of the Federal Government: Personnel, Equipment, and Contract Costs.

sition and use of ADP equipment. However, we are still faced with the problem of adopting measures that will provide consistent and accurate information which can be linked together effectively, that will make certain that our major frameworks and models are reliably constructed to guide the collection and analysis of such data and that will provide for the efficient flow of data to meet regular and special demands.

Because of the increasing number of computers being used by government and users, there is an increasing need for tapes for research needs, particularly in the way of special tabulations. Often, the users special needs require collation with other data and cut across agency lines. This problem will be dealt with later.

NEW USES AND NEEDS FOR STATISTICAL DATA

In the past several years, a great deal of legislation has been passed dealing with problems of manpower, housing, health, education, welfare, and poverty. New and expanded programs in these areas require a great deal of factual information, the latest of which often is available only from the 1960 censuses of which often is available only from the 1960 Censuses of Population and Housing. Certain aspects of a number of these programs have called for many kinds of new information as well as data on a local and regional area basis which has not been available or available only in rudimentary form.

A number of programs have been initiated and proposed by various agencies to fill some of the information gaps and includes collection of data on a small area basis. As we have often pointed out these programs call for a coordinated approach by the agencies involved and failure to do this could result in duplication of effort, waste of scarce resources, and unnecessary increased paperwork burden on respondents. In this connection, we have suggested to the Joint Economic Committee that the Federal Government should make a detailed study of the needs for subnational demographic and economic data and that the study should: (1) emphasize fuller utilization of presently available information without undertaking new programs, and (2) distinguish between those areas where the Federal Government should concentrate its efforts and those areas where other organizations, groups, and levels of government should concentrate their efforts.

There is another problem involved in the collection of new and additional data on a small area basis under the new legislation. This, too, has been pointed out in previous testimony before the Joint Economic Committee. In most cases, neither legislation nor administrative procedures prescribe standards to assure that the information is uniform as among areas. For some kinds of information, the lack of uniform standards may be unimportant in meeting the needs for a specific program. For some kinds of information, as, for example, data pertaining to such matters as population, employment, and income, the national interest in having comparable data, gathered and compiled by standard methods, using common definitions, transcends the immediate needs of any particular program. Such information is so important that steps should be taken to secure data which are comparable from one area to another. Without comparable data there will be no common measure to evaluate the success or failure of specific programs or to decide whether or not the existence, expansion, or curtailment of par-

ticular programs should be the order of the day. Here again, coordinated effort is vital and we do not know what steps are being taken to accomplish this objective.

NEW DIMENSIONS AND EMPHASIS OF ECONOMIC AND SOCIAL INQUIRY

Related to new uses and needs is the extension of statistical inquiry into new areas of economic and social interest because of changes occurring within the economy and increased emphasis of governmental policies in certain economic and social problems. Among the economic changes which have occurred are the shift to services and the decline in agricultural employment and changes in the labor force. With regard to the need for better economic statistics, Congressman Curtis has stated:

> Many of today's critical issues in economic policy involve manpower and human resource development, equal employment opportunity, elimination of poverty, regional economic development, and the problem of hard core unemployment.

Illustrative of one of these areas is the informational gaps pointed out in the recently issued annual report on manpower requirements, resources, utilization, and training by the Department of Labor:

> "More information on people outside the labor force—covering such questions as how many are neither working nor looking for work, who they are, why they have withdrawn from or failed to enter the work force, and how they and their families live—would help to guide job development and related programs, particularly in the urban slums.
>
> "The current population and labor force surveys fail to reach and count considerable numbers of people. Information on who these people are and how they can be reached is essential to program planning.
>
> "Regular collection of labor force data on a narrower geographic basis would help to pinpoint the problem areas of the Nation. Such data should be compiled by city, county, metropolitan area, and State.
>
> "Information on the scope and impact of rural unemployment and underemployment is practically nonexistent and much needed. * * *
>
> "On-the-job training in industry is the major path to skill development. Yet comprehensive information is lacking on the nature and extent of formal and informal training in industry. A broad survey and assessment of occupational training in this country are greatly needed, with emphasis on industry training activities.
>
> "The decennial census, in an economy as dynamic as that of the United States, does not provide sufficiently up-to-date information on basic demographic, economic, and social changes. A large-scale statistical effort to update the decennial census at mid-decade, together with more frequent sample surveys, would provide a more solid foundation for program planning and action."

Another development illustrating new directions of inquiry was pointed out by Ray Bowman of the Office of Statistical Standards to this subcommittee last year in its study of "Improved Statistics for Economic Growth." He called attention to the "wide-ranging nature of analysis and policy which has increasingly crossed the lines of traditional disciplines."

He said:

> Analysis of and policy concern for the performance and prospects of society have gone well beyond strictly economic considerations and multiplied the number of issues faster than our ability adequately to evaluate them. Such areas of inquiry as the changing nature of population expansion; the relation between education and economic growth; the motivations to enter or leave the labor force; the interplay between prices, productivity, and wages; the impact of economic changes on particular groups and communities; the increasing statisti-

cal study of disease; the research underpinnings for policy in the areas of transportation and urbanization—all of these and many more have both widened and intensified the scope of inquiry.

The complex interplay of such wide-ranging questions involves identification in many dimensions of individuals, groups, firms, administrative units, and communities—and their changes over time. This has greatly multiplied the technical requirements for consistency of data which will permit the appropriate analytical manipulations.

Another significant development has been a tremendous expansion in the collection of social statistics and extension into a greater number of social areas. This has created considerable problems of organization and coordination. This problem was dealt with briefly by Mr. Bowman in his statement to this subcommittee last year. This next section refers to a recent report of a House Committee on Government Operations in regard to social research. I don't want to quote it, but the essence of it is that the questions of value of a certain amount of social research indicated there is need for greater coordination in this area.

In this connection, I would like to call attention to a recently issued four-volume staff study prepared for the Research and Technical Programs Subcommittee of the House Committee on Government Operations.[2] The release accompanying the issuance of this study made several statements that we believe are significant and have a bearing on the problem now under review. The following quotes are extracted from that release:

Many Federal agencies and university social scientists have been more interested in the pursuit of knowledge for its own sake than in the use of research to evaluate or to improve programs directed to the Nation's major social problems. This divorce of much social science from the study of issues actually confronting the society reflects both an academic preference and a failure of Federal agencies to understand how social research can help them . . .

Although social research constitutes a modest proportion of the Federal Government's annual $17 billion expenditures for research and development, there has been a five-fold increase in Federal funds for social research from some $73 million in 1960 to over $380 million in 1967. Yet, there is no systematic review of, or even information on, the content and quality of Federally-financed social research; the extent to which it is made available and used by Federal agencies, the local governments, and other groups; and its effectiveness in helping the Nation to improve education, combat crime and delinquency, reduce poverty, and cope with other major social problems . . .

Federally-financed social research is too often trivial or irrelevant; usable, but not used; valuable, but buried in scholarly journals or government filing cabinets.

There is an excessive emphasis on undirected, small-scale research proposed by individual investigators, and not enough on large-scale, coordinated programs directed at specific objectives. This criticism was often made of research in education, social welfare, and the social aspects of medicine and health . . .

The current fashion of quantitative analysis and indiscriminate accumulation of large bodies of facts, made possible by the computer, often leads away from, rather than toward, greater social understanding. Yet the Federal Government encourages the first and permits the second in the name of "science". . .

The above needs no elaboration and certainly highlights another problem area where coordinated effort is necessary. We recognize that not all social research is statistical in nature, but certainly a great deal of it is and many studies which are primarily analytical in purpose must utilize statistical data whether from primary or secondary sources.

---

[2] The Use of Social Research in Federal Domestic Programs, a Staff Study for the Research and Technical Programs Subcommittee of the Committee on Government Operations, April 1967.

SUMMATION

The foregoing merely emphasizes our concern about the problem and our awareness of the complexities and increasing problems connected with developing an improved system for coordinating and integrating our Federal statistical activities. Because we have not researched or studied this problem in depth we are not in a position at this time to suggest specific ways and means for solving the problem, but hope that our comments and observations will further emphasize the need for improving the organization of the Federal statistical program.

We are inclined to agree with Mr. Bowman in his statement to this subcommittee last year when he said:

> We must redress the imbalance which assigns so much more resources to information gathering than to the effective organization and analysis of information.

He added:

> While it is natural for data collection to cost more than their processing and the thinking about their improved organization, the difference today is much too large.

Difficulties Encountered by Statistics Users

The committee indicated an interest in this area.

In general, there are three perennial basic problem areas for statistics users. These are: (1) Lack of timeliness, (2) need for more geographic detail, and (3) the need for more subject and product detail. Additional perennial problem areas are overlapping series and lack of continuity in series of basic data. Admittedly, improvements have been made and are being made in some of these areas, but more needs to be done. As one of our members has pointed out, one of the difficulties involved "has been the lack of full financing of ongoing statistical programs. In addition, there has been very little budget available to the agencies producing major statistical series to work on the series improvement."

Without meaning to skirt this question, I think the best way to emphasize our difficulties is to refer to the specific recommendations which we have repeatedly made for improvements in Federal statistical programs. This is so because our concern about improvements arises out of difficulties which we encounter. As improvements are made, our difficulties become less. The most recent specific outline of our priority needs for improvement will be found in FSUC testimony before the Joint Economic Committee in its hearings on the Economic Report of the President for the years 1964 and 1965.

By the same token, this subcommittee's examination of views and suggestions for "Improved statistics for economic growth" serves the same purpose in identifying difficulties.

However, I would like to direct attention to one major difficulty which users have that is not program oriented. That difficulty arises when users attempt to compare, relate, match and link data that cross agency boundaries. Researchers and administrators have for many years been making or attempting to make comparisons of this kind on

a manual basis and on a limited scale. Sometimes successfully, but more often not. The growing use and capabilities of computers, as well as the availability of new and expanded data, have stimulated an increased interest in and demand for making these comparisons.

The computer provides a means for pulling together a great deal of information in disaggregated form from different sources which can then be correlated and aggregated into new forms for analytical purposes. To do this successfully, the input data must be reasonably compatible, otherwise the output will be highly distorted and sometimes meaningless.

Among the things that make comparisons difficult are the uneven quality of data from different sources; substantial conceptual and definitional discrepancies among agencies collecting similar and related information; the lack of uniform practices concerning descriptions of data and methods used; and the absence of uniformity in disclosing technical and procedural aspects of the data collection and estimation process.

There is a real need for an intensive study of this problem to determine first the types of data from various agencies that are most used or needed for comparative purposes, second, to identify those factors that make comparisons difficult, and third, to find ways and means of establishing an instrument or mechanism for resolving the difficulties, where they can be resolved. We recognize that data collected for statistical series of various agencies are for different purposes and objectives that may require somewhat different concepts and definitions but our feeling is that there are cases where greater reconciliation and comparability might be effected.

The Federal Statistics Users' Conference plans to give greater attention to this problem and particularly to develop information regarding the first two needs mentioned above.

Another difficulty which users have results from the amount of revisions in data that have been taking place, particularly with regard to the monthly estimates of construction and retail trade. To those users who follow these series, it becomes an almost impossible task to make a forward estimate based on the latest information available because the latest information continues to be revised. It is particularly troublesome when substantial revisions are made.

Dr. Morton, who acompanies me today, has identified several additional difficulties that users encounter which I am sure he will be glad to discuss with you if you have any questions. These are as follows:

"The frequent necessity of approaching and negotiating with several agencies in connection with what is essentially one and the same data request.

"The inability to provide services needed to satisfy data requests, especially by minor regulatory and small agencies which are not equipped to perform such statistical services.

"The absence of a comprehensive and interagency uniform data inventory and indexing system. And, related to the above, the absence of a uniform and general data specification format and procedures for use for data requestors.

"The lack of uniform practices with respect to the preservation of records and to access to past information."

Responsiveness of Agencies to Users' Needs

As an organization, FSUC has received excellent cooperation from Government agencies in the way of keeping us informed of the details and status of current and proposed statistical programs, hearing our unsolicited views on problems or suggestions for improvement, calling on us for our views on statistical problems, and in working with us at our meetings by providing speakers.

FSUC has undertaken a series of 1-day conferences on current and proposed research and statistical programs of various Federal agencies. Thus, far, we have held conferences covering the work of the Office of Business Economics and the Bureau of Labor Statistics. On June 1, we will hold a conference on the research and statistical activities of the Federal Reserve System. Each of these agencies has been most cooperative and responsive in working out the details of the programs and in providing speakers.

These conferences are proving extremely valuable not only in helping our members become better acquainted with the programs of the Federal agencies, but in understanding the problems and difficulties such agencies have in meeting user needs. Also, agency officials find them useful in getting better acquainted with user needs and interests.

Many of our individual members have personal, close contacts with officials in the statistics-producing agencies and by and large have found them more responsive and cooperative with regard to information requests, particularly by the more specialized ones such as the Census Bureau. However, the data users presentation of needs to the Government is usually on an agency basis, and there is, of course, the all-pervasive problem of shortages of resources which, in fact, leaves it with a particular agency to comply or not to comply with a specific request.

Another aspect of the question relates to the extent to which agencies have introduced new, expanded, or improved data in response to users' needs. To accomplish this it must be established that there is sufficient need for the data and that the program will be designed to provide optimum usefulness for a number of purposes. This is one of the criteria for evaluating statistical programs which is spelled out in FSUC's "A Long-Range Program for the Improvement of Federal Statistics." The next factor is the ability of an agency to provide the needed data with its current resources. But more often, it requires additional resources that must be obtained through increased budgets. Accomplishing the latter is probably one of our greatest stumbling blocks.

Thus, satisfying user needs for statistics is a tedious, time-consuming job, and changes are not accomplished overnight. Even after approval and with funds available, the high quality standards set by the statistics-producing agencies, technical difficulties involved in setting up a program, and limited availability of skilled professional staffs combine for a long lag between the time an improvement is accepted and when data are published. A lag of 3 years is close to minimum; a gestation period of 5 years is not uncommon even with persistent efforts like those of our organization. Without continuous pressure, changes in Federal statistics frequently take a decade or more.

Nevertheless, since the inception of FSUC in November 1956, a number of statistical gains have been made in programs that were advocated and supported by the conference.

Coordination Between Producers and Users of Statistical Data

We believe there is much better coordination now than ever before between the producers and users of statistical data. Two major organized instruments exist to achieve such coordination. They are the Federal Statistics Users' Conference and advisory committees of the Federal agencies.

A number of our members are serving on the various industry and other advisory committees of the Federal Government. They are serving on 14 committees in four Federal agencies. This overlapping membership is advantageous in that it provides each group the opportunity to be apprised of the interests and activities of the other.

However, we have observed that certain advisory committees are utilized more frequently and more effectively than others. We believe that in certain instances greater and more effective use could be made of advisory committees.

As for our own organization, our difficulties involve limited resources which means restricting our activities to emphasizing major priority items. We are doing our best to expand our resources, to expand our activities, and to be as effective as we possibly can.

Proposals for a National Data Center

The proposal to establish a national statistical data center is one of several instruments that have been recommended to improve the organization and coordination of Federal statistical activities. While a great deal of attention is being directed toward this proposal, it is one that needs much more study and understanding.

At this time, FSUC has taken no position for or against the establishment of such a center. We have made every effort to inform our membership about the details of such proposal and to keep them apprised of discussions and developments connected with it.

Last year, our newsletter of September 7 gave a comprehensive report on the hearings held by Congressman Gallagher on the possible invasion of privacy through a national data center.

At our annual meeting last October, we held a special session devoted to the subject of a national data center. It was the most highly attended of any of our special sessions.

When the Kaysen committee report was issued, a copy was sent to each member organization in the conference, along with a summary guide and outline to the report.

Our newsletter of April 10, 1967, summarized the testimony of Charles J. Zwick, Assistant Director of the Bureau of the Budget, at recent hearings of Senator Long on the invasion of privacy.

Our board of trustees has appointed a special committee to study the national data center proposal. Thus far, it has held two full-day meetings and has scheduled a third meeting on June 2. More meetings will be held. As a starting point, the committee is examining in great detail the Kaysen committee report.

The committee has appointed a special subcommittee to study and report on shortcomings of the Federal statistical system. Another

subcommittee is considering possible functions of a national data center. Other subcommittees are to be appointed, including one to study possible benefits from such a center. Obviously, considerable attention will be given to the question of the invasion of privacy. The committee has already prepared a partial bibliography of articles and books dealing with this subject.

We can only say at this time that we are giving thorough, detailed consideration of the proposal. The objectives of our organization and obligation to our members requires that this be done. We expect that our committee will arrive at some well-thought-out conclusions and make constructive recommendations.

CONCLUSION

Mr. Chairman, the essence of this problem was summed up by the Office of Statistical Standards in its special study, "A Federal Statistics Program for the 1960's," prepared for this subcommittee in 1962. It said:

> The decentralized character of Federal statistical responsibilities is a source of strength and efficiency on one hand and, on the other, of weakness in our statistical system. . . In order to take advantage of the strength of this decentralization and minimize the effects of the weaknesses, central coordination and planning is imperative.

We certainly agree. We recognize, too, that continuing effort is being made in the way of coordination and planning and progress is being made which is evident from improvements in programs and the reorganization of certain statistical activities.

The increasing magnitude and complexity of our statistical activities has greatly expanded our need for, and problems connected with, strengthening and making more effective our efforts to achieve a coordinated information system. Finding a solution will not be easy, but this should be a high priority item on the agenda that will require much more intensive review and study. We appreciate the opportunity to present our views on this timely and important subject and promise our fullest cooperation in the future.

Chairman TALMADGE. Thank you very much, Mr. Aiken, for your very fine statement.

In the interest of time I think it might be well to proceed with the direct testimony of Prof. Frederick F. Stephan and then if both of you will stay available we can interrogate you both at that time because I think there will probably be a good deal of overlapping in the questions and in your testimony.

So, Professor Stephan, you may proceed as you see fit.

STATEMENT OF FREDERICK F. STEPHAN, PROFESSOR OF SOCIAL STATISTICS, PRINCETON UNIVERSITY

Mr. STEPHAN. Mr. Chairman and members of the subcommittee, I wish to thank you for inviting me to discuss with you some of the problems of making the Federal statistical system more useful. I appear as an individual and will speak out of my own observations and experience, rather than as a representative of any group or organization. Nevertheless, I believe my views are widely held among well-informed statisticians and users of statistics.

My experience has included direction of a Bureau of Social Research in Pittsburgh concerned with problems of population growth, urban development, health and welfare; service as full-time secretary-treasurer of the American Statistical Association and participant in numerous conferences and committee activities involving the improvement and coordination of statistical data; membership on the Central Statistical Board; director of statistical services in the War Production Board and the War Manpower Commission; operations analysis on the staff of the Army Air Force Evaluation Board in Europe; membership on the Board of Trustees of the Rand Corp. (which is an organization that pioneered in the greater use of mathematical and related techniques in the solution of new problems in the defense of our country); and independent statistical consulting for various private organizations, business firms, and government agencies.

From this wide-ranging experience in the production and use of statistical data I can summarize the major lessons I have learned in the following conclusions:

(1) Citizens of our great country have been well served by the statistical work of agencies of the Federal, State, and local governments. Individuals and families, as well as businesses and community enterprises, have benefited greatly from the use of statistical data. Statistical information about conditions and trends provided a foundation for better plans and provisions to meet their needs and to regulate the administration of services starting a century ago with marketing farm products and with sanitation, public health measures, streets and highways, and various municipal services. The beneficial uses of statistics have progressed in our time to medical research, prevention of economic depressions, efficient management of both public and private enterprises, and innumerable other phases of our national life.

(2) Public confidence is essential to the job of obtaining reasonably complete and accurate data of any sort. People rather remarkably do provide information generously even when they do not see any direct benefit to themselves. Nonetheless, the burden of requests for information falls heavily on some respondents. Over the years they have had good reasons to complain.

I might interpolate that in 1938 I was executive secretary of a committee of the Central Statistical Board that looked into the burden of paperwork and requests for data. Subsequently there have been other reviews, notably by the statistics users of the burden of reporting.

Now there is a wave of alarm at certain illegal or unethical practices which are not employed in statistical work but which threaten to undermine the readiness of individuals to provide statistical information. Statisticians have observed the same ethical standards that physicians and lawyers recognize in respect to confidential information and will continue to do so, since we are aware in our own lives of the importance of privacy and we know that only by safeguarding data obtained for statistical purposes can we continue to get the cooperation of those who report it to us. New and stronged steps must be taken to assure the public of the integrity of these safeguards and the protection of what is private and confidential.

(3) Many advances and improvements have been made in the statistical work of the Federal Government and many more are yet to be realized. It is difficult to resolve conflicting interests and to accomplish what clearly is desirable. In part this is due to resistance from

persons and groups who think their interests will be affected unfavorably by proposed changes. Much of the difficulty is due to the complexity of the matters to which the data refer. Much of its is due to the fact that data can only be produced at a considerable effort. Accuracy is not always easy to attain. This is true for the unemployed worker, the grocery store manager, the trade union official, the school superintendent, as well as for the controller of a large corporation. Statisticians have worked hard at lightening the burden of reporting data but more can be done to reduce the cost to the respondents of providing needed data.

(4) Through correspondence and conferences, statisticians have made great efforts to determine the needs of users of statistical data and to meet these needs as well as possible. Again and again it turns out that what one user needs is not what another desires. Consequently there are many compromises in selecting and defining the items of information to be obtained and in determining what tabulations and summaries will be produced. Even when the users are themselves the respondents, there are conflicting demands as to what information will be collected, how often, by what means, and for what kind of resulting compilations. Many of the difficulties encountered by users stem from this diversity and incompatibility of their needs. It is clear that government agencies could meet the needs of users still better by greater ingenuity and diligence in finding the best compromises, by obtaining larger appropriations and additional staff with special competence for statistical work, and by winning greater confidence and cooperation from the people who supply the original data. Administrators who need data collected by their own agencies enjoy first priority and users outside government service tend to get what is a byproduct of meeting the needs of those who pay the bill, hire the staff (and provide other incentives to the respondents for filing their reports. Congress may well consider giving greater support in appropriations to the needs of users not in government. This is as worthy of subsidy as preferential postal rates, tax exemptions, and other means of promoting the general welfare.

(5) There is not one best way to organize the statistical work of the Federal Government. I think the organization of the statistical work of the Government reflects the organization of the whole Government and should serve the purposes of the whole Government so that this dependence on the larger problems of organization is a very reasonable and appropriate one.

Certain parts can be centralized but most of it must be located close to the administrative functions it serves and the operations which provide the original reports. A succession of studies have agreed that we should not concentrate statistics in a central bureau. In the light of this consensus, the proposal of a national data center would not appear to be wise if it is to be a consolidation of statistical functions that can be done effectively by the separate agencies engaged in the production of statistical data.

If, on the other hand, it is to be a gradual development of interagency cooperation, it could contribute very substantially to the beneficial use of statistics. The form of such an interagency arrangement is, perhaps, less important than the fact of its acceptability to all involved and the skill with which it is managed.

(6) The proposal of a national data center is motivated by the desire of many economists and others to obtain data in detail appropriate for their scientific uses both in and out of government. It is also motivated by a desire to exploit the new computer technology promptly and economically. Many agencies are already using computers for statistical data processing. We can anticipate both advantages and disadvantages in concentrating dispersed statistical operations, as it evident in business and government experience with other types of operations. The appropriate pattern and degree of concentration should be determined by systems studies and cost-benefit analyses developed progressively over such period of time as may be necessary to accomplish them soundly. My guess, therefore, is that a limited and more deliberate program of combining data from diverse sources would be more productive than abrupt establishment of a central agency with functions not fully tested or widely accepted.

(7) Coincident with such a progressive development of a centralized data agency, it would be very worthwhile to proceed with certain other closely related developments.

I would recommend:

(*a*) A counterpart of the Council of Economic Advisers to be concerned with intensive study of noneconomic problems paralleling those economic problems to which the Council and the Joint Economic Committee have so effectively addressed yourselves. This would be another major user of the statistical product of the Federal Government. I might add at this point that only yesterday did I learn that a bill has been introduced by Senator Mondale of somewhat similar nature, and I offer this suggestion, on perhaps a somewhat broader basis, aimed at meeting what seems to be a seriously unfilled need for a thoroughgoing study of noneconomic problems.

(*b*) A national statistical index and library to serve users as the indexes of medical and legal literature serve their users.

(*c*) An interagency arrangement for data analysis which would facilitate the sharing of improved methods and facilities and permit cooperative work on common materials. It could also enhance the utilization of specialists by assigning them to temporary duty on major problems wherever they are located.

(*d*) Joint statistical services for obtaining and processing data when either the sources of the data or the users are benefited by such combinations of related statistical operations.

In a word I am recommending to you that improvement of the Federal statistical system may be achieved by a wisely instituted data center but that it warrants further thinking and experimentation to enhance the chances of its success, enlarge the gains to be derived, and make sure that it attains a high level of public understanding and confidence.

Chairman TALMADGE. Thank you. very much, Professor Stephan.

With your background and experience we are honored indeed to have the recommendations and advice.

My first question is for Mr. Aiken.

Do you have any specific recommendations for improving the lines of communication between users and producers of Government data?

Mr. AIKEN. I can only say that we are trying to utilize as effectively as we can the two means that I mentioned. We have no committee to study this subject and we have nothing in our policy or programs

that seek to achieve this. Certainly it is something we should always give attention to.

Chairman TALMADGE. How are priorities established?

Mr. AIKEN. I am sorry, I don't understand the question.

Chairman TALMADGE. How could you establish priorities?

Mr. AIKEN. Priorities for improvement in various statistical programs?

Chairman TALMADGE. Yes, sir.

Mr. AIKEN. We have a long-range program for the improvements of Federal statistics which outlines priority items that we have given attention to and our committees also, in the examination of problem areas, try to establish priorities. It is a continuing process within our organization—our board of directors, our committees, and our meetings highlight those items that they think deserve priority.

Chairman TALMADGE. When the public agencies decide to change the definition of a particular series, is there any mechanism which guarantees a full and effective hearing of the interested parties?

Mr. AIKEN. May Mr. Friedman answer that?

Mr. FRIEDMAN. My major experience is limited to the BLS, Mr. Chairman, and there the Bureau does work rather effectively through a system of advisory committees, mainy two. One is a Business Research Advisory Council on which representatives of the business community serve and another one is the Labor Research Advisory Council on which representatives of the labor movement serve.

Most of the activities of the Bureau are discussed with those committees, hopefully in time for those committees to make their views felt in advance of the decision, and from my experience I would say this is a rather healthy approach to getting the views of the users and getting the attitudes, getting the experiences of some of the outside expert users in the statistical programs with which the Bureau is concerned. I would heartily endorse this kind of activity for a number of reasons.

Chairman TALMADGE. Do you think the relationship between the Government agencies and users is relatively good in the field of statistics?

Mr. FRIEDMAN. I can't talk for all of the agencies. I think there are obviously certain weaknesses. I suppose this may be due in part to some of the users as well as to some of the agencies.

It is my observation that if you decide that you want to get your views felt or heard in many of the Government agencies there are ways to get this accomplished.

What I am saying about advisory committees, however, I think this provides a more orderly form through which this can be accomplished. I can say that so far as the Bureau of Labor Statistics is concerned, my feeling is that the views of the users, and I think I am safe in speaking for the business community on this as well as the labor community, I think the views of the users are heard. I would hesitate to talk with any degree of confidence about many of the other agencies.

Mr. AIKEN. Mr. Moor will be glad to comment on that question, too, Mr. Chairman.

Mr. MOOR. Mr. Chairman, we all have our specialties and therefore none of us can speak about all of the agencies, but I have had contact with a number of them since going with a private business firm and

my impression is that those users who tend to specialize in particular fields have good communications with the agencies involved and that is a two-way street, in that the agencies come to the users to discuss possible changes in the nature of the data as well as the users expressing their own concern about difficulties.

As a matter of fact, I think that something that is emerging now in the use of statistics is simply that we are getting sufficiently sophisticated so that we are changing the nature of the data, improving them, I think, in almost every instance, but improving—changing them sufficiently frequently so that it is a little awkward to maintain a series over time. I don't want to be placed in the position here of arguing against improvements, but I wish that they would come along perhaps a little less frequently than they do.

In respect to an earlier question, in fact, that first one that you asked, I think one further point should be made.

You asked, if I understood you correctly, how improved communications might occur between the Federal agencies providing data and the users of the data and I think one potential answer to that question is through the medium of a national data center. We are not in a position in the Federal Statistics Users' Conference to take a specific position on a data center at this time. But certainly one of the functions which it can provide is as a medium or a channel by which users could be placed in connection with, in contact with the producers of statistics.

Mr. FRIEDMAN. I just want to comment, Mr. Chairman, in connection with the question of the relationship of users to the agencies and how this is sometimes informal—without mentioning agencies—a year or so ago a problem arose in connection with a statistical program and we got in touch with the agency and explored it with them. We had very little contact with them before, but in the process of exploring it, the suggestion was made that it would be a pretty good idea for the agency to have an advisory committee on the statistical program and the person with whom we were talking said it was a wonderful idea, the point being that in many cases these things just aren't thought of until the problem arises directly.

Chairman TALMADGE. Yes, sir.

Mr. AIKEN. I can give another recent example. One of the agencies came to me and said that they had assembled a great deal of statistical data from various sources and consolidated them in a group of tables for their own internal use and they had a feeling that this might be valuable to those outside of the Government if they were put in published form. They asked me if I could give them a list of our members. They didn't want to solicit all of them, but would contact a selected number and tell them what they had in mind in order to find out if it would be valuable to them to put this out in published form. They came to us and I think the communications were good. And we were glad to be able to participate in that regard.

Chairman TALMADGE. What are some of the major uses of Government statistics on the part of business firms and labor organizations?

Mr. MOOR. Mr. Chairman, there is one very simple and very direct answer to that which I discovered very quickly after I got into business.

I would say the great overwhelming use, at least by users of Federal statistics of all sorts, is for the purpose of forecasting. The economist, at least as identified in business, is one who is considered to know some-

thing about the future even though he doesn't, and he tries to do the best job he can by providing information forecasts based on current data. It can also be used in a number of other ways, but I think these others are subordinate.

There is another organization which is not directly related here and that is the National Association of Business Economists. I have done a lot of talking to fellows in this group and there is no question about the use of statistics and the use of forecasting.

Chairman TALMADGE. Can you give us some illustrations where a more coordinated and integrated system would better serve the needs of business and labor?

Mr. MOOR. I will give a very generalized answer, Mr. Chairman, by saying while I didn't do as much homework as John Aiken and some others in preparation for this statement, I know that in prior hearings of this subcommittee during the last 2 or 3 years, a number of specific examples have been given.

There are some examples that come to mind quite readily. There are differences in some of the GNP concepts where we begin to talk about—some of the newer concepts, as a matter of fact, on consumption, for example, as contrasted with personal outlays. I think without going into detail on these, these are justifiable distinctions but that the conceptual nature of these distinctions needs to be clarified more to the users, and I have no proposal at this time as to how that clarification can be provided, although I would say hearings such as these are very useful.

Chairman TALMADGE. Congressman Bolling?

Representative BOLLING. Professor Stephan, I would like to pursue your recommendations a little to make sure I understand them.

Your recommendation (a)—and I think you mentioned it was somewhat along the lines of a bill recently introduced by Senator Mondale?

Mr. STEPHEN. Yes, sir.

Representative BOLLING. Will you be a little more explicit? What might be the title of this particular group of advisers? Would they have a counterpart in the Congress in an overall committee? Is it a council of advisers on social problems or what? What would be the concept?

Mr. STEPHAN. I think there is considerable room for shaping such activity and it would be most effective if it did have a committee of the Congress with which it worked and to whom it reported its results initially, along with reporting them to the President and the Cabinet.

What I had in mind is this: There are many problems that are not thought of originally as being essentially economic, although they may have economic consequences. For example, crime is a problem in our life today. In the Wall Street Journal a couple of days ago there was a story that members of criminal syndicates were establishing themselves in legitimate businesses and this could have very serious consequences.

Now, I think that one hardly needs to think very far on a problem like this before he sees that we need to have some facts to tell us whether this is a scare story blown up out of all proportions or whether it is, in fact, a very serious threat to our economic stability and to those legal and ethical problems that arise in business in various instances, but would be greatly aggravated, including the corruption of public officials.

I am sure that much of this problem is a problem for local government and for State governments, but I am also sure that the Federal Government has a major concern with a problem of this kind. Therefore, while it might be investigated by a special committee or agency, a continuing study of this type of problem would be appropriate. It would be an appropriate matter for statistical study as well as legal and other modes of study, and the results of such intensive studies, using the best scientific procedures we have available as well as our common body of procedures in the law and otherwise, would contribute toward the control of such situations and possibly to their solution.

That's only one particular example. We could find many in the area of poverty and in the area of education. I am sure that education is a major matter of economic significance. As a matter of fact, as a capitalist country we are making a capital investment in the education of our population which ranks with any other kind of capital investment we are making. Our strength and preeminence in the world is in many ways a reflection of that wise decision many years ago to have an educated citizenry. There are many problems connected with education that are not economic, and not problems of the international implications of education, but which have to do with the very basis of the lives of individual human beings and their parents and families. I don't look on the Federal Government as having responsibility for regulating the lives of individuals and families, but I think that insofar as general conditions may create serious problems and difficulties for families and individuals, the Federal Government has an interest. The soundness of our society and our economy depends upon the soundness of our family life and community life and individual life.

I could go on. I don't have much sympathy for the attempt to dress up all of these questions in the form of social issues and social problems. I would like to see them also as scientific problems. It is not only important that we have equality of opportunity for our citizens, but that we make the most rapid social advances along with economic advances. We need inventions in the social sphere just as much as we have needed them in the spheres of manufacturing and weaponry and technology generally.

As a matter of fact, I think we are at the beginning of an era where we will turn some of our best brains to the problems of how to improve the conditions of human freedom and human life, apart from those on which we have put the major emphasis so far.

Representative BOLLING. The concept would be, in effect—if I may interject—really quite a lot like the Council of Economic Advisers and Joint Economic Committee in a different field in that those two entities were instituted as groups having much broader jurisdiction than any single department of the Executive or any single committee in the Congress, and thus, were in a position to take a broader and more—if the word is correct—integrated view of problems which are extraordinarily broad and complex.

Mr. MOOR. I don't know what Professor Stephan has in mind exactly, but I would call your attention to a particular new type of innovation which is being introduced at the local level. The best example I know of is in New York City, where a new group now reports to Mayor Lindsay. This group, which is essentially an operations research group, makes used of a variety of scientific techniques and applies these to broad city-social problems, as well as other things.

Representative BOLLING. Thank you. Now, with regard to your recommendation (b), for a national statistical index and library. Where and under whose jurisdiction?

Mr. STEPHAN. I am perhaps a skeptic about the possibility of finding the best place to put any activity, because I strongly believe what may be a good place with one type of people performing the activity may be a bad place for a different type of people.

I think there are many good places to put such an index and library. One of them that would be suitable would be in the Office of the Statistical Standards which already has connections and communications with the statistical agencies of the Government. I would be very happy to see this in the Library of Congress, which I think also has many of the facilities and the traditions that would make it operate effectively.

I would like to see the possibility explored of having it developed so that there would be in all major cities an office to which users could go and get information. I am cognizant of the fact that some of the statistical agencies now do perform this function through offices in various major metropolitan areas.

So my answer can only be I would have great confidence in the good sense of the Congress to find one of the very good places to put it and wouldn't worry too much about the question of where would be the best place.

Representative BOLLING. You would be satisfied with the good sense of Congress but you would not argue that one was the best.

Mr. STEPHAN. I could conceive of the possibility that what is a good place now turns out in 10 or 15 years to be an inappropriate place and that changes in location may be made from time to time.

Representative BOLLING. I would like to be as confident as you seem to be that Congress would pick a good place. What about the other two? I would like some indication as to location.

Mr. STEPHAN. The other two recommendations are purposely stated in a somewhat vague form because I could see them as consisting of a sort of federation type of arrangement, rather than a type of formal organization. If we could get the statistical agencies to come together as they do to some extent in the Office of Statistical Standards and they did under the old Central Statistical Board, to work out agreements and to staff it for a time as an ad hoc group to carry out a function, I think that is most likely to meet the needs of the particular problem at the particular time.

Representative BOLLING. I am glad to see that you seem to be a little bit the same way as I am. A sort of anti-institutionalist.

Mr. STEPHAN. I am happy to find somebody who thinks the same way.

Representative BOLLING. Thank you very much.

Chairman TALMADGE. Getting back to Mr. Aiken, do you feel that there is much duplication in the information requested from the public?

Mr. AIKEN. Well, we have been giving some thought to this in this special committee which we have and we have not tried to identify areas of duplication. But it certainly is something we are going to give consideration to.

In this connection, at our recent meeting, this was the subject of considerable discussion. Our committee recognizes that, while duplication does exist to some extent, it is sometimes unavoidable and that

there are cases which exist because of different needs which must be served for specialized purposes. How to identify all the areas of duplication is quite a task and that all agencies should continue to be alert to the problem of avoiding duplication. It is a subject that our committee will give a lot of further consideration to.

Mr. Moor would like to add something to that.

Mr. MOOR. Mr. Chairman, I think my answer to your question would be, essentially—and I am sure I will be in trouble with this with my business colleagues—that the amount of duplication is not great. The amount of complexity, particularly for specific types of studies, can be immense, but the amount of duplication I don't think, as such, is great.

Chairman TALMADGE. Are there any estimates of the costs incurred by the public in complying with requests for data?

I frequently have small businesses telling me they have to keep two or three secretaries working on such information all the time.

Mr. MOOR. I don't have any estimates of cost. But I would like to qualify my last answer, particularly with respect to small businesses, since you have raised it.

I think there is substantial duplication as far as small businesses are concerned, particularly as between various levels of government—Federal, State, and local. The classic example is particularly with respect to alternative types of tax collection and alternative types of classification.

Mr. FRIEDMAN. Dr. Moor just touched on a point that I wanted to make that I think we have to separate here the reporting systems that flow out of legal requirements, that is, tax laws, whether they are local, State, or Federal. In many respects the complaints that I have heard relate to this kind of problem. I think we have to make a distinction between that and the collection of data by the Government on the basis of voluntary responses.

Chairman TALMADGE. Do you think it might be wise to have some provision for notice or hearings before the Government changes its procedures, systems of reporting, and so on?

Mr. FRIEDMAN. I am not sure I follow you.

Chairman TALMADGE. Do you think it might be wise if Congress passed legislation providing that Federal agencies give notice and an opportunity to be heard to the users and those who provide the data before an agency may change definitions or collecting procedures? Something similar to the Administrative Procedure Act.

Mr. MOOR. Mr. Chairman, first of all, I think, in general, this is done extensively by the individual agencies at present. This is not done, obviously, with all the users or reporters of data, but it is done with samples of them. As a matter of fact, a genius frequently conduct hearings at which they work out mutually satisfactory arrangements. It is done quite effectively at present, especially by regulatory agencies, FCC, FAA, and so on. Obviously, any time any changes are made people are going to be discontented because that is the nature of change, Generally, the amount of communications with the responders is quite good.

Chairman TALMADGE. To get back to Mr. Stephan. Hasn't the concept of a national data center already received intensive study?

Mr. STEPHAN. I believe a great deal of study has been given to the proposal of the national data center and there are several very excel-

lent reports that discuss major aspects of it. However, it is of such magnitude, and the Kaysen report, in particular, involves such a sweeping reorganization of all the statistical work of the Government in connection with the establishment of such a center that I feel it has not been adequately studied. It has not been adequately studied, I think, from the standpoint of defining rather precisely what types of information would be consolidated in such a center. We could readily imagine a collection of statistical data so huge that Congress would be quite unwilling to pay the cost of collecting it, let alone processing it, so huge that there would not be space enough, even for the reels of tape and the personnel engaged in processing and providing to users the results of such data.

On the other hand, if the scope of the activity of the data center is very, very small, then some of the dreams about its great contribution to the needs of users would fall by the wayside. I do not find an adequate statement of the extent to which the center would throw added burdens on the existing statistical agencies and require increases in their budgets.

If the user, for example, were to go to the national data center and ask for data that were not already available in precisely the form that he wishes, within the consolidated files, then the national data center would, through its personnel, go to a statistical agency or several and make requests for additional data.

This would then involve activity on the part of the staff of such statistical agencies to get the data available and even before they were made available, to find out what problems might be involved in making them comparable and matching them to the needs of the users.

I think in instance after instance this would develop the fact that the agencies did not have the data that was needed, and that new inquiries would have to be addressed to business firms or to individuals to obtain the data that is needed.

Chairman TALMADGE. It would make more work in addition to what has already gone on.

Mr. STEPHAN. Yes, sir. I would like to approach this much more realistically in terms of the total expansion and development of statistical work in the Government that would be involved if there is a great increase in provision of data for users of statistics to find out the extent to which users of statistics would finance these additional activities. Under some arrangement with statistical agencies the user now pays the additional cost of getting the special tabulations he requires. He does not pay the original cost of getting the data since that has already been provided by congressional appropriations. That is why I speak of a systems analysis and cost benefit study of this problem. Clearly there is an intermediate point at which we will get thoroughly desirable benefits from such a centralizing of important statistical data, but beyond which it would be unwise to carry the development at this stage of the demonstration. Ten years from now, 50 years from now the situation may be quite different.

Professor Kaysen's plan is in some ways a blueprint for long-range development that could well guide a series of future steps toward improving the statistical system in the Federal Government.

Chairman TALMADGE. Do you think that we might start off by having a pilot program, a pilot project in trying to find an optimum point where it might work and might not?

Mr. STEPHAN. Yes, sir; I think this is a very wise approach. It is wise in industrial development and it is wise in government.

Chairman TALMADGE. You don't think we ought to have a Fourth Commission study on the project, do you?

Mr. STEPHAN. I don't recommend that we postpone action, but I recommend that our actions start on what is a practical, realistic basis of action and proceed as we learn more about what can be done, what it costs, about what the needs of the users are, and then proceed to add to it.

I would like to make a great plea for the users who are only potential users now—who are not using statistical data but could if they were made available to them in a more easily obtainable form. I would include among them the mayors of our major cities, councilmen, major heads of departments, planning and zoning agencies, similar departments of State government, Governors, legislators, and many private individuals and business executives and their staffs who are not now taking advantage of what could benefit them a great deal, and thereby benefit the whole Nation by improving the soundness of our economy.

I am impressed by the need for forecasting but there are also needs for planning and for allocating and for measuring the efficiency of one's operations against a suitable standard provided by other members of the same industry, many other uses that can be improved to the benefit of everyone concerned.

Chairman TALMADGE. Do you think that a national data center would constitute a threat to privacy to the people of this country?

Mr. STEPHEN. On this point I think it could be a threat or it could actually operate to increase the protection of privacy of individuals, depending on how it was set up and operated.

Chairman TALMADGE. Are there adequate safeguards now?

Mr. STEPHEN. I think the safeguards we have now are very substantial. We have on the statute books laws that provide penalties for Government employees who disclose private information about private individuals and firms, as well as laws concerning conspiracy and fraud and other activities on the part of persons who may not be governmental employees.

Possibly, with the change in price levels, some of the limits placed on the fines might be elevated and that a stronger statement to the public is very much needed, pointing out the nature of these safeguards and reporting the experience of the Government with respect to them. I have not seen any summary that tells us how frequently these laws have been violated. It would be informative to the man on the street if he knew the extent to which these laws are effective in preventing the type of disclosure and invasion of privacy that has been a source of anxiety to him.

Chairman TALMADGE. Do you think a national data center would impair the cooperation of the respondents?

Mr. STEPHEN. It all depends on how it is set up and operated. You could have a wave of distrust of Government—perhaps exaggerated in a spectacular way by some author of a book or some newspaper columnist or somebody else who points out the possibilities one can imagine of the misuse of centralized data.

Historically the most notorious example is the use by the Nazi of recorded data for the persecution and extermination of Jews in Ger-

many and other persons. But on the other hand, I think if a national data center is operated properly and the public is kept informed of what is being put into the central files and how it is being safeguarded, it is possible to reassure even people who are more anxious than the average person about what the Government knows about them and what possibly false information is being put into files and may be used against them. Anyone who has been called up for investigation under the Security Act knows that there is information that he is not permitted to see that may be derogatory and at the same time may be false. Any of his enemies or people with a grudge against him may have fed information into the system and he had no way of knowing what it is and whether it is being taken seriously. I personally would trust Government officials to use their good sense and judgment in the interpretation of such material, but I can understand the anxiety of the man on the street who feels that he has a hopeless task of trying to defend himself against backbiting gossip and malicious slander. A centralized data file may be seen by him to be of that character. We have not only to assure him but give him good reasons for accepting our assurances that a centralized data file would not have that character.

One possibility is that you might provide that any individual who so wishes or any firm that wishes can get a printout of what is in the computer concerning him. He would have the right to be told what is in the central file and to correct it if it is incorrect.

Chairman TALMADGE. Do you think that the present burden of requests for information falls too heavily on some respondents?

Mr. STEPHAN. Yes, sir; I think it does. We have been too much committed in statistical work to covering the whole group concerning which we want information. One of the great advances of the last 15 or 20 years, has been the development of sampling methods, very soundly and scientifically set up to extract accurate data from a very small fraction of the total population and provide estimates, and forecasts if you will, of considerable dependability and reliability. The prime example you have is the series on the unemployment rate, based on 50,000 households. This is a major indicator watched by Congress as well as by the Council of Economic Advisers, by business executives and others, based upon a fraction of less than 1 out of 1,000 households in the United States.

It needs further development because it can relieve populations and business firms of our country of a good deal of reporting. It does not meet the needs of localities for detailed data about their own populations and businesses. But this is the point at which we most need constructive work by statisticians and public officials and by the respondents to achieve the kind of information that is needed with a minimum necessary burden on the respondent.

Chairman TALMADGE. We see published from time to time the so-called leading business indicators, and then you will find trained economists differ on these views as to what they mean. Can we take statistics, for instance, and project, 30, 60, or even 90 days from now what business conditions are likely to be?

Mr. STEPHAN. Well, sir, this is a fundamental problem to which the statisticians have addressed themselves. In the progress of our statistical competence we have increasingly stressed the fact that the farther you make a projection the less certain it is and the greater variation

you may expect in the actual facts when they are exhibited and the events have occurred. This is a point on which the users of statistics are still in need of further development in their thinking. If we take projections as being approximate, as being less dependable the farther we put them out we become aware that we need the best possible set of data on which to base them because we can improve the accuracy of the projections by having good, sound, accurate data about the present and immediate past of recent facts. But we can never get the full story until the future has been exhibited, recorded, and incorporated into the statistics. Therefore, we must always think in terms of a band, if you will, or a spray of possibilities within which we can expect the actual events to occur.

Sometimes we have an unforeseen development that makes a tremendous change. The most outstanding of these, I think, was the baby boom that occurred after World War II. It had a profound effect on our school system and it is having profound effects on the labor force and our whole economy. I suggest to you, as you certainly have observed, that it may have profound effects on our political life in years to come.

Chairman TALMADGE. In other words, these statistics give you the tools with which to make an educated guess; that is about right?

Mr. STEPHAN. Yes, sir. It remains a guess, but by being based upon the best information we could get it becomes a more dependable guess.

Chairman TALMADGE. Thank you very much, gentlemen, for your contributions.

This concludes the hearings today and we will stand in adjournment until Wednesday, June 7, at which time they will continue.

(Whereupon, at 11:35 a.m., the hearings recessed to reconvene on Wednesday, June 7, 1967, at 10 a.m.)

THE COORDINATION AND INTEGRATION OF GOVERNMENT STATISTICAL PROGRAMS

WEDNESDAY, JUNE 7, 1967

CONGRESS OF THE UNITED STATES,
SUBCOMMITTEE ON ECONOMIC STATISTICS OF THE
JOINT ECONOMIC COMMITTEE,
Washington, D.C.

The subcommittee met, pursuant to recess, at 10:07 a.m., in room 1202, New Senate Office Building, Hon. Richard Bolling presiding in the absence of Chairman Talmadge.

Present: Representatives Bolling, Rumsfeld, and Curtis; and Senator Miller.

Also present: James W. Knowles, director of research; George R. Iden, staff economist; and Donald A. Webster, minority staff economist.

Representative BOLLING. The subcommittee will come to order.

This morning the Subcommittee on Economic Statistics of the Joint Economic Committee begins its third day of hearings on the subject of the coordination and integration of Government statistical programs. I wish to insert in the record as an appendix to these hearings the proceedings of the recent conference organized by the Washington chapters of the American Statistical Association, and the American Marketing Association, on purposes and uses of Federal statistics. At this conference the ranking minority member of this subcommittee, Representative Thomas B. Curtis, presented a very excellent paper entitled "Strengthening the Tools of Economic Policy."

This morning's witness, Mr. Raymond T. Bowman, Assistant Director of Statistical Standards, Bureau of the Budget, is no stranger to this subcommittee, since he has been very helpful to us for many years. Mr. Bowman, we appreciate your coming to discuss with us the goals, achievements, and outlook of the Federal statistical program. I would like to interject a thought that this seems like "old home week" to me, because years ago I was chairman of this subcommittee and Mr. Bowman and I were sort of coconspirators in an attempt to improve the funding by the Congress of the statistical programs of the executive.

Mr. Bowman, we are delighted to have you here. You may proceed as you wish.

STATEMENT OF HON. RAYMOND T. BOWMAN, ASSISTANT DIRECTOR FOR STATISTICAL STANDARDS, BUREAU OF THE BUDGET; ACCOMPANIED BY MILTON MOSS, MARGARET MARTIN, AND MRS. ROSE CASSEDY

Mr. BOWMAN. Thank you, Mr. Chairman. It is a real opportunity for me to appear before this subcommittee again to discuss Federal statistical programs which you know are very close to my heart. I call

your attention, Mr. Chairman, to the fact that I have several members of my staff with me who might be called upon to assist in answering any questions, if necessary. They are Mr. Milton Moss, who works in the national accounts area, Miss Margaret Martin, who works in the labor force and population area, and Mrs. Cassedy, who works with Mr. Moss in the area of the national economic accounts.

I have a prepared statement which, with your permission, I will submit for the record.

Representative BOLLING. Without objection, it will be included.

Mr. BOWMAN. But, to save the time of the committee, I will try to summarize the important items in that statement orally rather than read it in the hope that this will provide more time for questioning if the committee wishes.

In introduction, I would like to call attention to the fact that the last two decades are decades in which major improvements have taken place in Federal statistics. I would like the record to show that in my opinion, this subcommittee has had a great deal to do with the improvements that have taken place.

I am particularly interested in, and strongly support, the present emphasis of the subcommittee on the need for strengthening coordination and integration efforts, and also the need for greater attention to the availability of the information in interrelated forms. Only in this way can the information be of maximum usefulness for economic analysis and for the broad problems of social improvement and the various action programs in that direction which the Government is undertaking.

In my presentation, Mr. Chairman, I would like to do several things. First of all, I would like to review some of the progress that has been made in the various statistical fields and do it in such a way as to make it possible for the committee to see the way in which my office attempts to coordinate and improve the integration of the various statistical programs. In doing this, I will call attention to some of the major methods which the office uses in achieving its purposes.

First, as you will see as I illustrated with various programs, the office utilizes the authority which is given to it in the Federal Reports Act and in the Budget and Accounting Procedures Act to shift statistical programs or basically statistical series from one agency to another when this seems desirable in the interest of a better development of a statistical series.

Secondly, I would like to call attention to the fact that another way which a coordinated and integrated statistical program is promoted is through the review of budget requests and the way in which the statistical programs are incorporated into the President's budget. This committee deserves considerable credit for having suggested the "special analysis" which is presented with the budget each year and in which the statistical programs are put together into a single document.

A third method which the office uses and which was also developed to some extent by this committee, is the use of outside task forces which bring expert guidance in the development of statistical programs. Internal task forces, interagency committees of one sort or another are also used to develop ideas concerning the necessary changes that should be made in statistical series and how they should be better integrated and made more consistent one with the other.

Another way in which the Office works to insure better coordinated data is through its review and improvement of the data requests which the various agencies address to the public.

I believe these are the major methods. I have outlined them more fully in the paper which I have submitted for the record.

I will now give some illustrations of how these methods work in practice, and in my final remarks, indicate some of the issues which I think must be faced in the near future.

I know that you will remember, Mr. Chairman, that one of our early efforts recognized that the best guidelines that we have for deciding what kind of statistics are necessary and in what form they should be produced to improve economic analysis were the national economic accounts. Here we have a logical framework which puts data together in such a way as to make it possible to examine the operations of the economy and to determine where the weaknesses are and what actions should be taken.

In order to get a greater insight into what is needed to be done in this area, you will remember that the Bureau of the Budget contracted with the National Bureau of Economic Research for a report in this area. That report was reviewed in this committee and many of its elements have now been implemented.

In addition to improving the basic flow of data into the income and product accounts, we have also incorporated into the activities of a single agency the responsibility for the general input-output accounting structure, sometimes called purchases and sales accounting. This type of accounting gives us a greater insight into the way in which intermediate products enter into the total output of the economy.

I think you will also remember that the United States had enjoyed early leadership in this area, had then lost it, and it was a considerable effort to bring it back and into focus so that it would be integrated clearly and concisely with the income and product accounts.

An input-output table for 1958 has been produced with 87 industrial sectors. A table for 1963 is in preparation. Our objective is for a table every 5 years—we hope a little more promptly after the close of each census year—but it is a complicated procedure.

It is also our hope that a table in between the 5 years can also appear as an interpolative input-output table between the 5-year intervals for the major table.

We also have paid considerable attention in the development of the statistics program to the need for better statistics in the balance-of-payments area. The balance-of-payments accounts are integrated with the national income and product accounts, the input-output accounts. Here we also utilized the idea of a committee to examine the work that was being done in this area and to make recommendations. This committee is commonly referred to as the Bernstein committee and its recommendations have been implemented to a considerable extent but efforts continue. We now have a complete set of reorganized tables on balance-of-payments statistics. We punish the estimate of the balance in two forms following one of the reconmmendations of the Bernstein committee. These bring out the balance in terms of what is often called the liquidity principle and the other the official settlements principle.

Recently we have published in the Statistical Reporter a general review of the progress that has been made in implementing the recom-

mendations of the Bernstein committee and I have directed my office to send to each member of this committee a copy of that article so that you will be up to date on what it is that we have done to date. (See p. 95.)

We are still working on incorporating into the system of economic accounts a set of accounts on national balance sheets and sector balance sheets. We hope that major benchmarks may be set up by work that will be done near the close of this decade. This illustrates another method which my office uses to improve coordination; namely, the establishment of focal agency responsibility in certain statistical areas.

In this area of wealth statistics and balance sheets we have asked the Office of Business Economics to be the focal agency to outline the work that should be done in the area of the development of national and sector balance sheets and the Census Bureau to take focal responsibility for the basic body of data that is required in order to implement a system of naitonal balance sheets and sector balance sheets.

I think this illustrates as well as I can in a few minutes, the way in which we are using the national income and product accounts, the input-output accounts, the balance-of-payment accounts, and we hope eventually the wealth accounts to guide and reinforce our efforts to develop a coordinated and integrated body of information.

I would not like to miss the fact that we have also paid particular attention to improving the consistency within the general system of accounts of one other system of accounts which stresses the money flow aspects of the economy and is called the money flow accounts, or flow of funds accounts. This is the one element in the set of accounts which is not centered in the Office of Business Economics but is in the Board of Governors of the Federal Reserve System. All the other elements of the accounts are in the Office of Business Economics of the Department of Commerce.

Another area to which we have been devoting a great deal of attention, but have not made as much progress as we had hoped, is the area of prices. This is an area in which we took early action by having a committee of nongovernmental experts work in this area, through a contract with the National Bureau of Economic Research, as early as 1959. The report of these experts was presented before this committee in 1961. Since that time we have achieved a revision and improvement in the Consumer Price Index (CPI), and also in the Wholesale Price Index, but we have yet to achieve certain other objectives which we consider very important.

One is a set of price indexes which will delineate our competitive position in world markets. It is commonly called export-import price indexes. We have not made the progress in this price area which we believe essential. Funds for this improvement are in the 1968 budget request, but have not been approved in the House report. The second is more attention to quality problems in price index construction. The request, I believe, has been supported in the House report.

We have made some progress in developing sector price indexes but we have not gone nearly as far as I had hoped we would have been able to go by this time. Sector prices really are the essential ingredients to better measures of real product by industry, because without appropriate price indexes by industry, it is impossible to make the necessary adjustments in order to measure real product by

industry. We will continue our work in this area with existing resources, but to make significant further progress in 1968 will require the funds requested in the President's 1968 budget.

Mr. Chairman, you have been particularly close to the work that we did in the early years with regard to labor statistics. There has been, of course, continuing attention to this area of statistics. There have at times been doubts as to the accuracy of the statistics on unemployment. We have been continually reviewing this area and today we probably have the finest body of information in the world on employment and unemployment, yet there are many problems, many areas, in which we still do not know the sorts of things we would like to know about the U.S. labor force.

Here again a committee commonly called the Gordon Committee appointed by the President at the recommendation of the Department of Labor and the Bureau of the Budget, developed an excellent body of guidelines and recommendations. Many of these recommendations have been implemented, but the report was of a character which gives us guidance for many years to come. The highlights of what has been done to date can be summarized something as follows: As a result of that report an experimental panel of about 17,000 households was set up as a basis for studying some of the conceptual problems associated with the measurement of employment and unemployment. As a result of that study, we recently incorporated new definitions, not significantly different from the older ones conceptually, but of such a character as to more specifically delineate the characteristics of people who are in and outside the labor force and make it possible for us to give more consistent estimates of various sectors of the labor force and their employment-unemployment characteristics.

We have also expanded the labor force survey from 35,000 households to 52,000 households. This expansion of sample size will make it possible to take the emphasis off the single figure, total number of unemployed or total number of employed persons, and give better data with greater accuracy with respect to the age, sex, color and composition of both the employed and the unemployed.

We have had little success in achieving a national program on job vacancies. This committee has had reports in that area and I will not dwell on it. However, experimental programs are still going on within the Federal Government and we would hope that sometime in the future that job vacancy series designed to meet both operating program needs and basic analytical needs could be added to our general body of statistical information.

In the areas of construction, production, and distribution, we have also been taking major actions to improve the interrelationship and validity of measurements. Here again, there are illustrations of the different methods which the office uses to improve statistical series. Construction statistics is one of the areas to which I devoted early attention when I first came to take charge of the Office. It had been clearly recognized for some time that one of the difficulties that the responsibility for this type of statistics was distributed among three agencies, the Bureau of Labor Statistics, the Business and Defense Service Administration of the Department of Commerce, and the Census Bureau. One of our early actions, therefore, was to consolidate this program in the Census Bureau.

With the consolidation of the program in the Census Bureau, efforts were made to improve the measurement of housing starts by reports on actual starts, to improve the measurement of value put in place by progress reports on value put in place, particularly in the nonresidential construction area. It is still true, however, that this is not the strongest element of our statistics program. It is a difficult area. It is diverse. It is costly to develop information which is both prompt and accurate. Recently we have had to revise the publication schedule in the interest of not publishing inaccurate data. We hope that as time goes on, we will be able to develop methods that will improve the accuracy of the data so that various elements in the program of construction statistics can be brought forward to more prompt publication.

It had also been recognized for some time that the current report on manufacturers sales inventories, and orders required special attention. This important series was on a very minimum budget in the Office of Business Economics. It involved the collection of data from enterprises rather than from the establishments of enterprises, so that it was not possible to get estimates that were consistent with other estimates for individual industries. Here again, we used the device of transferring a series, so far as the collection of the data was concerned, from the Office of Business Economics to the Bureau of the Census.

Work was then undertaken to expand the sample, to collect the data along divisional lines so that more attention could be paid to the industry characteristics of the information, and to tie it into and make it consistent with the annual survey of manufacturers. In this way a body of information collected monthly was made consistent with information available annually by our annual survey of manufacturers and also the quinquennial censuses.

I would not like to leave the impression that all problems have been solved in this area. This is not an easy area. It is a difficult one and I will mention one or two special difficulties which arose.

It was recognized that when we began to do more work in the national accounting area one thing that should be done was to develop ways of estimating real product by industry consistent with the overall GNP real product. We knew when this was done that the index of industrial production computed using somewhat different data and in a different agency might not agree exactly with real product by industry measurements within a GNP framework. We now are working diligently—Mr. Moss, who is with me now, is working with an interagency group and is also privately working as a scholar at the Brookings Institution in this area—to see what can be done to bring these two bodies of data closer together and give us more consistent measures of manufacturing real product by industry. This is an illustration of the type of problem with which my Office wrestles.

The retail trade area provides another illustration of statistical coordination. Retail trade statistics—except for department store statistics which had a long history of association with the Board of Governors of the Federal Reserve System were the responsibility of the Bureau of the Census. After several years of effort arrangements have now been worked out so that the full body of retail trade statistics is in charge of the Census Bureau. And in doing this, we have not lost the advantages of local data on retail trade. In fact, we have improved the local data because in many instances the only local data available under the earlier arrangements were for department store

sales. Such sales came to be poorer and poorer indicators of the total retail trade of a particular locality. Now overall measures of retail sales in the major metropolitan areas are being made available.

We still have a long way to go in the area of retail inventories, but we do have a program underway now which will attempt to give us measures of the retail inventories of large consumer durables. If we are successful with that program, we will have at least built one more link in the chain of adequate information at both the manufacturing and the retail level.

Mr. Chairman, major advances have also been made in the social and demographic statistics field. It has been, therefore, one of fastest growing areas of statistics in the Federal Government, and particularly this is true of the health statistics area and the education statistics area where little was done in this area 10, 15 years ago, and now we have extensive surveys in the area of health and the eduactional statistcs program has made major advances and these are described in my paper for the record.

In the population area, also, major advances have been made, not only in the basic censuses but in the current development of estimates for States and for metropolitan areas. The Census Bureau is working on a program now in cooperation with the States where it would be our hope that the States would take responsibility for estimating county populations consistent with Census Bureau estimates of State populations and national population, and that these could be brought together by the Census Bureau as an overall presentation annually of not only the population of the United States and each of its States, but the population of each of the counties of each of the States. This will be an economical program and a program which should have major uses and eliminate major contradictions in the population figures that are used by a wide variety of people. In fact, we have even found States in which three or four agencies were producing population estimates within the same State.

The committee might like to know that there was a difference of opinion for some time between various agencies of the Federal Government and to some extent with the Congress with regard to a middecade census of population. It was not that we did not recognize the need for more population information, but we were rather reluctant to see another census started that would not give sufficient attention to the need for more current data. However, recently the administration has agreed that a middecade census bill that allowed considerable flexibility in the design of such a major effort to measure population and associated characteristics at the middecade but did not represent a mere repetition of the decade census at the middecade point is certainly with the general objectives of the administration. Such a bill is now being worked on by a committee of the Congress and it would be my hope that it might be part of our statistical paraphernalia.

I think, since my paper deals in detail with this area, I need not illustrate further.

This is another area that I would like to say a word or two about which is neglected because it is rather dry and it is hard to get attention for it. That is the importance of standard statistical classifications and other standards for a well coordinated, well integrated body of statistical information. I am sure this committee recognizes that

one thing the office has done and has been doing for many years, is the development of the standard industrial classification.

Without such a standard industrial classification, the comparison of statistics put together by various agencies is virtually impossible. But, what some people do not realize, a standard classification by itself does not necessarily insure the comparability of data because we have to make certain that all of the different agencies of the Federal Government are classifying the establishment in the same industry in accordance with the standard industrial classification. This we have not achieved entirely to our satisfaction to date and so one thing that I am still pushing for, and I will mention it in another connection, is that there be for the Federal Government a directory of establishments interrealted with enterprises, with associated industrial classification codes so that it will be easier for the different agencies of the Federal Government to find out to what extent establishments are classified differently and may, therefore, cause certain differences in different bodies of information for a given industry.

For example, employment versus output. Investments versus employment and output. All problems of where you classify individual establishments or individual firms.

Now, in this classification area we have also recognized that some data have to be collected for enterprises and other data for their establishments. We want, therefore, a classification which will make it easier to cross the bridge between an enterprise classification and an establishment classification, so we are developing an enterprise industrial classification of that sort.

Similarly in the area of foreign trade statistics, exports and imports, we want a classification of commodities which can also be related to a classification of industries and we are working in that area as well and have achieved some success to date and have tried to use that success also in connection with the classification of commodities for internal transportation.

One of the areas in which there is now a great deal of interest is the Federal-State-local relationship, not only in statistics but in many other respects. We are trying not to neglect this in the statistical area as well. And, I think if I may, Mr. Chairman, I would like to read the section in my report on this because it is probably better organized than I can do it orally.

The relationship of Federal-State-local governmental statistical activities has received an increasing amount of attention during the past few years. This new surge of interest was largely aroused by the Governors of the several States themselves. In each year since 1964, the National Governors' conference has by resolution noted the need to develop valid and comparable statistical information in order that programs and program operations in different States may be appraised properly.

Acting in cooperation with the Council of State Governments, the National Governors' conference sponsored a National Conference on Comparative Statistics, which was held on February 23–25, 1966. The Advisory Commission on Intergovernmental Relations, the U.S. Conference of Mayors, the National Association of Counties, the National League of Cities, the Municipal Finance Officers Association, and the Bureau of the Budget served as cosponsors for the conference. I served as a member of the steering committee of that conference.

The conference's agenda was a wide ranging exploration of the information needs for decisionmaking by State and local governments. The conference's principal recommendations were that:

Each State should establish (or designate) a statistical coordination/standardization unit.

There should be a continuing forum for the development of improved statistical data in functional areas where appropriate.

There should be a continuing body to provide the necessary links between the States and the Federal Government to assure the continuing and persistent attention necessary to secure improvements in the comparability of statistical information.

On another front, the National Association of State Budget Officers and the Council of State Governments have been concerned for some time about the interrelationships between increasing requirements for information arising out of Federal grant-in-aid programs, the increasing use of electronic data processing equipment, and the comparability of information designed to serve the needs of program managers and overall executive management. Originally oriented primarily to problems related to data processing, this interest has shifted to the general area of information systems.

We have welcomed both of these efforts. We have added a member of the staff of the Office of Statistical Standards who has the primary responsibility of working on matters relating to Federal-State statistical coordination. This is Mr. Roye Lowry of our staff, formerly associated as executive secretary of the Federal Statistical Users' Conference.

The Bureau of the Budget through OSS has also agreed to provide the secretariat for the continuing body recommended by the National Conference on Comparative Statistics. We hope that another conference on comparative statistics will be held later this year.

An intergovernmental Task Force on Information Systems has been established by the Director of the Bureau of the Budget and by the Council of State Governments to consider the problems and opportunities which arise out of the flow of information between the various levels of government in the federal system. It is chaired by Mr. John Kennedy, Special Assistant on Intergovernmental Relations to the Governor of Illinois. It consists of two representatives from State governments, two from county governments, two from local governments, a representative from the Advisory Council on Intergovernmental Relations, and two members of the staff of the Bureau of the Budget. A member of my staff is one of the Budget Bureau representatives on this task force.

Since the National Conference on Comparative Statistics, some eight States have created or designated a statistical coordination/standardization unit or have such a step under active consideration. There may be similar activity underway in other States about which I am not aware. While it had been hoped that substantially more progress in this direction would have been accomplished by now, the steps that have been taken represent significant gains. I hope that the States which have not yet established statistical coordination/standardization units will soon do so, for efforts to promote the statistical comparability desired by the Governors are seriously hampered when there is no single point of contact within each State charged with overall responsibility for statistical development.

I would like to emphasize, if I may, Mr. Chairman, the fact that it is this overall coordination which is being striven for now. A lot of attention has been devoted in cooperative efforts between Federal and State agencies but there has been no single unit in a State government to which attention could be directed to the overall statistical program of the Federal Government and the State government. What is being proposed is something like a unit such as my office in the States, but I have tried to stress that this does not have to be a big office and it does not have to be a big unit. All it requires is a center for responsibility for statistical programs in the States.

Now, Mr. Chairman, I have reviewed broadly some of the things that we have done, some of the things that still remain to be done, some of our problems in coordination.

I want to turn now, and I will have to be brief, to the major problems as I see them for the future.

I think of the major problem as one of making our body of data really serve analytical uses. This means that statistics can no longer be used as individual items of information. They have to be used as a body of interrelated information.

The reasons why the Bureau of the Budget has been looking into the problem of a Federal statistical data center involve this point. Several issues have to be met.

First of all, what statistical data in detail should be kept? It costs money to keep things around.

Secondly, when you keep it, can anybody use it or do you just keep it because you do not want to throw it away, which is what I do to a considerable extent with many of my pieces of paper and various activities over the years.

We believe that machine technology permits us to devise a better way of storing in accessible form a great deal of essential statistical information which the Federal Government has gathered. That is one of the objectives of the data center idea.

Notice this storage objective really has two parts. Not only better decisions on what to keep but keeping it so that it can be used. Much of our archival materials cost so much to use that you cannot use it even if it is kept, because it costs too much to get at it.

The second advantage for a data center is that in organizing the data files for analytical uses, we will learn certain things which cannot be learned otherwise. This is what we call the feedback principle. We would hope that a data center along these lines would make it possible for us to find out where in various statistical series inconsistencies exist, help us to learn where they are and what they are, and then we can take the necessary actions to eliminate those difficulties for the future.

These are two main objectives that are often neglected.

The third is we believe that users, both within the Government and outside the Government, should have one place where they could come to, to get help in their statistical data needs and could get help in a way which will allow the different bodies of information to be made available to them and not have to contact three or four different agencies in order to get information of this sort.

Now, I am really not—I want to make it clear, Mr. Chairman, I am not talking about isolated pieces of information. I am not talking about a telephone call as to how many people were employed last

month. That still comes from the agency that collects the data. But, I am talking about people that are using historical series over long periods of time and who want to associate employment with output, with investment and with a great many other measures.

So that is another aim of a data center.

Now, this notion of a data center, when it was first presented in some reports which had been made to us by consultants, was seized upon by some people who had a fear that what we were trying to do was to set up in the Federal Government a dossier of information on every individual and firm in the country which could then be used to do all sorts of things about these individual units.

Since the letter from this committee asked me to deal in some detail with this problem of confidentiality, I shall try to do it in a way which will show what we are trying to do in the data center. The statistical programs basic inception recognizes that information collected for statistical purposes should be restricted to statistical purposes. It should not release any information concerning any individual. But, there was a legitimate need for our attention to be called to the possibility that a data center might so centralize information that it could be a threat to the privacy or the confidentiality of information about individuals.

We have, therefore, said to two committees of Congress that we will not propose a specific data center unless we can satisfy ourselves that it does not infringe upon the confidentiality or privacy of individuals, and with your permission I would like to sketch in some of the things we are thinking about, although we have no definite proposal to date.

I have tried to indicate, first of all, why we think the data center is necessary to carry on the job. The way we are thinking about it now is that a data center could be organized within one of the existing agencies of the Federal Government, that such a center should clearly not have in it information whose main use is with respect to individuals. So, we would clearly not think of putting into the data center personnel files, files which are basically associated with individuals.

Secondly, we think that no universe data need be put into the center as an actual storage of such data. This means, therefore, that if anybody tried to use the center for obtaining information about individuals, they would have a chance, say, one in a hundred, of finding that individual's record or that firm's record in the center.

Third, we would recognize that the Congress by law would have to make it positively clear, as it is now in the Census Act, that the center would never release any information to anybody about an individual or an individual business firm. We believe the restriction on the universe data and other restrictions that can be built in the center would also take care of surreptitious uses of the center contrary to the law. And we would pay particular attention to that.

We believe that another device that might prove useful would be an arrangement whereby such a center would have associated with it an advisory committee under the chairmanship of the Bureau of the Budget and made up of the principal statistical officers of the Federal Government that would pass upon all data to be put into the centers as well as the general procedures for using data from the center.

These are the elements that we have thought of as usable for protecting the confidentiality and privacy of individual records. We are working on this problem. I will be glad to discuss it further with the

committee if you have questions but, in general outline, that is the kind of a center that we are talking about.

I will say two other things about it. One is, a center must meet the needs for two kinds of analysis. One is macro, and that provides very little difficulty with regard to confidentiality. But the other and also very important is microanalysis where people are trying to discover what are the factors which account for the various conditions such as why are some establishments expanding and others are not. Why is it that some people are poor and others are not? Why is it that some kinds of poverty are persistent over time while others evaporate quickly with age and education? This type of microanalysis where the data base has to provide for the association of data about individual units, both longitudinally and latitudinally, this is where the great difficulty is. We think this need can also be met without releasing information about any individual, but it would be wrong to consider a data center that did not provide for this kind of analytical use as well as the macro type of analytical use.

Mr. Chairman, I think I have taken more time than I had really anticipated. It might be better, then, if I stopped at this point so that you could ask the kind of questions which are pertinent to your interests. I will say before closing that for the information of the subcommittee, I have attached to the report for the record a table showing the detailed expenditures of the Federal statistical program for the year 1950 and for each fiscal year from 1956 to 1968 by program and by agency because I thought this committee might like to have that kind of a record.

Thank you.

Representative BOLLING. Without objection, that will also be included. Mr. Bowman, we are grateful to you for your comprehensive and thoughtful statement. We will include your prepared statement at this point in the record.

(The statement referred to follows:)

PREPARED STATEMENT OF RAYMOND T. BOWMAN

Mr. Chairman and Members of the Committee, I am glad to have the opportunity to again appear before this Subcommittee to discuss the Federal Statistical Program: its goals, its recent achievements and the outlook for the future. The Joint Economic Committee through this Subcommittee has played an important role in securing the substantial improvements of Federal statistics over the past two decades. Your hearings have provided a forum at which the relation of statistics to policy making has been made explicit—a meeting place where the producers and users of statistics can search together for better ways to bring more and better quantitative information to bear on crucial policy issues.

The subject of the current hearings are particularly pertinent because they focus attention on areas of central importance to the further improvement of Federal statistics—the coordination, integration and appropriate availability of Government statistics.

In 1956, shortly after shifting from academic pursuits to my present responsibility for statistical coordination, I attempted to set forth in an address before the American Statistical Association some ideas for the improvement of Federal statistics. I entitled that address: "Philosophy of an Inegrated Federal Statistical Program." My emphasis in that paper was on the need for developing and improving appropriately interrelated, accurate, timely, consistent and relevant statistical information for a better delineation of the economic and social functioning of the Nation and its parts.

I attempted to focus attention on three requirements:

The need to better understand, in quantitative terms, how the economy operates;

The need to be able to diagnose corrective actions where appropriate;

And the ability to evaluate the success of corrective actions taken.

We have come a long way toward meeting these requirements in the ten years since that paper was written. We have a better understanding, in quantitative terms, of the working of the economy in its various sectors than we had 10 years ago; we have better data to guide us in both public and private actions; and we have learned to use these data more effectively in analyzing and solving our problems. The economic success which we have enjoyed as a Nation is at lease partly due to the increased availability and greater use of better statistical information in decisionmaking. Our success on the one hand permits us, and on the other requires us, to deal with new problems which were of less concern to use ten years ago. The problems to which we are turning more attention today call for new kinds of information, greater attention to area details and new ways of organizing it. Despite our progress, we still have much to do, to achieve the goal of appropriately interrelated, accurate, timely, consistent, and relevant statistical information for today's problems.

I would like on this occasion to review for the Committee's attention what has been accomplished in the recent past, what is in process currently and the general methods which we are using to reach our goals. In this accounting for the work of my Office and the statistical agencies generally, I shall note the shortcomings as well as the successes. I shall recognize the need for performing the statistical task of the Government economically by considering not only costs to the Government, but the burden which the collection of information places on respondent businesses, individuals, and nonprofit institutions. I will emphasize the need to maintain the privacy of individual reports while at the same time underscoring the importance of maximum accessibility of the data to all classes of users for statistical purposes.

I want also to note that even the most precise statistics cannot determine substantive goals or choose means to achieve them. Good statistics cannot eliminate differences of opinion about objectives. Good statistics can, however, provide a framework of knowledge which increases the probability that choices reflect value judgments and not ignorance.

National Economic Accounts

The national economic accounts are the most systematic and comprehensive statistical measure we have of the Nation's economic activity. These accounts include the national income and product accounts, interindustry purchases and sales, flow of funds accounts and the balance of payments. An expansion of the accounts to include national and sector balance sheets is planned. By their very nature, the national economic accounts call for bringing together statistical information from a vast number of separate statistical series which are collected by a large number of agencies for a variety of purposes. Data from the Bureau of the Census, from the Department of Agriculture, from the Department of Labor, the Internal Revenue Service, the Bureau of Customs, the Immigration and Naturalization Service, the Securities and Exchange Commission and the Federal Trade Commission, among others, are integrated into the national economic accounts.

Just as the statistics which provide data for the national economic accounts developed at different times, in different places, in response to different needs, so the various acounts themselves grew up originally without too much regard for their integration into a common framework for economic analysis.

When we first directed attention to the problem of working toward the goals outlined in my remarks to the American Statistical Association, it was natural that we should turn first to the national economic accounts. They are our most important analytical tool for appraising the Nation's economic position and its course of development. Such an accounting structure provides guidance for determining what different elements of statistical data are necessary and the consistent definitions required for their integration.

Our first step was to request and contract with the National Bureau of Economic Research to form a committee to review and appraise the accounts and to make recommendations for their improvement both by integrating and extending the accounts themselves and by strengthening the statistical information which is incorporated into the accounting structure. The National Economic Accounts Review Committee, as it came to be known, made a number of important recommendations, many of which have been translated into reality. The Committee's report and recommendations were the subject of hearings before this committee in the fall of 1957. Without attempting an exhaustive review of

the 29 recommendations by the Review Committee, I would like to note some of the major recommendations which have been implemented:

Interindustry purchases and sales studies have been established as a regular part of the national economic accounts. The construction of a 1958 interindustry purchase and sales table, integrated with the income and product accounts, permits consistent analysis of intermediate and final product outputs. With this table, published in November 1964, official research on interindustry relations was reestablished. It will be recalled that the United States had taken early leadership in this area and then lost it. An interindustry purchase and sales table is now in preparation for 1963 under a program for such a Table every five years.

The flow of funds accounts have been more closely integrated with the national income and product accounts and have provided for quarterly presentation of these accounts. But more is required to strengthen the integration for general analysis.

Quarterly national income and product data have been published in constant dollars—a change which considerably facilitates current analysis of total output, investment, and consumption.

Annual estimates of real product by industry have been introduced, a change which makes possible the detailed consistent analysis of productivity and price changes within a GNP framework.

The foreign transactions account has received particular attention because of the Nation's persistent balance of payments problems. In 1963, the Director of the Bureau of the Budget appointed a committee, under the chairmanship of Dr. Edward M. Bernstein, to make an intensive review of balance of payments statistics. The Committee's report was the subject of hearings before this Committee in May and June 1965. As a consquence of this review, a major reorganization of the presentation of the balance of payments has been achieved. This considerably improved analysis of foreign transactions and provided two official presentations of a payments balance: the liquidity balance, and the balance based on official transactions.

I would like to note two other significant developments which embody the use of the national economic accounts in the analysis of emerging economic problems. A long-term growth model for projection to 1970 of labor force, productivity, GNP, industry output and employment within a framework consistent with the income and product and input-output accounts has been constructed. Results were published early this year. This work was done in the Departments of Labor and Commerce, under the direction of an interagency committee represented by these Departments as well as the Council of Economic Advisers and the Budget Bureau.

A short-run quarterly model of the economy by OBE utilizing 13 equations based mainly on accounting identities and 36 equations based on historical behavior covering all key economic processes and activities has been constructed. A description of the model, its past behavior and the listing of the basic equations were published in the May 1966 *Survey of Current Business*.

While this Committee, the Office of Statistical Standards, the statistical agencies, and users of Federal Statistics, may look back with some feeling of accomplishment, there is no room for complacency. The task has just begun.

The continuing work of extension, integration, and data improvement involves efforts of many agencies and requires much greater attention to the accuracy of the several interrelated parts. The need to resolve inconsistencies and to promote the fullest use of the fruit of separate efforts requires bringing the several producers and the interest of different users of data into more productive relationships.

The following coordinating operations, now going forward, exemplify such efforts:

1. The Committee on Measurement of Real Output—The Bureau of the Budget has recently organized a committee which it chairs to include FRB, OBE, Census and BLS. Its immediate task is to resolve differences in measurement of economic growth and fluctuations provided by FRB Index of Industrial Production and OBE measures of real output by industry. Its longer run goal is to integrate the main bodies of information on inputs, outputs and prices in order that study of productivity, prices, wages, and final demand can be done in detailed and consistent fashion and the data base for such studies more efficiently organized.

2. Assignment of focal responsibility to the Department of Commerce to develop a comprehensive statement on the Nation's tangible wealth. Frameworks

for this statement are being developed by OBE on the basis of which improvements in Census surveys are to be made to obtain basic data. The objective is to achieve time series which should have a benchmark at or near the end of each decade. Improved time series on private fixed capital formation have been published in the *Survey of Current Business* for *December 1966* and *February 1967*.

3. The Technical Advisory Committee on Balance of Payments Statistics, chaired by Bureau of the Budget to implement recommendations of the Bernstein Committee, includes representatives of Treasury, OBE, FRB, and CEA. This Committee arranged for the introduction of the official quarterly presentations in the *Survey of Current Business* of the two balances—liquidity and official reserve transactions. Under its aegis the basic tables have been reformulated to improve analysis of receipts and payments and the committee continues in an effort to resolve difficulties in this complex area of statistics as they arise; for example, in presenting unusual financial or other transactions which may have special effects on balance of payments position. A more detailed statement on actions to date on the Bernstein Committee report has been prepared and published in the *Staistical Reporter*. Copies of this statement have been sent to all members of this Committee.

Prices

The need to overcome deficiencies in price statistics has been a matter of continuing concern because of price developments in the postwar period. These include:

The marked inflationary pressures associated with the aftermath of World War II and the onset of the Korean conflict.

The puzzle of increasing prices under conditions of less than full employment, particularly during 1957–58.

More recently, the pressures on prices which have accompanied the move toward full employment and the rising outlays in Vietnam.

The need, partly because of our balance of payments difficulties, to improve understanding of our competitive position in world markets with better data on export-import prices.

In 1959, the Bureau of the Budget contracted with the National Bureau of Economic Research to form a committee to review the Government's Price Statistics. A committee of distinguished economists, under Professor George J. Stigler as chairman, made a thorough examination of price data, and made its report and recommendations in the fall of 1960. They were discussed before this Committee in January and May 1961.

Since the appearance of the Review Committee's report a number of significant improvements have been made. I would like to cite the following:

1. The Consumer Price Index (CPI) has been updated with new weights and broadened to include new series covering single as well as family workers. The CPI has been placed on a probability basis incorporating a replication system permitting measurement of sampling error. The CPI indexes were extended to all SMSA's with over one million population.

2. The commodity coverage of the Wholesale Price Index (WPI) has been widened.

3. Work has begun to organize the WPI data along industry sector lines. Such indexes, developed according to the Standard Industrial Classification (SIC) better serve as deflators for gross product originating in manufacturing and other sectors. Much further work needs to be done on this, however, to permit consistent analysis of price and output changes.

4. On export-import prices we are working closely with the Departments of Labor, Treasury and Commerce to

(1) Improve unit value data, and

(2) Develop new approaches to improved pricing of exports for U.S. and other countries including possible cooperation with the U.N. on an international reporting system.

While a number of actions to improve the body of price statistics in the United States have taken place, many of the recommendations of the Stigler Committee have not been fulfilled largely because it has not been possible to obtain the resources for the needed work. There is presently pending before the Congress a request for needed improvements in price work which include:

a. Pilot and testing work in preparation for a consumer expenditure survey to be taken around 1970 to update the weights for the Consumer Price Index (CPI);

b. Research on improved techniques to take into account changes in the value of goods purchased by consumers which result from changes in quality as distinct from "pure" price changes;

c. The more rapid development of price indexes by industry in order that it may be possible to determine which industry or industries have contributed to the change in prices of goods purchased by final consumers. Such indexes, called "sector price indexes," are to be arranged according to the Standard Industrial Classification in order that consistent analysis or prices, wages, production, and productivity may be made—and a framework may be provided for analysis of demand and cost influences upon price changes.

Labor Statistics

More than ten years ago—in November 1955—I appeared before this Subcommittee to report on the status of employment and unemployment statistics. A number of recommendations for improvement were made which recognized the importance of these data in measuring the input of labor as well as the utilization of human resources. Most were shortly put into effect—revised definitions of unemployment in 1957, an increase in the coverage and detail of our industry and area employment statistics, information on overtime hours and on reasons for part-time work. But implementation of the recommendations for increased analytical studies of labor force participation, for information on the types and sources of unemployment was deferred for lack of funds.

In the summer of 1959, a major realignment of function occurred. Responsibility for planning, analysis and publication of the employment and unemployment statistics obtained from the household survey was transferred to the BLS, while the Census Bureau continued to collect the data as agent. In effect, BLS became more completely the responsible agency for employment and unemployment statistics. Meanwhile, the wealth of detail expanded, and seasonal adjustment of the series was introduced to enhance its use as an economic indicator. Nevertheless, the continued high level of unemployment fostered continued doubt over the reliability and adequacy of the unemployment series.

The President appointed an outside group, the Committee to Appraise Employment and Unemployment Statistics (the Gordon Committee) to review the whole field and recommend improvements. The Committee suggested many changes or directions for future work. Of them all, I should like to touch on two. The first is the major development and testing program for improving the questions in the monthly household survey. Revised questions on unemployment and nonparticipation in the labor force were devised, tested, tentatively adopted and used during a year's overlap period before being incorporated in the official series. At the beginning of this year the basic sample for the monthly labor force reports was expanded from 35,000 households to 52,000. This permits more accurate measures of the data by age, sex, color and other characteristics of labor force participants. The Office of Statistical Standards played a major role in encouraging this work on the CPS in response to the Gordon Committee recommendations and in focussing interagency technical and policy advice on the adoption of the specific proposal which finally evolved.

Efforts to carry out another Gordon Committee recommendation have been less successful. I refer to the proposal to develop job vacancy statistics. The Labor Department, encouraged by some exploratory private attempts, has developed plans for such surveys and has engaged in large-scale testing for several years. As Labor Department representatives reported to this Subcommittee last summer, results have been promising. However, despite the considerable interest in such data, as a means to assist in employment service operations, to throw light on the operations of the labor market and to measure the current conditions of demand in the labor market, funds have not been provided by the Congress to make such surveys "operational." Meanwhile, the Budget Bureau is continuing to work with the Labor Department on improvements in these pilot surveys.

Before leaving the Gordon Committee Report, I might remind you that its Report, like those of other review committees, was far-ranging and not limited solely to immediate changes. Its influence will continue to be felt. For example, BLS is in the process of developing current information on occupational employment from employers, the Department of Agriculture is in the midst of a major revision of its farm employment series and the Census Bureau is continuing some of its methodological tests in connection with the Current Population Survey.

Major changes in other types of labor statistics have occurred in the past decade. Again, I will be selective and mention only two. The first is the strength-

ening of information on the training and employment of scientific manpower. Information on scientists and engineers is collected in a variety of ways, often as part of a data collection system which cover much broader groups—for example, the population censuses or reports on college enrollments. In order to obtain help in our coordinating functions and in recognition of its unique and specialized interests, we asked the National Science Foundation to undertake a focal agency assignment in this area.

The other major development is the strengthening of occupational wage rate information. The BLS has been conducting a special survey of salaries in selected professional, administrative, technical and clerical occupations in private industry to provide wage data for comparison with Federal employee compensation. The Bureau of the Budget has cooperated with BLS and CSC in planning these surveys, and about three years ago it employed a consultant firm to evaluate the adequacy of the survey. This evaluation resulted in an expansion of the survey to include non-metropolitan areas and smaller establishments, as well as other improvements.

Construction, Production, Distribution, and Transportation

Our statistical information relating to construction, production, distribution, and transportation is a continuing description of the Nation's economic activities. It should provide consistent interrelated information.

Estimates of the value of construction activity are incorporated directly into the product side of the national economic accounts. This component of economic activity accounts for over ten percent of GNP. Reliability of estimates for various types of construction—residential, industrial, and commercial—and by ownership—public and private—is of paramount importance in analyzing the state of the economy and in the study of economic growth and additions to our capital stock. Our statistical measures of construction for a variety of reasons have long been among the weaker of our statistics. The inherent difficulty of collecting meaningful data on construction activity was heightened before 1959 by a division of responsibility for collecting the data among three agencies: Business and Defense Services Administration, Bureau of Labor Statistics, and the Bureau of the Census. At that time we recognized, promoted and accomplished the centralization of responsibility for construction statistics in the Bureau of the Census as a necessary prerequisite to a consistent, long-range program to improve these data.

Because construction activity itself is very decentralized, there are numerous difficult conceptual problems to be faced in developing a meaningful set of measures. In addition the collection of data is expensive and burdensome on respondents. Progress in improvement has been disappointly slow. Nevertheless, we have made some noteworthy progress. The series on building permits is far more inclusive than it was in 1959. The measure of housing starts is a more accurate reflection of reality than it was, because of the instituting of actual reporting for such starts. The monthly measure of private, nonresidential construction *put-in-place* is even more difficult. Repairs and rehabilitation of residential structures also presented special difficulties but despite some setbacks, progress in measurement of this element has been improved. Nevertheless, construction activity statistics leaves much to be desired and our measures of current construction activity are still unsatisfactory in promptness sand accuracy. Furthermore, we have made almost no progress in developing measures of construction prices and the failure to secure appropriations for a mid-decade Housing Inventory cost the Nation an important mid-decade benchmark of changes taking place in the stock of housing for its citizens.

The program for further improvements in the construction activity series in the near future includes updating the patterns which are used to translate value of work started on one- to four-family residential structures into work put-in-place, and the initiation of progress reporting for larger residential structures. Also funds are requested in the 1968 budget to expand reporting from building permit jurisdictions so that monthly construction figures can be published for 100 standard metropolitan statistical areas.

When the Bureau effected the transfer of the monthly *Survey of Manufacturers' Shipments, Inventories and Orders* from OBE to Census in 1958, Census was charged with responsibility for strengthening the series and for achieving consistency with other measures for individual manufacturing industries. The sample of reporters has been expanded and many of the large companies report on a "divisional" basis thus permitting better and more comparable measures by industry.

Benchmarking of the *Inventories* has been shifted from corporate tax returns to the *Annual Survey of Manufactures*. As a result of these changes, consistency with the Annual Survey and the quinquennial Census of Manufacturers has been enhanced, and the amount of detail by industry groupings and stage of fabrication has been enlarged.

The link of Census *establishment* data from the 1963 consensus and IRS *corporation* data will be substantially expanded over that for 1958. The number of corporations matched will be nearly doubled and many more data items will be linked. Most of the procedures for the matching of IRS cases and their equivalent Census company records have been computerized, whereas the 1958 project was largely a hand operation. The linked record files thus created constitute an important data source for analytical uses and a base for providing company-establishment matrices required in preparing input-output tables, national income and GNP estimates.

I would like also to call your attention to the fact that the use of administrative records has substantially reduced the reporting burden for businesses covered by the Economic Concensus. Some 900,000 small businesses were freed from the burden of reporting for the Census in 1963 because the Census Bureau was able to get the data it needed from IRS records. Another 1,000,000 small businesses will be relieved from the chore of filling in questionnaires for the 1967 Censuses for the same reason.

Consolidation of responsibility for the development of the program of retail trade statistics in the Bureau of the Census recognized the necessity to bring department store sales data which had been compiled by Federal Reserve into better alignment with other retail trade data and the need for metropolitan area trade data broader than just department store sales. Efforts of the Bureau of the Budget extending over a number of years to accomplish this are now bearing fruit. Department store sales data are now completely integrated into the total retail sales series. Furthermore, an annual series of merchandise line data in retail trade has been initiated with collection of data for 1966. The lines covered are similar to those obtained in the 1963 census of retail trade, but the sample coverage will not permit publication of the full breakdown for certain kinds of stores. The 1968 budget for Census requests funds for an expanded data collection program to enable the Bureau to provide uniform breakdowns of retail sales in the limited number of SMSA's for which such data are published.

For several years, the Bureau of the Census has been engaging in research and development work in the collection of inventory data at the retail level. The 1968 budget requests funds to provide monthly national indexes of the dollar and physical volume of total retail inventories of large consumer durable items.

Publication of a monthly series on service trade receipts has been initiated. This filled an important gap in our economic intelligence for a rapidly expanding sector of our economy. Further development of the series will entail more comprehensive coverage and improved classification and breakout of service trades by type. This is an area which is becoming more important but is most difficult to cover statistically. The limited data we now have became available only after some five years of intensive efforts on the part of the Bureau of the Census.

A little over two years ago, an interagency Petroleum Statistics Study Group organized and chaired by the Bureau of the Budget submitted its report to me. The Study Group was concerned with the domestic petroleum industry—reserves, productive capacity, wells, deliverability, expenditures and revenues. Its report laid out the general framework for a coordinated program to provide needed informaiton pertaining to these areas and pointed up the necessity for further study required to develop the details for the program before it could be implemented. In March of last year, the Bureau assigned to the Department of the Interior "focal agency responsibility" for carrying out eight recommendations of the Study Group. That Department has recently prepared its first report which we believe indicates statisfactory progress in the various initial tasks.

Social and Demographic Statistics.

This is the fastest growing area of statistics and information-gathering. Expansion has taken place by specific developments in fields of special interest, such as the nationwide health surveys, the large-scale reorganization of educational statistics, and information on scientists and engineers and research and development expenditures, which is still in progress. Expansion has also occurred as the development of new action programs required new administrative reporting procedures which produce statistical by-products and which require considerable coordination under great difficulties. Another type of expansion is the growth in the number of contract research projects (requiring review under the Federal

Reports Act) which are used to plan or evaluate on-going Government programs. Not only is coordination important in these last two areas but we have little practical experience to guide us.

During the past ten years a whole new area of health statistics has been planned, tested and developed, including health interview surveys of the population, health examinations, and summaries of seleotive health records. These new statistical series, for which many concepts and methods had to be developed, have been placed in a new oragnizational unit, the National Center for Health Statistics in the Public Health Service, together with the responsibility for vital records and reports. Much of the contribution of this Office to this major development came in the form of review of specific proposals for data collection and tabulation, although at one time we convened a small group of experts to evaluate the scope and effectiveness of the new surveys.

The success of the NCHS has led to the planning of other statistical centers, notably the Center for Educational Statistics, which is working on improving and coordinating educational statistics with the guidance of an "educational model" to indicate gaps and inconsistencies. This work has not progressed far enough to make much impact visible to users of educational data, but I anticipate the results will become evident in the near future. We had urged the Office of Education to undertake such basic review and improvements during most of the early part of this decade, so that better information will be available for policy guidance.

Statistics on crime and delinquency remains a difficult area, in which I cannot report much progress at this date. We are working closely with the National Crime Commission, particularly its Task Force on Crime Assessment, on developing better statistics and establishing a national crime statistics center.

If you have been following the basic population statistics produced by the Bureau of the Census, you are already aware that this area of statistics has been strengthened. More current estimates are made, on more characteristics, for smaller areas, and the program of population projections has expanded. At the present time, we are encouraging the Census Bureau to establish a cooperative program of county population estimates with the various States. Census would provide technical guidance and research and publish the results so that comparable data would be widely available, but the States would produce the actually county estimates in between Census dates, in the light of local knowledge and conditions but consistent with National and State estimates by the Census Bureau. If this program succeeds, it will be a notable step forward in our policy of developing more Federal-State statistical relationships.

Of course, the kind and quality of population and housing statistics depend to a large degree on the scope and adequacy of the decennial censuses. In preparation for the 1960 Census, the OSS undertook the specific function of coordinating Federal agency needs for Census data. The questions to be asked, and the tabulations to be made, were considered by this group and a variety of useful recommendations made and adopted. We are undertaking the same function in connection with planning for the 1970 Census.

Population statistics, including the social and economic characteristics of the population, are important as a general-purpose statistical resource. They are also needed as indicators of problems, as measures of progress, for many of the new anti-poverty, manpower development and educational assistance action programs. To coordinate these interests and start the development of longitudinal studies for Government-wide use, the OSS developed a plan for a large-scale population survey in 1968, in cooperation most particularly with the Census Bureau and the OEO, but with the advice and assistance of all interested agencies, working through the Federal Council on the 1970 Censuses. A plan for a survey of about 2.6 million households has been developed and is now before the Congress for funding.

Our future work in the demographic and social statistics area may lay more emphasis on the development of standard classifications (on occupations, for example) and on an overall framework of "social indicators" against which statistical progress can be measured and weaknesses pointed out.

Standard Statistical Classifications and Other Standards

Standard statistical classifications are a necessary means for achieving a body of coordinated and integrated statistical data. Among the most important of the classifications developed and maintained under the auspices of the Bureau of the Budget are the Standard Industrial Classification (SIC); Standard Metropolitan Statistical Areas (SMSA); Standard Enterprise Classification; Commodity Classification for Transportation Statistics; and Statistical Classifications of

Commodities Imported into (Schedule A) and Exported from the United States (Schedule B). At the present time the Budget Bureau is undertaking the preliminary review to determine whether it should attempt to develop a standard occupational classification for statistical purposes. A revision of the Industrial Classification manual was just released this year as was a revision of Standard Metropolitan Statistical Areas.

While standard classifications are essential to comparable data, more efforts must also be devoted to a closer examination of the consistency in fact of actual classifications. For the industrial classification, in particular, there is clear need for an official Federal directory of establishments and firms which can be used by all the statistical agencies of Government so that checks can be made to insure uniform classification in accordance with the manual.

The Bureau of the Budget also issues other standards with respect to methods of compiling statistical reports, establishing a common reference base for index numbers, establishing uniform reporting periods for employment and payrolls, and certain definitions. It is my impression that we have not done enough in this area for present needs.

Federal-State-local relationships

The relationship of Federal, State, and local governmental statistical activities has received an increasing amount of attention during the past few years. This new surge of interest was largely aroused by the governors of the several States themselves. In each year since 1964, the National Governors' Conference has by resolution noted the need to develop valid and comparable statistical information in order that programs and program operations in different States may be appraised properly.

Acting in cooperation with the Council of State Governments, the National Governors' Conference sponsored a National Conference on Comparative Statistics, which was held on February 23–25, 1966. The Advisory Commission on Intergovernmental Relations, the U.S. Conference of Mayors, the National Association of Counties, the National League of Cities, the Municipal Finance Officers Association, and the Bureau of the Budget served as co-sponsors for the Conference. I served as a member of the steering committee of that Conference.

The Conference's agenda was a wide-ranging exploration of the information needs for decision-making by State and local governments. The Conference's principal recommendations were that:

> Each State should establish (or designate) a statistical coordination/standardization unit.
>
> There should be a continuing forum for the development of improved statistical data in functional areas where appropriate.
>
> There should be a continuing body to provide the necessary links between the States and the Federal Government to assure the continuing and persistent attention necessary to secure improvements in the comparability of statistical information.

On another front, the National Association of State Budget Officers and the Council of State Governments have been concerned for some time about the interrelationships between increasing requirements for information arising out of Federal grant-in-aid programs, the increasing use of electronic data processing equipment, and the comparability of information designed to serve the needs of program managers and over-all executive management. Originally oriented primarily to problems related to data processing, this interest has shifted to the general area of information systems.

We have welcomed both of these efforts. We have added a member to the staff of the Office of Statistical Standards who has the primary responsibility of working on matters relating to Federal-State statistical coordination. The Bureau of the Budget through OSS has also agreed to provide the secretariat for the continuing body recommended by the National Conference on Comparative Statistics.

An Intergovernmental Task Force on Information Systems has been established by Director of the Bureau of the Budget and by the Council of State Governments to consider the problems and opportunities which arise out of the flow of information between the various levels of Government in the Federal system. It is chaired by Mr. John Kennedy, Special Assistant on Intergovernmental Relations to the Governor of Illinois. It consists of two representatives from State governments, two from county governments, two from local governments, a representative from the Advisory Council on Intergovernmental Rela-

tions, and two members of the staff of the Bureau of the Budget. A member of my staff is one of the Budget Bureau representatives on this task force.

Since the National Conference on Comparative Statistics, some eight States have created or designated a statistical coordination/standardization unit or have such a step under active consideration. There may be similar activity underway in other States about which I am not aware. While it had been hoped that substantially more progress in this direction would have been accomplished by now, the steps that have been taken represent significant gains. I hope that the States which have not yet established statistical coordination/standardization units will soon do so, for efforts to promote the statistical comparability desired by the governors are seriously hampered when there is no single point of contact within each State charged with overall responsibility for statistical development.

I have mentioned some of the major efforts which have been undertaken in the recent past to bring about an integrated body of statistical information which can be used to better understand existing and emerging problems and to provide a basis for developing action programs to deal with them.

It may be useful at this point to first note the more specific formal authority under which the Bureau of the Budget currently carries out the statistical coordinating function and second, to list the major methods that it uses.

The Budget and Accounting Procedures Act of 1950 provides in Title I, Part I, section 103 as follows:

"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 for any purpose by the various agencies in the executive branch of the government. Such regulations and orders shall be adhered to by such agencies."

This provision of law is carried out under Executive Order 10253.

Another specific grant of authority is provided by the Federal Reports Act of 1942. It empowers the Director of the Bureau of the Budget (a) after certain procedures specified in the Act to transfer the responsibility 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. This review of reporting requirements is broader than statistical inquiries and relates to all requests for information by the use of indentical questions addressed to 10 or more respondents.

As I have indicated illustratively in my remarks these duties are carried out flexibly and with reference to the particular problems at hand. The methods used include:

The direct transfer of responsibility for the collection, preparation, and maintenance of statistical series from one agency to another.

The focal agency principle—the assignment of responsibility for exercising leadership for the planned and coordinated development of a subject matter field to a particular agency, without any transfer of function relating to the actual collection and preparation of statistical series.

Basic guidance of the statistical program through the President's budget requests for statistics programs. This is a particularly effective way to determine priorities for an integrated program. A special budget analysis, setting forth the details of statistical programs in a single place and explaining the interrelationships of recommended improvements has strengthened the effectiveness of this action. We are grateful to this Committee for suggesting it.

Authority to review data requests directed to the public is also extremely important. It assures that each request for information follows sound statistical procedures, and prevents duplicating of requests.

The use of interagency task forces and consultants and consultant groups from outside Government is essential for thorough examination of issues to determine sound and well considered courses of action for the improvement of statistics that can be supported by both the executive and legislative branches. Here again we owe a large debt of gratitude to this committee for promoting the use of outside consultative groups by the example of the groups set up by the Federal Reserve Board at this committee's request more than a decade ago to review statistics on Inventories, Plant and Equipment Expenditures, Consumer Anticipations and Savings.

The methods which we have employed are related to our desire to maintain only a small but experienced staff of well trained persons. It has not increased in size over the last decade and presently numbers 37 persons of whom 25 are professional staff. Our ability to carry on the job is due to the dedication and experience of this staff.

In my concluding remarks I would like to stress the needs of the future.

The proliferation of statistical materials, the rapid development of computer technology, and the increasingly wide spread and more detailed use of quantitative information have created certain very important needs for the further improvement of compatible statistics for interrelated uses.

Among the outstanding needs are the following:

We need a better way to determine what details of basic statistics should be stored so that they can be retrieved efficiently.

We need a better way of serving the needs of governmental and nongovernmental users of statistics for comprehensive analysis which require the use of historic data files and/or interrelated files providing cross section data for analysis of a micro character.

We need an operation that can examine the current and prospective data files to make certain that they are compatible and where they have deficiencies suggest steps that might be taken to assure compatibility of future files.

We need a service which can make it easier *to check* whether the industrial classification of establishments or firms by different agencies is uniform and in line with the standard classifications prescribed.

It is for these reasons that we have currently been examining the feasibility and propriety of establishing a Federal (or National) Statistical Data Center. We have been conscious from the outset of our investigations into these matters that a Statistical Data Center can be a feasible way to meet these needs only if we can maintain the long accepted practice of statistical agencies of protecting the confidentiality of information about individuals or businesses units collected for statistical purposes. Statistical agencies and the Bureau of the Budget have long been convinced that the U.S. statistical program enjoys a major advantage over the statistical programs of many other countries because of the greater willingness of businesses and individuals to cooperate in providing information.

We are also convinced that the ready cooperation in providing basic data for statistical purposes requires that it be used for statistical purposes only and that so far as the individual respondent is concerned, it will not be used against him or be revealed with reference to him in any way.

Recent reactions to publicity about a Data Center and Congressional discussion of the general problem of privacy of individuals have strengthened our conviction that a statistical data center if it is to be proposed must protect the confidentiality of information provided for statistical purposes and must not be a vehicle for invasion of individual privacy. It is for these reasons that we have indicated to two committees of Congress that any proposal for a statistical data center will be presented for Congressional approval and then only after we are convinced that it will not violate or provide a significant basis for violating individual privacy.

I believe that many of our requirements now and in the future for a body of compatible and accessible statistics for interrelated analysis on both a micro and macro level could be materially advanced by a Federal Statistical Data Center. I also agree such a Center must not sacrifice confidentiality of reports for statistical purposes or the privacy of individuals. It is for this reason that I want to close my remarks with a brief outline of our current thinking of how such a Center might be designed.

Let me make it clear at the outset that no decision on a specific proposal has been reached. All I can do here is share our thinking with you.

We believe that if a Federal Statistical Data Center is to be established, it must be clear that such a Center will not provide any information about individual persons or businesses. This prohibition must be established by law. Moreover, the organization of data files and the concentration of files in such a Center must not make it feasible to abuse the system.

We believe this means that certain types of records should not be in the Center at all. Examples of records we think should be excluded are (1) individual personnel records (letters of reference, performance ratings, test scores, etc.) of Federal employees and applicants; (2) military personnel records; (3) files compiled by FBI, regulatory or other agencies as a result of investigations of individual persons, or businesses or other organizations; (4) FBI fingerprint

files and files on persons convicted of crimes; (5) files of revoked drivers' permits; (6) medical records on Government employees or applicants, and patients of Government institutions; etc.

It seems obvious to us that since this type of information is largely useful to provide information about individuals *per se* it has no major place in a statistical data center which, by definition, does not provide data about individuals.

We believe it will also be necessary to take another step in restricting the files of a Federal Statistical Data Center. If the Center is given authority to utilize the files of the data originating agency, when necessary, then the files the Center actually possesses can be restricted to samples that no universe data would be in the continuing possession of the Center. This will largely foreclose temptations to organize files along individual dossier lines and would make attempts to use the Center to obtain information on individuals have very little pay off. If the Center is to achieve its purposes, however, it must have authority to use such universe data outside its own files under appropriate conditions and also to have some impact on the type of file maintenance of such data.

Perhaps equally important with these general principles is the need for providing some way of authorizing what goes into the Center. It might be provided that transfers of data into the Center could only be by direction of the Director of the Bureau of the Budget in consultation with appropriate advisory groups. This could be arranged by supplementing the authority now vested in the Director under existing law.

We believe that these types of arrangements supplemented by an annual report of the Center's file inventories and uses could provide the protection to privacy which is desired. These arrangements would not significantly impair the usefulness of a Center in making data and data services available, and in improving the compatibility of data when used in association.

It must be remembered, of course, that the development of a successful center, if proposed, would be difficult and could not be immediate. I estimate that the first two years would be required to determine what files are pertinent for center use; to edit them for storage and retrieval uses and to arrange for their transfer or for access to them when transfer seems undesirable.

I have attempted to cover most of the major topics in which I believed this Committee is interested. For the general information of the Committee, I am attaching two tables showing expenditures for the principal statistical programs and agencies for 1950, and by years since 1956. I will be most happy to try to answer any questions you may have.

TABLE 1.—*Obligations for principal statistical programs by broad subject areas*

[In millions of dollars]

Program	Fiscal years													
	1968 estimated	1967 estimated	1966	1965	1964	1963	1962	1961	1960	1959	1958	1957	1956	1950
Grand total	162.7	124.9	107.5	110.0	93.5	82.4	68.0	72.6	141.2	56.7	46.9	40.8	43.3	85.6
Total current	122.2	111.2	96.3	87.8	78.5	70.1	57.1	51.2	45.7	43.1	40.5	36.6	34.1	28.4
Labor statistics	29.2	26.9	26.0	23.6	21.4	18.1	13.8	11.7	11.1	9.4	8.9	8.3	7.6	5.6
ERS	.4	.3	.3	.1	.1	.1	.1	.1	.2	.2	.2	.2	.1	.1
SRS	.3	.3	.2	.2	.2	.2	.2	.1						
BES	3.2	3.1	3.0	2.5	1.7	1.7	1.5	1.3	1.3	1.2	1.2	1.1	1.1	.8
B S	18.4	16.8	16.6	14.9	13.2	11.8	10.3	8.7	8.3	5.7	5.3	5.0	4.5	3.6
OMPER (OMAT)	4.5	4.0	4.0	3.7	4.2	2.5								
Mines	.4	.4	.4	.3	.4	.4	.4	.4	.3	.3	.3	.2	.2	.1
NSF	2.1	2.0	1.4	1.9	1.6	1.4	1.3	1.1	1.0	1.0	.9	.8	.7	
Census										1.0	1.0	1.0	1.0	1.0
Demographic and social statistics	37.0	31.9	24.2	18.3	15.9	14.3	10.7	9.7	8.3	8.8	8.3	7.1	6.0	4.3
ERS	.6	.4	.4	.3	.1	.1	.1	.1	.1	.1	.1	.1	.1	.1
Census	2.6	2.4	1.9	1.8	1.6	1.5	1.2	1.1	1.1	1.3	1.2	1.1	1.2	1.0
NCHS	9.8	8.2	6.9	6.3	5.8	5.1	4.5	4.0	3.0	3.6	3.5	2.6	1.9	1.5
NCES	5.4	3.9	2.7	2.1	1.5	1.3	1.1	.9	.8	.7	.6	.5	.2	.2
SSA	9.9	8.7	6.0	4.6	4.3	4.1	2.8	2.6	2.4	2.3	2.1	2.2	2.1	1.5
WA	2.2	1.9	1.7	1.5	1.2	1.0								
NSF	2.5	2.4	2.3	1.6	1.4	1.2	1.0	1.0	.9	.8	.7	.6	.5	
OEO	4.0	4.0	2.4											
Prices and price indexes	7.4	6.2	5.6	5.7	5.0	4.6	3.9	4.0	3.8	3.5	3.4	3.1	2.8	2.5
ERS	.7	.7	.6	.6	.6	.6	.6	.6	1.6	1.6	1.6	1.4	1.2	1.2
SRS	1.7	1.6	1.5	1.4	1.3	1.1	1.0	1.0						
BLS	4.7	3.8	3.4	3.6	3.0	2.8	2.3	2.4	2.2	1.9	1.8	1.7	1.6	1.3
Census	.3	.1	.1	.1	.1	.1								
Production and distribution statistics	33.1	32.0	28.6	27.2	24.9	23.1	20.2	18.6	16.3	15.5	14.5	13.4	12.8	11.4
ERS	2.1	2.0	1.8	1.8	1.5	1.5	1.4	1.2	6.1	5.8	5.2	4.8	4.5	4.5
SRS	11.9	11.4	10.5	10.3	9.0	8.0	6.9	6.5						
Census	11.1	10.9	11.0	10.7	9.6	9.3	7.7	6.9	6.2	5.7	5.5	4.9	4.7	4.4
Engineers	1.1	1.1	1.1	1.0	1.0	.9	.9	.9	.9	.9	.9	.8	.8	.5
Mines	2.4	2.5	2.3	2.0	2.0	1.7	1.6	1.6	1.6	1.6	1.5	1.5	1.4	.9
ICC	1.2	1.2	1.1	.9	1.4	1.3	1.3	1.2	1.2	1.2	1.1	1.1	1.1	.9
CAB	.7	.7	.6	.5	.4	.4	.4	.3	.3	.3	.3	.3	.3	.2
Transportation	2.6	2.3	.1											

Construction and housing statistics	3.7	3.4	3.2	3.1	3.1	2.5	2.3	1.6	1.3	1.2	1.0	.9	1.0	1.1
Census	2.4	2.3	2.0	2.0	1.8	1.6	1.5	1.3	1.0	.3	.2	.1	.1	.1
FHLBB	.9	.8	.7	.7	.9	.5	.4	.3	.3	.3	.3	.3	.3	.2
HUD	.5	.4	.5	.4	.4	.4	.4							
BLS										.4	.4	.4	.5	.7
BDSA										.2	.1	.1	.1	.1
National income and business, financial accounts	11.8	10.8	8.7	9.9	8.2	7.5	6.2	5.6	4.9	4.7	4.4	3.8	3.9	3.5
ERS	.3	.3	.3	.3	.3	.2	.2	.2	.2	.2	.2	.2	.2	.1
Census	.8	.8	.7	.6	.5	.4	.4	.3	.3	.3	.3	.3	.3	.3
OBE	3.2	2.8	2.5	2.3	2.1	2.0	1.6	1.5	1.4	1.2	1.1	1.0	1.0	1.1
IRS	6.8	6.1	4.5	6.1	4.6	4.3	3.4	3.1	2.6	2.6	2.4	1.8	1.9	1.8
FTC	.4	.4	.4	.3	.3	.3	.3	.3	.2	.2	.2	.3	.3	.1
SEC	.4	.4	.3	.3	.3	.3	.3	.2	.2	.2	.2	.2	.2	.1
Total periodic	40.5	13.7	11.2	22.2	15.0	12.3	10.9	21.4	95.5	13.6	6.4	4.2	9.2	57.2
Decennial censuses [a]	7.7	2.8	2.2	.9	.5	2.8	6.7	18.0	90.1	6.1	4.3			49.1
Economic censuses [b]	7.7	3.3	.8	5.8	7.1	2.9	1.2	2.0	5.1	7.0	1.4	1.6	4.8	7.1
Censuses of agriculture [c]		1.8	5.1	15.5	1.5	.5						.8	4.4	
Censuses of governments [d]	1.1	1.3	.2		.6	.8	.9	.1	.1	.5	.7	.8		
National housing inventory												1.0		
Registration and voting			1.2											
1968 sample survey	20.0	1.0												
CPI revision														1.0
Computer equipment	4.0	3.5	1.7		4.5	4.0								

[a] 1950, 17th Cecennial; 1958–63, 18th Decennial; 1964–68, 19th Decennial.
[b] 1950, 1948 Census of Business; 1956–58, 1954 Economic Censuses; 1957–62, 1958 Economic Censuses; 1962–1967, 1963 Economic Censuses; 1966–68, 1967 Economic Censuses.
[c] 1956–57, 1954 Census of Agriculture; 1963–67, 1964 Census of Agriculture.
[d] 1957–60, 1957 Census of Governments; 1961–64, 1962 Census of Governments; 1966–1968, 1967 Census of Governments.

Explanatory Notes

Figures represent obligations, actual for 1966 and earlier, estimated for 1967 and 1968. Where possible, data are based on Special Analysis for Principal Federal Statistical Programs printed with the Budget. However, since agencies included have changed at times, principally by expansion in 1961, 62 and 63, estimates have been made for certain agencies not shown in that document in earlier years in order to obtain a reasonably comparable series for the entire period shown here. When new agencies have been organized to perform new functions, they have been added—OMPER in 1963, OEO ($2.4 million) in 1966. When new agencies were organized to perform functions previously performed elsewhere, with or without new functions, data are shown under the present name, but carried back to be continuous. (For example, SRS and ERS in Department of Agriculture are shown separately since 1961, but only a single figure for these programs is shown for earlier years. NCHS was organized in 1960; figures for earlier years represent work on vital statistics and public health reports.) In other cases, program content and budgeting has been changed without reorganization. Most notable example is that in 1960 Census took over construction and housing starts, most of which had been in BLS; BLS took over the labor force portion of the CPS.

Abbreviations

BDSA—Business and Defense Services Administration,[a] Commerce Department
BES—Bureau of Employment Security,[b] Labor Department
BLS—Bureau of Labor Statistics, Labor Department
Census—Bureau of the Census, Commerce Department
CAB—Civil Aeronautics Board [b]
Engrs.—Corps of Engineers,[b] Department of Army
ERS—Economic Research Service, Agriculture Department
FHLBB—Federal Home Loan Bank Board [b]
FTC—Federal Trade Commission [b]
HUD—Department of Housing and Urban Development [b]
ICC—Interstate Commerce Commission [b]
IRS—Internal Revenue Service,[b] Treasury Department
Mines—Bureau of Mines,[b] Interior Department
NCES—National Center for Educational Statistics, HEW
NCHS—National Center for Health Statistics, HEW
NSF—National Science Foundation [b]
OBE—Office of Business Economics, Commerce Department
OEO—Office of Economic Opportunity [b]
OMPER—Office of Manpower Policy, Evaluation and Research,[b] Labor Department
SEC—Securities and Exchange Commission [b]
SRS—Statistical Reporting Service, Agriculture Department
SSA—Social Security Administration,[b] HEW
Transportation—Department of Transportation [b]
WA—Welfare Administration [b]

[a] Construction statistics only.
[b] Statistical programs only.

TABLE 2.—*Obligations for principal statistical programs total and selected agencies*

[In millions of dollars]

Agency	Fiscal years													
	1968 (estimate)	1967 (estimate)	1966	1965	1964	1963	1962	1961	1960	1959	1958	1957	1956	1950
Grand total	162.7	124.9	107.5	110.0	93.5	82.3	68.0	72.6	141.2	56.7	46.9	40.8	43.3	85.6
CURRENT PROGRAMS														
Total, current	122.2	111.2	96.3	87.8	78.5	70.1	57.1	51.2	45.7	43.1	40.5	36.6	34.1	28.4
Census	17.1	16.4	15.7	15.2	13.6	12.8	10.7	9.6	8.6	8.6	8.2	7.4	7.3	6.8
OBE	3.2	2.8	2.5	2.3	2.1	2.0	1.6	1.5	1.4	1.2	1.1	1.0	1.0	1.1
BLS	23.1	20.6	20.0	18.5	16.3	14.6	12.7	11.1	10.5	8.0	7.5	7.1	6.6	5.6
BES	3.2	3.1	3.0	2.5	1.7	1.7	1.5	1.3	1.3	1.2	1.2	1.1	1.1	.8
OMPER	4.5	4.0	4.0	3.7	4.2	2.5								
ERS	4.1	3.7	3.4	3.1	2.6	2.5	2.4	2.2	8.2	7.9	7.3	6.7	6.1	6.0
SRS	13.9	13.3	12.3	11.9	10.5	9.3	8.1	7.6						
NCHS	9.8	8.2	6.9	6.3	5.8	5.1	4.5	4.0	3.0	3.6	3.5	2.6	1.9	1.5
NCES	5.4	3.9	2.7	2.1	1.5	1.3	1.1	.9	.8	.7	.6	.5	.2	.2
SSA	9.9	8.7	6.0	4.6	4.3	4.1	2.8	2.6	2.4	2.3	2.2	2.2	2.1	1.5
WA	2.2	1.9	1.7	1.5	1.2	1.0								
IRS	6.8	6.1	4.5	6.1	4.6	4.3	3.4	3.1	2.6	2.6	2.4	1.8	1.9	1.8
NSF	4.6	4.4	3.7	3.5	3.0	2.6	2.3	2.1	1.9	1.8	1.6	1.4	1.2	
Mines	2.8	2.9	2.7	2.4	2.4	2.2	2.0	2.0	1.9	1.9	1.8	1.7	1.6	1.0
Other	11.6	11.2	7.2	4.1	4.7	4.0	4.0	3.2	3.1	3.3	3.1	3.1	3.1	2.1
PERIODIC PROGRAMS														
Total, periodic	40.5	13.7	11.2	22.2	15.0	12.3	10.9	21.4	95.5	13.6	6.4	4.2	9.2	57.2
Census	40.5	12.1	9.5	22.2	14.2	10.9	8.8	20.1	95.3	13.6	6.4	4.2	9.2	56.2
All others		1.6	1.7		.8	1.4	2.1	1.3	.2					1.0

Explanatory Notes for Table 2

Figures represent obligations, actual for 1966 and earlier, estimated for 1967 and 1968.

Where possible, data are based on Special Analysis for Principal Federal Statistical Programs printed with the Budget. However, since agencies included have changed at times, principally by expansion in 1961, 1962, and 1963, estimates have been made for certain agencies not shown in that document in earlier years in order to obtain a reasonably comparable series for the entire period shown here. When new agencies have been organized to perform new functions, they have been added; OMPER in 1963, OEO ($2,400,000) in 1966. When new agencies were organized to perform functions previously performed elsewhere, with or without new functions, data are shown under the present name, but carried back to be continuous. For example, SRS and ERS in Department of Agriculture are shown separately since 1961, but only a single figure for these programs is shown for earlier years. NCHS was organized in 1960; figures for earlier years represent work on vital statistics and public health reports; the latter is not in later years but the National Health Survey is. In other cases, program content and budgeting has been changed without reorganization. Most notable example is that in 1960 Census took over construction and housing starts, most of which had been in BLS; BLS took over the labor force portion of the CPS. Although these changes could be, and were costed out for budgeting purposes at the time they occurred it would have been very difficult to carry them back, so this has not been done.

Census periodic includes funds for computers. Other periodic includes in 1966 and 1967 funds for computers for SRS and NCHS; in 1960 through 1964, funds for last revision of the Consumer Price Index, and in 1950, 1 year's funds for the previous CPI revision.

ABBREVIATIONS

Census—Bureau of the Census
OBE—Office of Business Economics
BLS—Bureau of Labor Statistics
BES—Bureau of Employment Security *
OMPER—Office of Manpower Policy, Evaluation and Research *
ERS—Economic Research Service
SRS—Statistical Reporting Service
NCHS—National Center for Health Statistics
NCES—National Center for Educational Statistics
SSA—Social Security Administration *
WA—Welfare Administration *
IRS—Internal Revenue Service *
NSF—National Science Foundation *
Mines—Bureau of Mines *

Agencies included in "Other" current:
Department of Housing and Urban Development *
Civil Aeronautics Board *
Federal Home Loan Bank Board *
Federal Trade Commission *
Interstate Commerce Commission *
Office of Economic Opportunity *
Corps of Engineers *
Securities and Exchange Commission *
(Business and Defense Services Administration, construction statistics only included in total for years prior to 1960)
Department of Transportation *
WA—Welfare Administration *

* Statistical programs only.

Representative BOLLING. Prior to beginning the questioning, I would like to do a little housekeeping and without objection, have included in the record at the appropriate place extracts from the Federal Reports Act of 1942; from Public Law 784 of the 81st Congress, second session, Budget and Accounting Procedures Act of 1950; from report of the then Committee on Expenditures in the Executive Departments of the 81st Congress, second session, an extract; and from the Federal Register of Wednesday, June 13, 1951, the President's Executive Order 10253; and also the article mentioned by Mr. Bowman in his testimony, "Recent Developments in U.S. Balance of Payments Statistics," which appeared in the May 1967 Statistical Reporter.

Without objection, those will be included.

(The articles referred to follow:)

COORDINATION OF FEDERAL REPORTING SERVICES

§ 139. Declaration of Congressional policy.

It is declared to be the policy of the Congress that information which may be needed by the various Federal agencies should be obtained with a minimum burden upon business enterprises (especially small business enterprises) and other persons required to furnish such information, and at a minimum cost to the Government, that all unnecessary duplication of efforts in obtaining such information through the use of reports, questionnaires, and other such methods should be eliminated as rapidly as practicable; and that information collected and tabulated by any Federal agency should insofar as is expedient be tabulated in a manner to maximize the usefulness of the information to other Federal agencies and the public. (Dec. 24, 1942, ch. 811, § 2, 56 Stat. 1078.)

SHORT TITLE

Section 1 of act Dec. 24, 1942, provided that: "This Act [which enacted sections 139–139f of this title] may be cited as the 'Federal Reports Act of 1942'."

APPROPRIATIONS

Section 9 of act Dec. 24, 1942, provided: "There are hereby authorized to be appropriated annually, out of any money in the Treasury not otherwise appropriated, such sums as may be necessary to carry out the provisions of this Act (sections 139–139f of this title)."

§ 139a. Collection of information.

(a) Duties of Director of the Bureau of the Budget.

With a view to carrying out the policy of sections 139–139f of this title, the Director of the Bureau of the Budget (hereinafter referred to as the "Director") is directed from time to time (1) to investigate the needs of the various Federal agencies for information from business enterprises, from other persons, and from other Federal agencies; (2) to investigate the methods used by such agencies in obtaining such information; and (3) to coordinate as rapidly as possible the information-collecting services of all such agencies with a view to reducing the cost to the Government of obtaining such information and minimizing the burden upon business enterprises and other persons, and utilizing, as far as practicable, the continuing organization, files of information and existing facilities of the established Federal departments and independent agencies.

(b) Designation of central collection agency.

If, after any such investigation, the Director is of the opinion that the needs of two or more Federal agencies for information from business enterprises and other persons will be adequately served by a single collecting agency, he shall fix a time and place for a hearing at which the agencies concerned and any other interested persons shall have an opportunity to present their views. After such hearing, the Director may issue an order designating a collecting agency to obtain such information for any two or more of the agencies concerned, and prescribing (with reference to the collection of such information) the duties and functions of the collecting agency so designated and the Federal agencies for which it is to act as agent. Any such order may be modified from time to time by the Director as circumstances may require, but no such modification shall be made except after investigation and hearing as hereinbefore provided.

(c) Independent collection by an agency is prohibited.

While any such order or modified order is in effect, no Federal agency covered by such order shall obtain for itself any information which it is the duty of the collecting agency designated by such order to obtain.

(d) Determination for necessity of information; hearing.

Upon the request of any party having a substantial interest, or upon his own motion, the Director is authorized within his discretion to make a determination as to whether or not the collection of any information by any Federal agency is necessary for the proper performance of the functions of such agency or for any other proper purpose. Before making any such determination, the Director may, within his discretion, give to such agency and to other interested persons an adequate opportunity to be heard or to submit statements in writing. To the extent, if any, that the Director determines the collection of such information by such agency is unnecessary, either because it is not needed for the proper performance of the functions of such agency or because it can be obtained from another Federal agency or for any other reason, such agency shall not thereafter engage in the collection of such information.

(e) Cooperation of agencies in making information available.

For the purposes of sections 139–139f of this title, the Director is authorized to require any Federal agency to make available to any other Federal agency any information which it has obtained from any person after December 24, 1942, and all such agencies are directed to cooperate to the fullest practicable extent at all times in making such information available to other such agencies: *Provided*, That the provisions of sections 139–139f of this title shall not apply to the obtaining or releasing of information by the Internal Revenue Service, the Comptroller of the Currency, the Bureau of the Public Debt, the Bureau of Accounts, and the Division of Foreign Funds Control of the Treasury Department: *Provided further*, That the provisions of sections 139–139f of this title shall not apply to the obtaining by any Federal bank supervisory agency of reports and information from banks as provided or authorized by law and in the proper performance of such agency's functions in its supervisory capacity. (Dec. 24, 1942, ch. 811, § 3, 56 Stat. 1078.)

CHANGE OF NAME

The official title of the Bureau of Internal Revenue was changed to the Internal Revenue Service by Treas. Dept. Order 150–29, eff. July 9, 1953.

TRANSFER OF FUNCTIONS

All functions of all officers of the Department of the Treasury, and all functions of all agencies and employees of such Department, were transferred, with certain exceptions, to the Secretary of the Treasury, with power vested in him to authorize their performance or the performance of any of his functions, by any of such officers.

Excerpt from Public Law 784, 81st Congress, 2d session, "Budget and Accounting Procedures Act of 1950," Part I, Section 103.

GOVERNMENT STATISTICAL ACTIVITIES

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 for any purpose by the various agencies in the executive branch of the Government. Such regulations and orders shall be adhered to by such agencies.

The report of the Committee on Expenditures in the Executive Departments, 81st Congress, 2d session [Senate Report No. 2031], respecting this provision reads as follows:

Section 104—*Government Statistical Activities*

This section clarifies the present law in accordance with the recommendations of the Commission on Organization of the Executive Branch with respect to the supervision and coordination of the Government's statistical activities, and is intended to be in addition to, and not in substitution for, the existing authority of the Bureau of the Budget with respect to Government statistical and reporting activities.

The following Executive Order is reprinted from the *Federal Register*, volume 16, number 114, Washington, Wednesday, June 13, 1951.

TITLE 3—THE PRESIDENT

EXECUTIVE ORDER 10253

PROVIDING FOR THE IMPROVEMENT OF THE WORK OF FEDERAL EXECUTIVE AGENCIES WITH RESPECT TO STATISTICAL INFORMATION

By virtue of the authority vested in me by Section 103 of the Budget and Accounting Procedures Act of 1950 (31 U.S.C. 18b), and as President of the United States, and in order to carry out the purposes of said section, it is hereby ordered as follows:

SECTION 1. The Director of the Bureau of the Budget (hereinafter referred to as the Director) shall develop programs, and issue regulations and orders, for the improved gathering, compiling, analyzing, publishing, and disseminating of statistical information for any purpose by the various agencies in the executive branch of the Federal Government.

SEC. 2. In order to carry out the provisions of Section 1 of this order, the Director shall maintain a continuing study for the improvement of the statistical work of the agencies in the executive branch of the Federal Government with a view to obtaining the maximum benefit from the funds and facilities available for such work, giving due consideration to the constantly changing character of the various needs for statistical information both within and without the Government and, where the statistical work is primarily concerned with operating programs, giving due consideration to administrative needs, statutory requirements, and the needs involved in the development of administrative and legislative recommendations. The Director, either upon his own initiative or upon the request of any such agency, shall (a) provide for the interchange of information calculated to improve statistical work, (b) make appropriate arrangements for improving statistical work involving relationships between two or more agencies, and (c) assist the agencies by other means, to improve their statistical work.

SEC. 3. The following shall be included among the objectives sought in carrying out the provisions of Section 1 hereof:

(a) To achieve an adequate program of statistical work in the agencies of the executive branch, in relation to over-all needs for statistical information, including the filling of gaps and overcoming of weaknesses in presently available statistical information.

(b) To achieve the most effective use of resources available for statistical work by the agencies, in relation to over-all needs.

(c) To minimize the burden upon those furnishing statistical data needed by the various Federal agencies.

(d) To improve the reliability and timeliness of statistical information.

(e) To achieve maximum comparability among the several statistical series and studies.

(f) To improve the presentation of statistical information and of explanations regarding the sources and reliability of such information, and regarding the limitations on the uses that can appropriately be made of it.

SEC. 4. Regulations and orders issued pursuant to Section 1 hereof shall be signed by the Director. When so signed, such regulations and orders shall require no further approval and shall be adhered to by all agencies in the executive branch. Any such regulation or order may pertain to a single agency, a group of agencies, or all agencies in the executive branch.

SEC. 5. In the development of programs and the preparation of regulations and orders for issuance pursuant to Section 1 hereof, the Director shall consult Federal agencies whose activities will be substantially affected, and may consult non-Federal groups to the extent he finds it necessary to carry out the purposes of this order.

SEC. 6. The authority outlined in this order is in addition to and not in substitution for the existing authority of the Director, or of the Bureau of the Budget, with respect to statistical and reporting activities. To the extent, however, that this order conflicts with any previous Executive order affecting statistical or reporting activities, the provisions of this order shall control.

SEC. 7. Nothing in this Executive order shall be construed to apply to the obtaining or releasing of information by the Bureau of Internal Revenue, the Comptroller of the Currency, the Bureau of the Public Debt, the Bureau of Ac-

counts, and the Division of Foreign Assets Control of the Treasury Department, or to the obtaining by any Federal bank supervisory agency of reports and information from banks as provided or authorized by law and in the proper performance of such agency's functions in its supervisory capacity.

HARRY S. TRUMAN.

THE WHITE HOUSE, *June 11, 1951.*

[F.R. Doc. 51-6883; Filed, June 11, 1951; 3:31 p.m.]

The article below is reprinted from *Statistical Reporter*, May 1967, Office of Statistical Standards, Bureau of the Budget:

RECENT DEVELOPMENTS IN U.S. BALANCE OF PAYMENTS STATISTICS

By JOHN BABYLON, Office of Statistical Standards, Bureau of the Budget

This statement summarizes the changes which have been taking place over the past 2 years in the compilation and presentation of the balance of payments statistics of the United States—changes which have been made to a considerable extent in response to the recommendations of the Review Committee for Balance of Payments Statistics as set forth in its report of April 1965.

Sources of Data

The balance of payments statistics comprise a statistical summary of the international economic transactions of the United States. Their collection is the primary responsibility of the Balance of Payments Division of the Office of Business Economics (OBE) of the Department of Commerce. The official figures are released quarterly by the Balance of Payments Division in the *Survey of Current Business*. Much of the data that enter the balance of payments accounts is not collected by the Office of Business Economics, but is provided by the Census Bureau, the Treasury, and other Government agencies, including all with foreign transactions. The OBE does collect data directly from the public on travel and transportation receipts and expenditures, on institutional and personal remittances, on dollar deposit liabilities to foreigners, and on capital flows associated with foreign direct investment by American firms and various other information on the operations of foreign branches and subsidiaries of U.S. corporations.

The Review Committee

In April 1963, the Director of the Bureau of the Budget appointed the Review Committee for Balance of Payments statistics, citing "the heightened interest in the U.S. balance of payments problem and in the adequacy of our balance of payments statistics as a measure of the problem and a framework within which to consider policy alternatives." The Committee, composed of eight academic and business economists, was chaired by Edward M. Bernstein. Its assignment was to review basic conceptual problems, problems of presentation and analysis, and technical statistical problems of data collection, estimation, and related matters. It was not, as the final report of the Committee emphasized, "asked to determine the causes of the U.S. balance of payments deficit or to recommend remedies for dealing with it." The final report of approximately 200 pages, published in April 1965 as *The Balance of Payments Statistics of the United States: A Review and Appraisal*, contained some 65 recommendations for improvement in the U.S. balance of payments statistics.

These recommendations range widely over all aspects of the accounts. Some are in fairly general terms; others are quite specific. Some involve important conceptual and policy issues while others are concerned with technical improvements in the data or with more effective presentation. In the approximately 2 years since publication of the recommendations, some have been wholly or virtually adopted while a number of others are in process of adoption. Still others are not yet acted upon and await administrative decisions or budgetary support.

Measurement of Surplus or Deficit

In terms of public policy, the most significant development has been the qualified adoption of the recommendation of the Review Committee that a change be

made in the method of measuring the surplus or deficit. The balance of payments statistics are arranged according to a double entry type of accounting system in which each transaction appears as a debit as well as credit entry. For example, transfers of goods or services usually have a counterpart on the other side of the account in capital transactions. For all transactions the accounts must balance; a failure of all known transactions to do so is remedied by inserting an "errors and omissions" entry to achieve the equality. A balance may be struck between the debit and credit entries attributed to any particular category of transactions or group of categories (such as merchandise trade, or all goods and services). Various methods of grouping have been used for the purpose of computing a single balance to indicate whether the international transactions of the country are in surplus or deficit.

The Committee recommended the use of what it called the "official settlements" method whereby surplus or deficit is measured by reserve transactions and, when appropriate, special intergovernmental transactions. This recommendation was a principal subject of hearings held by the Subcommittee on Economic Statistics of the Joint Economic Committee in May and June of 1965. The Subcommittee recommended the adoption of the Review Committee's recommendation. In August 1965 the Cabinet Committee on Balance of Payments, following consideration of proposals regarding Review Committee recommendations formulated by a special committee of Cabinet-Committee agencies, also (*inter alia*) recommended adoption of the "official settlements" measure. The presentation of U.S. balance of payments statistics showing the official settlements balance first appeared in December 1965. The new measure did not replace the previous balance, referred to as the "liquidity balance". Both appear in the official presentation of the balance of payments.

Technical Advisory Committee

A second important recommendation which has been adopted concerns the organizational arrangement within the U.S. Government for the improvement of the balance of payments statistics. Noting that the Bureau of the Budget, by law, is assigned responsibility for coordination of Government statistical activities, the Committee stated that: "In the broader role of seeing that the quality of the data and their presentation are maintained and improved, the OSS (Office of Statistical Standards in the Bureau of the Budget) has not taken the firm lead that is needed to make the system work most effectively." Accordingly, the Committee recommended, as one remedial measure, that the OSS "should chair a permanent inter-agency committee in which issues of data adequacy, statistical classification, and presentation may be discussed." As a result the Technical Advisory Committee on Balance of Payments Statistics was established in October 1965, and has met frequently. In addition to the Budget Bureau, the Committee includes representatives from the Departments of Commerce and Treasury, the Federal Reserve Board, and the Council of Economic Advisers.

Changes in Tabular Presentation

A third recommendation has led to a substantial revision in presentation of the balance of payments statistics. The Committee recommended that the then existing tabular presentation in the *Survey of Current Business* be replaced by a new group of tables, along lines outlined in detail in the Report. Work on revision of the tabular presentation of the balance of payments statistics was carried out within the Technical Advisory Committee. New tables were published in the June 1966 issue of the *Survey*. The new presentation served mainly to eliminate from the summary table any explicit combining of items which would seem to provide a single measure of the U.S. payments balance, and to clearly provide in the summary table a framework from which the supplementary analytical table could be derived.

Five-Year Cycle of Investment Censuses

The Review Committee supported requests made to Congress by OBE in 1963 and 1964 that the staff of the Balance of Payments Division (BPD) be augmented The Office of Business Economics received funds for 13 additional positions in fiscal year 1967. This permanent increase in staff will permit carrying out a project endorsed by the Committee, namely, that the Balance of Payments Division conduct a five-year cycle of censuses of U.S. assets abroad and foreign assets in the U.S. Work on planning for the census of U.S. investments abroad,

the first since 1957 and the principal part of the 5-year cycle, is in an advanced stage. A census of foreign financial assets in the United States will follow.

Foreign Trade Statistics

Several recommendations on which action has been taken are concerned with the merchandise trade statistics in the balance of payments. These recommendations were:

(1) "The Census Bureau (with BPD cooperation) should promptly make studies to discover and measure sources of error in valuation, coverage, timing, and other aspects of trade data for balance of payments purposes." With the cooperation of the Customs Bureau considerable activity has gone into improving the timing of the reporting to Census of import data. A study of sources of error in export reporting by commercial sources on the Shipper's Export Declaration is presently being made. The Review Committee went on to state that, "These studies should include . . . an inquiry as to the adequacy of export reporting on the Canadian border." A study by the Census Bureau of the comprehensiveness of reporting of truck exports along the border is nearing completion.

(2) "The Census Bureau should take steps immediately to make instructions for reporting the value of exports and imports insofar as possible consistent with balance of payments concepts." Amendments to the *Foreign Trade Statistics Regulations* will be issued by the Census Bureau to eliminate the language "cost if not sold" from the instructions for reporting exports on consignment and substitute the correct concept of an approximate sales value of the materials shipped.

(3) The Review Committee in several ways called for improvement in the quality of trade statistics. In addition to the actions noted above which were taken by the Census Bureau and Customs Bureau, the Balance of Payments Division has completed a revision of its end-use classifications for imports and exports. These classifications group U.S. trade by probable end-use in American industry and commerce.

(4) ". . . the findings and recommendations of the Stigler Committee (the Price Statistics Review Committee of the National Bureau of Economic Research, which reported to the Bureau of the Budget in 1960) be reviewed and acted upon promptly" in connection with the construction of true price indexes (as opposed to the currently published unit value indexes) for imports and exports, both in aggregate and by five "economic classes." The Bureau of Labor Statistics currently is conducting a comprehensive theoretical review of the problems involved in constructing such indexes. Funds are included in the President's 1968 budget to further this work.

(5) The Review Committee recommended that the Bureau of the Census "accelerate its efforts to present its foreign trade statistics in a way more suitable for economic analysis." The Census Bureau's Foreign Trade Division has devoted considerable effort to achieving this objective. The publication program of the Division has been changed so as to provide more general-purpose foreign trade data as distinguished from highly detailed commodity-by-country data. The new FT–990 series *Highlights of U.S. Export and Import Trade* reflects this effort. Gaps in the data for Special Category commodity exports, those not shown in full detail because of security restrictions, are being reduced insofar as possible without sacrificing national security. Finally, the publication of an annual foreign trade reference volume, *Foreign Commerce and Navigation of the United States*, as been resumed after an interval of 15 years.

Travel and Transportation

Statistics on travel and transportation receipts and expenditures in the balance of payments were the subject of several Review Committee recommendations and action is in progress on several of these recommendations.

The Balance of Payments Division has recently completed preliminary studies of Mexican border area transactions, and of the characteristics of the border area labor force, with the objective of establishing a system for more accurately estimating expenditures by both Mexican and U.S. border crossers. The travel receipts and expenditures estimates for the United States vis-a-vis Mexico have been "the most unsatisfactory" of the estimates entering the travel account.

The Review Committee also recommended that the Balance of Payments Division should seek to improve the sampling coverage of both U.S. travelers and foreign visitors by more effective ways of soliciting cooperation, by discussion with transportation companies of the possibility of distributing question-

naires to Americans on board planes on their return journeys, or through international agreements for the joint collection of travel data. Approaches to the problems of low response rates and unknown biases in response have taken several forms in the past, and recently, in addition to studies made in Mexico, sampling of United States-Canada travelers has been modified to cover 1-day travelers.

Two Review Committee recommendations dealt with the need to improve the liaison with United States and foreign ocean freight carriers. Staff in the Balance of Payments Division have increased their efforts to obtain reports on Form BE–30, *Ocean Freight Revenues and Expenses, United States Carriers*, on a voluntary basis. If this should not prove rewarding, efforts will be made to make BE–30 reporting mandatory, as the Review Committee recommended.

The Review Committee further recommended that, "The BPD should enforce a satisfactory filing of Form BE–29 (*Foreign Ocean Carriers' Ocean Freight Revenues and Expenses in the United States*) ; if this should prove unobtainable, a pilot study should be undertaken as a step toward collecting both f.a.s. and c.i.f. valuations for U.S. imports to provide a benchmark." The report explains that, "For ocean freight payments, the major payment item in the transportation account, the BPD requires annual reports (Form BE–29) from foreign shipping companies or their U.S. Agents." At present, most major foreign ocean fleets' receipts for hauling U.S. imports are reported by the embassies of the countries to the Balance of Payments Division under an arrangement developed soon after the reporting program was introduced. These "embassy reports" are defective in that they may include double-counting for vessels operated by nationals of one country (one report) but chartered by a shipping company of another country (perhaps another report), and because the reports are not broken down by class of vessel. A satisfactory filing of amended Form BE–29 should correct the latter defect.

U.S. Government Transactions

Steps have been taken to implement the Committee recommendations with respect to the reporting of Public Law 480 (Food for Peace) grant shipments by the Department of Agriculture. These shipments, which are not a result of commercial market transactions, had been valued at CCC full reimbursement cost prices, reflecting the cost of the products to the Commodity Credit Corporation in the oldest crop year from which any substantial quantity of the commodity remains in the CCC inventory plus accumulated storage charges and transportation and other related charges. As a result of this practice, grant shipments had been over-valued (above what they might have brought in commercial transactions) by from 15% to 50%. Instructions issued in December 1966 should result in all Food for Peace grant shipments, which have amounted to about 10% of the value of total Food for Peace shipments, being recorded in both foreign trade and balance of payments data at a close approximation of what such shipments would have been worth in commercial sale.

Significant progress has been made in introducing improvements in the reporting of U.S. Government transactions that affect the balance of payments, a subject of several Review Committee recommendations. One of these, affecting the valuation of P.L.–480 grant shipments, has been discussed above. Another called for "The Bureau of the Budget, the reporting Government agencies, and the BPD . . . to make further efforts to improve and regularize the reporting of the international transactions of the Federal Government." Bureau of the Budget Circular A–65 was issued in July 1964 during the period of the Committee's investigation, replacing a reporting regulations issued in 1944. It provides a legal basis and detailed instructions for reporting to the Balance of Payments Division by the various agencies of their international transactions. This circular, described by the Review Committee as the basis for a fresh start, has provided the focal point for a continuing effort on the part of the Balance of Payments Division to improve the quality and internal consistency of agency reporting. A major step in that direction was recorded last summer when the Department of Defense implemented a new balance of payments reporting system geared to the needs of the statistical program.

The general recommendations concerning the Government transactions account in the balance of payments are being pursued. The Review Committee stated that: "The Defense Department should seek to improve its data on the foreign expenditures of personnel stationed abroad . . ." The DOD balance of payments reporting system referred to above should, over time, substantially increase the quality of such data. In recommending that, "Government agencies, particularly the AID, should improve their data on the extent to which Government grants

and loans are initially spent in the United States or initially spent abroad, and the corresponding data on aid-financed exports of goods and services . . .", the Review Committee's interests and those of the Agency for International Development were closely related, and continuing efforts are being made to improve and maintain the quality of the subject data.

Progress on Other Recommendations

Progres is being made with respect to a number of other Committee recommendations such as:

(1) That Treasury, the Federal Reserve banks, and the Balance of Payments Division should improve review and consultative procedures with each other and with public respondents in the field of investment and other capital transactions. Discussions have been initiated with commercial banks to develop means of improving the estimates of receipts of interest and commissions from foreigners and payments of interest to foreigners.

(2) That improved consultation among the various reporting Government agencies be achieved. The consistency and comparability which Circular A-65 requires has greatly stimulated such consultation.

(3) That "Automatic data processing should be used where it would increase efficiency," a recommendation which certainly was in the spirit of the times as the subsequent 2 years have witnessed an increase use of ADP equipment by the Balance of Payments Division, the establishment of a machine-based "data bank" on U.S. Government foreign grants and credits, discussion of another "data bank" to be used for the collection and dissemination of data on foreign countries' trade and trade barriers, and near completion of arrangements for machine consolidation of Treasury foreign exchange reports.

(4) That "better coordination of the analytical work done by the BPD, the Treasury, the Council of Economic Advisers, the Federal Reserve Board, the Federal Reserve Bank of New York, and other agencies concerned" be established. This recommendation is being achieved through a number of institutions and through the operations of the Technical Advisory Committee on the Balance of Payments.

(5) That Government agencies "must accept greater responsibility for accurate and complete reports".

No action has as yet been taken on several recommendations concerning the capital accounts, on the preparation by the Balance of Payments Division of methodologies describing the compilation of several of the accounts, and on the preparation of other descriptive texts, particularly yearbooks. As mentioned above, the staff increase to date has been directed to work on the capital account in order to handle the various investment censuses. The preparations of methodologies and yearbooks are projects which require the attention of senior Balance of Payments Division staff for an extended period. To date it has not been possible to schedule such projects for early completion. It is anticipated that one or two of the additional positions recently granted to the Division will be utilized for the preparation of a bridge yearbook (covering, for example, developments in the period 1950–66) and subsequent annual yearbooks detailing activities in the U.S. balance of payments. The recommended preparation of a manual of reporting instructions for direct-investment reporters probably will follow a thorough review of all the direct-investment reporting forms which is planned following the analysis of the results of the various investment censuses.

In summary, both the record of the past 2 years and current efforts indicate that the Review Committee's recommendations have received major attention and that considerable change has occurred in the U.S. balance of payments statistics.

Representative BOLLING. Mr. Bowman, prior to the beginning of the questioning may I suggest that the associates whom you mentioned as accompanying you come forward to the table.

Then, prior to recognizing Mr. Curtis to begin the questioning, I would suggest without objection, we dispense with the 10-minute rule. I have always found that it was better in a subcommittee not to have it. So, Mr. Curtis, if you will proceed, unless somebody objects to that, you may carry on for as long as you like.

Representative CURTIS. Thank you, Mr. Chairman, and I am not going to abuse that license. There are a couple of specific questions

that I have in mind. I ask in behalf of Senator Miller, who expressed his regret that he has to leave, How can the the Federal Government aid in the coordinating of Federal, State, and local statistical programs? What can be done that is not already being done? Senator Miller pointed out that the Council of State Governments has now established themselves here in Washington. They have a staff. Couldn't this be utilized and developed along the statistical lines? Are you aware of this and would you comment on this possibility?

Mr. BOWMAN. Yes, Mr. Curtis. I would be happy to do that. The Council of State Governments has been working with the Council of Mayors and the State Governors in our effort to set up a conference on comparative statistics. At one time it was felt that they might have been the secretariat for this conference, but after discussing it, they preferred that we undertake that responsibility rather than that they undertake it themselves. The understanding, however, is that they will maintain an active part in the steering committee of the conference and that all we will do is act basically as the secretariat to the conference. They are actively in this and they are very much interested in its development.

Representative CURTIS. Very good. In our work in the Ways and Means Committee we deal so much with welfare programs, involved as many of them are in the social security laws. One of the things that we have been seeking data on for years is the amount of money that is spent in the field of health and welfare by the nonprofit private sector, Community Chest agencies, church groups, and so on. HEW says that they do not have statistics for this. What has been done toward developing complete statistics of what our total society spends in the field of health and welfare?

Mr. BOWMAN. Well, Mr. Chairman, I would have to check on that, but so far as the major nonprofit institutions like the Community Chest, they are organized on a national basis and their expenditures in this area should be easily available. I would be very much surprised if they are not. I have met with them. I have discussed with them methods which they can use in allocating their funds in various ways and I know that they have an organization which must develop this information, but I do not know it in detail.

(The following additional material was later supplied:)

ESTIMATES OF SOCIAL WELFARE EXPENDITURES

Estimates of social welfare expenditures, private as well as public, have been prepared by the Social Security Administration (see the Social Security Bulletin, December 1966 for the latest figures). State and local funds are included, as well as Federal, in the public sector. It is true that estimates for the private sector are not based on as complete reporting as for the public sector. Additional information on expenditures of non-profit institutions may be obtained from *Giving USA*, an annual report of the American Association of Fund-Raising Councils. From time to time, data for individual communities have been assembled by the United Community Funds and Councils. These provide additional detail for a limited number of communities.

There are many difficult problems in connection with the complete estimating of expenditures for social welfare, not only because basic accounting records may not be maintained upon the most desirable basis, but also because in some programs it is difficult to separate expenditures by purpose—health, welfare, education—and estimates of the contributions made by most volunteers, professional as well as nonprofessional, are not included.

In the health field, the most complete estimates of which I am aware were prepared by the Social Security Administration for the Committee on Ways and

Means on medical resources available to meet the needs of public assistance recipients. These were contained in "Social Security Amendments of 1960." House Report No. 1799, 86th Congress, 1961. It is my understanding that the Social Security Administration has just completed an updating of the part relating to health care through private charity. Furthermore, the whole problem of hospital finances is now under study by the Department of Health, Education and Welfare in connection with the administration of the Medicare program.

Representative CURTIS. Right now Ways and Means is going over this whole field again, particularly with Medicare coming in. We are getting into some of the cost accounting problems with the hospitals. It is quite apparent that statistics have never really been kept, and are not available on the capital investment, let us say, in hospitals. A great deal of the capital investment came from charitable donations and was never put on a cost accounting basis.

Now, where we are seeking to set the fees that hospitals can charge, questions come up concerning the depreciation of capital assets and replacement and possible growth, and so on. There is a real question of whether community chest agencies or, in specific cases, the hospitals themselves, ever had the data.

But I will move on. I know this is taking you far afield but I am concerned with this problem.

We could not even get from HEW the figures on what local governments spent in the field of health. The HEW figures told what State governments spent. For Texas and Missouri, because most of these expenditures occur at the county and municipal levels, there were no statistics. And, of course, the statistics of other States were limited largely to what the State governments spent as opposed to or in relation to what the local government might spend.

Well, I just wanted to open the record on those points and I would appreciate any comments that you could give us.

Now, there is a specific matter you have mentioned in your prepared statement. You point out that Congress has been given responsibility for developing wealth statistics. As you know, this subcommittee held hearings on the wealth of the Nation. "Measuring the Nation's Material Wealth," a report of the Subcommittee on Economic Statistics of the Joint Economic Committee, was issued September 10, 1965. The House Committee on Government Operations has for a number of years published data on Federal real and personal property inventory reports, civilian and military, of the U. S. Government, covering its properties located in the United States, territories, and overseas. The latest I have was as of June 30, 1966. The question, though, is: How far have you or the Department of Commerce gone in developing wealth statistics, and will the inventory be well underway, say, by 1970?

Mr. BOWMAN. I think I would have to answer the question by saying our effort is directed to a program of data collection—where we do not now have data near the end of the decade. That was the way the program was laid out. At the present time the only active program we have is in estimates of wealth in the business sector which are mentioned in my report for the record and which have recently been published in the Survey of Current Business.

Now, it is our expectation to use all the kinds of information which you have just described with regard to Federal wealth.

Representative CURTIS. Yes.

Mr. BOWMAN. But, a considerable amount of work has been done in this area, but much of the basic data needs to be strengthened. It was our hope that this could be done near the end of the decade in connection with the population census and presumably the censuses of 1972, the economic census.

I must call your attention, however, to the fact that our ability to get expansions of our statistical programs in some of these areas makes us wonder whether or not this will be supported when the time comes. I certainly hope it will because I think it is an important enterprise. But, it is a burdensome type of reporting and it is apt to involve some costs—how much I cannot say at this time—because until we get the report from the Office of Business Economics with regard to the general program, which will have to be implemented near the end of the decade, we cannot really estimate it.

Representative CURTIS. I know it does cost. I would make the observation that generally speaking a dollar well spent on economical statistics is a hundredfold returned, at least, and this is an area in which I think we badly need to move forward.

This subcommittee's hearings in 1964 and 1965 were largely based on this report of the wealth inventory plan and study group under the direction of Dr. John Kendrick, then professor of economics at George Washington University. Professor Kendrick's testimony indicated that we had kept up wealth statistics until the early 1920's, but beginning about then we allowed this to fall into disarray. This study that he conducted was one of the first comprehensive studies that had been undertaken in several decades.

Is this about the way the picture has been or is it not quite that bleak?

Mr. BOWMAN. It is the way the picture has been. It is a little more complicated than that and I guess the only thing I would add is to say the reason it was discontinued is because so many people felt that the data we had up to that time were no good.

Representative CURTIS. Yes. That is right. I recall that point in there.

Mr. BOWMAN. And, what we are saying now is to get good data costs money and careful planning. It would also be wrong to say that we have not really developed a great deal of data of broad character in private sources. The National Bureau of Economics Research and Raymond Goldsmith have published balance sheets of the wealth of the United States, but he has to use resources which are still suspect in many areas and we were hoping that with a moderate expenditure at the end of each decade or at the beginning of each decade we could establish the necessary benchmarks which would make our wealth estimates much better.

The two countries that have done most in this area to date are the U.S.S.R. and Japan. These two countries have taken this quite seriously. The U.S.S.R. has it a little easier to do it but Japan has also done it and their problem is just as complicated as ours.

Representative CURTIS. Then, I would like to make a point that we are talking about—and there we were talking about, too—physical wealth. I think most students agree that physical wealth is really the lesser part of the wealth of a society and that the greater wealth lies in the skills and knowledge of the people. How you measure it is somethink else, but it is certainly a very real thing. So I question about

your work here which is limited to measuring physical wealth. You are not attempting to move into this very difficult field of measuring the wealth of the knowledge and skills of people.

Mr. BOWMAN. Not in this immediate project, although we think of this as being a part that has to be added on in any analytical use of a wealth base.

Representative CURTIS. I could not agree more. A number of scholars have been attempting to see how we can possibly measure educational input and things of that nature. I have seen a number of studies along that line. When we talk about physical wealth we should keep uppermost in our minds that this is really the lesser portion of the wealth of the society.

Thank you, Mr. Chairman.

Representative BOLLING. Mr. Rumsfeld?

Representative RUMSFELD. Mr. Bowman, I certainly want to thank you for your very interesting statement. I wonder if you could sketch two or three of the principal areas where you feel there is the most significant data problem or need, for users, both in Government and out of Government. You mentioned—in answer to Mr. Curtis—certain areas that need some thought and attention down the road. What are the areas that are needed most by users within Government and users out of Government that we do not have the capability of providing right now?

Mr. BOWMAN. Well, I will mention one. I will state my own bias in a sense. I generally have an overall view of the statistical program and I think of gaps in terms of the way they are to be filled in in terms of analysis. At the present time the area where we have the greatest need for knowing things requires comparisons over time, what I would call longitudinal studies. I thought we proposed to the Congress this time one of the most important ventures that we have outlined for many years and this was a large sample survey of the population which would get a great deal of characteristics of the population in 1968. We would then pick up the same sort of information in 1970 in the census and find to what extent these characteristics had changed.

The whole area of longitudinal studies of households or individuals, I think, is an area which we must go into. It is a very difficult area, much more than we ever have had in the past. A more pedestrian area, and one which we all know about; we know our data on inventories are not nearly as good as they should be. They are not nearly good enough in the interrelationship of inventories at the raw material level, the manufacturing level, and at the retail level and the way they move, so that for cyclical analysis and other types of analysis of that sort, we get some notion of what is happening to the inventory picture. But, as you can see, these are all very difficult areas and very costly areas in which to work. But, that is another one.

Another area, I think, is more data on the service trades.

Now, my paper does indicate we have started a series on service trade receipts. We have had a series for a long period in manufacturing and in retail trade, but we have very little current data on the output of the very fast growing services sector of the economy. We now do have a series—this series was arranged for 5 years ago, I think. We thought we knew how to produce it at that time. We just published a series about 6 months ago because the difficulties associated with producing

that series were so much more difficult than we had anticipated. But, it is now being produced and, I think, will go forward.

I think another area in which we need data is in labor force for local areas. See particular participation or lack of participation in the labor force in some of the poverty areas of the United States. Here we are in difficulties in really understanding what are the characteristics of nonparticipation in the labor force.

Now, you may have read something recently in the newspaper, that we have known for a long time, but it is now being publicized more than it ever has before and it is becoming more important for analysis: our inability in the census to count all the people in the United States. We think we are missing—we think we missed in 1960 about 5 million people. We think if we took the census the same way in 1970 we would miss about 6½ million people.

You say, Is that very important? It would be——

Representative RUMSFELD. If you will excuse me, that is not what I would say. What I would say if you missed them, how are you able to know that you missed 5 million?

Mr. BOWMAN. Well, I will tell you. We have vital statistics reports. We know how many people were born. We know how many people died. We can tell between two periods how many people there ought to be given age and sex groups. If we do not find that many people, we have a suspicion that we missed them.

When we do not find them in particular categories where we know they should be our suspicion is increased.

Now——

Representative RUMSFELD. So, that figure is a hard figure. Five million.

Mr. BOWMAN. Oh, not real hard.

Representative RUMSFELD. That figure comes from the vital statistics, the gap between what you found and what you should have found by the births and deaths?

Mr. BOWMAN. That is one of the places from which it comes; yes. And on the other hand, when you say "a hard figure," it is as hard as many other figures are, which are estimates of this character and I would not think it would be extremely far off, but it does not do us very much good because it does not tell us what the characteristics of these people are except broadly and it does not tell us where they are. We think most of them are in large cities but we do not know. And the death—the birth and death statistics do not tell us anything about that.

Representative RUMSFELD. Well, every article I have seen, the ones you are referring to, talking about this gap, have drawn some rather firm conclusions as to the characteristics of those people. You are suggesting that you cannot legitimately draw those conclusions?

Mr. BOWMAN. We do not believe we have sufficient evidence to really draw those conclusions firmly with regard to the locations of the country in which these people are located or in some instances just exactly what their other characteristics are. We do believe that it is proportionately larger in the Negro group, especially among Negro males, 20 to 40. But, we know there is a large segment of white population in those same age groups not counted, but just exactly where these peoples are located and what the reasons are, why we do not count them, we do not know, and we are trying to find out.

Representative RUMSFELD. Well, let me ask you a question. If your conclusion is that the 5 million are males, Negro or white, between ages 20 to 40, what has your study shown you when you have checked your estimates against the selective service registrations? There is and has been a Federal law that requires registration of all males in that general age group throughout the 20-year period that would be encompassed in the ages 20 to 40.

Mr. BOWMAN. We are paying more attention to this and checklists are going to be one thing we are thinking of using in connection with the census of population. Do you want to add anything?

Representative RUMSFELD. Let me put it this way. Have you checked this against selective service registers?

Miss MARTIN. There has been a proposal to do this in some test areas, but to my knowledge it has not been completed as yet and I am not sure whether it is even underway at the moment. There are certain problems of confidentiality of the selective service lists which have been a problem in this area.

Representative RUMSFELD. You are suggesting that the Bureau of the Budget is not able to gain access to the selective service registrant lists?

Miss MARTIN. Well, when you say not able, I do not know that we have carried it that far but the Census Bureau has had some conversations with the head of Selective Service and it was indicated that these files are not supposed to be used for any other purpose. This was an assurance that was given to the selective service registrants at the time that the files were established.

Mr. BOWMAN. I think Mr. Rumsfeld, I can answer your question quite positively. No checks, I believe, have been made to date of the selective service records in connection with this particular problem. However, we are working now to try to find out what methods we should use in connection with the 1970 census to try to lick this problem of undercounting and to find out what are the factors which make for it and what are the characteristics of the people that do not get counted, and I am sure the selective service lists, if they can be used will be thought of as one of the resources and are being discussed at the present time.

Representative RUMSFELD. Has your organization been making plans for the effective date of the so-called public records law which will become law on July 4, 1967?

Mr. BOWMAN. Well, the Bureau of the Budget, along with all of the other Federal agencies, has been discussing this issue, I think with the Attorney General's office, with regard to general guidelines. We have had some particular problems that have been raised with us by the business advisory committee which we work with, the Advisory Council on Federal Reports, as to whether in any way this Public Information Act would infringe upon the confidentiality of statistical reports to the Federal Government.

Representative RUMSFELD. This is why I raised the question, to see what kind of problems you might be having with it.

Mr. BOWMAN. Our answer to date has been—we see no reason for changing it—that we think the Public Information Act does have provisions within it which indicate that information provided the Government in confidence by business firms is not subject to release with respect to——

Representative RUMSFELD. There is a specific provision in the bill, exempt from the public disclosure requirements, certain types of information, for the most part voluntarily given.

Mr. BOWMAN. And, until we are instructed otherwise, we will advise the other Federal agencies and the business community that this Public Information Act does not infringe upon the confidentiality of reports to the Federal Government of a statistical character so far as the individual is concerned.

Representative RUMSFELD. This is such a fascinating area. I have a great many questions. Let me just ask one other which is a little off the beam. I serve on the House Committee on Science and Astronautics. We have had before the subcommittee a bill which will bring the Federal Government, specifically the Bureau of Standards, into the business of developing standard reference data, specific information, some of which the Bureau has in the past developed, other information which has been generated through the private sector. It will then serve as a focal point for the development and storage and dissemination of this information. There have been statements made to the effect that this will save in excess of $500,000,000 for private concerns, possibly billions, depending on one's imagination. It struck me if, in fact, the development of this information, and storage, and dissemination of the information is gonig to save the private sector so much money, why does not some ambitious person in the private sector undertake to develop this information or pay the Federal Government for that information that it develops, and then sell it at a substantially reduced rate from the cost per item because it will be disseminated to a substantial number of private concerns that theoretically, by the testimony of everyone before the committee, will have use for the information. I suppose the same question can, in some areas of your work, be transferred over here.

Are there private organizations that do a good job in this area where there is a market for the information and is this private activity growing or dying out? Is there a future for it? Is there some flaw in our private system that prevents or inhibits people from moving into this area? I happen to know the Standard Rate & Data Co. in Illinois has done a good job of accumulating information that is available and selling it at a reasonable profit. They have done a fine job of providing a service for people, of accumulating information and selling it, and in addition, selling advertising in their very fine publications and making additional money on the advertising.

Now, what is going on? Why is not this happening more?

Mr. BOWMAN. Well, it is happening quite a lot and, of course, we recognize the fact that many statistical—most statistical and all Government publications are not copyrighted. I know several areas in which materials have just been taken, reorganized, put into a better package for usefulness of a customer list, and sold. But, I cannot tell—I cannot comment really on the particular Bureau of Standards project that you talked about but I can turn it around to my data center and say why I do not think anybody else, even though there is a big demand for this information, why I do not believe anybody else could take the information and package it through a data center and sell it. It is because we could not give to such a unit the detailed information about individual respondents.

They do do much of this in terms of what I call the macrodata but even there we have a big consumer right within the Federal Government and I think what I would like to do is to serve them economically with the overall body of Federal data and I think we would save money. I do not claim—you know, I do not claim any large figures like you cited but I think we would do a better job at a less cost if we could have the kind of data center I tried to describe to you a little earlier.

Representative RUMSFELD. Returning to the costs, then, for a minute, in our patent programs we have tried to the extent possible, to have them paid by the user, I am told. In the public records law I mentioned previously there is a provision that will permit, and in fact in some cases require, that users of records pay for the cost of producing that record. So user fees are not unknown.

This bill I was referring to, the standard reference data bill, the Federal Government will be paying millions of dollars and will not be receiving from users but a fraction of the cost of accumulating that information.

What is your view on this data center? Do you have in mind that it should be a break-even operation for the Government and that to the extent that people want to purchase this information, that they will in fact pay and pay something approximating the cost, not a profit, maybe, to the Government, but at least the full cost of accumulating it?

Mr. BOWMAN. This is a point I should have mentioned. We would think of the data center serving its users, charging what I would call as an economist marginal cost of the operation. In other words, the direct cost associated with servicing a customer. We would not think of them as being charged an overhead which went right straight back to the original collection of the data. We think that is a function that the Federal Government, for a wide variety of reasons, is making available to the public generally. But any special costs associated with special tabulations, with the development of special tapes, with special searches, we would think of as being paid for by the user.

Now, we would have—there are difficult problems here. In other words, suppose you as a professor at Harvard University come in with a special tabulation and we tell you it will cost you $10,000. You probably got a grant and you say, "OK, I will pay the $10,000." Suppose later on somebody else comes along and wants the same kind of a tabulation. Do we charge him $10,000, too, when maybe it only costs a hundred dollars to run the same tape that we had prepared for you originally?

So, a data center will have to try to work out whether or not there are some packages in which the costs can be distributed.

Now, remember, my interest is making information as available as possible because in my opinion, this is what promotes the growth of the American economy. So, I am not trying to hide it once we have it. But, I do agree there is no reason why the special user should not pay all the costs of the special use and the only thing I am saying is there are some special problems even in fixing the special costs.

Representative RUMSFELD. Thank you. I will stop there, Mr. Chairman.

Representative BOLLING. Mr. Bowman, I discovered a long, long time ago when I was chairman of this subcommittee for a number of

years, that after all, all the things that we talked about in the subcommittee and all the hearings that we held, our unanimous recommendations and total agreement between this subcommittee and the full committee and the executive agencies involved in this particular problem, came to naught if we did not get adequate appropriations. We could have perfect unanimity, seemingly, and if we did not get any money to funds the programs, we did not get very far in terms of improvements in individual programs and in terms of better coordination of the program.

When I was chairman of the subcommittee, I followed what happened in the Appropriations Committee and on the floor of the House and Senate very carefully. However, I would confess that this year I have not done so; and I wonder if you and your associates are prepared to give us a rundown of what has happened so far to the recommendations that were made to the Congress and the actions to date of the Appropriations Committee in the House of Representatives? I do not think the Senate has acted on this.

Mr. BOWMAN. I think I can do that. I may have to call upon my associates a little bit. So far as the census budget is concerned in statistics, it has been acted on by both the House and the Senate. And it was a small budget in all respects except one, and there we had this large sample survey which I said I thought was one of the best things that we had developed for a long time in the light of our present problems. It called for a survey of 3 million households. And it would have cost about $23 million, with some funds required in fiscal year 1969.

We put it into a double package. There was a request for a supplemental to the 1967 budget of $1 million in order to finance—to do the advance financing on this and then I think an item in the regular 1968 budget of $20 million.

The situation is that the supplemental has been turned down by both the House and the Senate. The regular request, the 1968 budget request, has been turned down by the House but the Senate has not reported yet. And, I believe it will——

Representative BOLLING. Please clarify that.

Mr. BOWMAN. They are meeting later this month.

Representative BOLLING. You said they have not acted yet?

Mr. BOWMAN. On the regular budget request; no.

Representative BOLLING. I see.

Mr. BOWMAN. And, I believe the appeal will be for approximately half that amount and a curtailed survey to about a million and a half households as contrasted with 3 million.

Representative BOLLING. Before you proceed to another area, if I remember correctly what you said, this survey of 3 million households in 1968 was planned so that particular characteristics of individuals could be compared with the quinquennial census of 1970. So, in fact, if money is not appropriated there will be no reasonable comparison.

Mr. BOWMAN. That is right. I only stressed one element. There are a lot of other things where the survey would serve as well as that.

Representative BOLLING. I understand that, but the point is that if this money is delayed now, it will be impossible to implement the program.

Mr. BOWMAN. That is right. The only thing we could do later on is propose a survey in 1972 for which we could make similar comparisons, but——

Representative BOLLING. Well, now, would that work as well? Will your going from the larger to the smaller work as well?

Mr. BOWMAN. Yes.

Representative BOLLING. It would?

Mr. BOWMAN. I think it could be done that way.

Representative BOLLING. I just wondered if there are any technical problems in going from the big census to the smaller survey.

Mr. BOWMAN. I think the reason, Miss Martin is explaining to me the reason why we are curtailing the appeal to the Senate from the larger survey to a somewhat smaller is because so much time has elapsed now and without getting the supplemental funds we would not be able to conduct it.

Representative BOLLING. OK. Go on on the other one. I did not mean to interrupt you.

Mr. BOWMAN. The other one is the Labor Department request. We had what we thought was an excellent program in there. One of them—one item was a million dollars for some further detailed study of employment and unemployment in large metropolitan areas, similar to the type of survey which is made in the CPS monthly. This was denied. We think this is very important. The second item that I can remember that was denied, in fact, I think there was only $340,000 of increases approved out of a budget request for increase of about $2 million, but the other main item, which was missed, which is very dear to my own heart, although I recognize there are different positions on it, is the sector price one.

The House did approve some money for quality improvement in price work but this and the preparatory work for revisions of CPI weights were the only two items that were approved by the House.

Representative BOLLING. Based on the figures that you appended to your statement, showing a very substantial increase from 1950 to the present, this is not what you would call one of the good years.

Mr. BOWMAN. Not——

Representative BOLLING. So far.

Mr. BOWMAN. So far, that is right.

Representative BOLLING. Thank you very much, Mr. Bowman. That is all I have.

Mr. Curtis?

Representative CURTIS. I could not agree more with the Chairman on his comments. This leads me to a point that I have been very much interested in. You comment on the job vacancy statistics. This subcommittee actually held specific hearings in this area because it was our judgment that this is very important. We wanted to be sure that the series would be practical and that industry would cooperate. The hearings clearly revealed that it is practical. It is a necessary thing for the Manpower Training Act or any of the training programs to be meaningful. In fact, we wrote it into the Manpower Training Act. To my regret the administration withdrew its request—I think that was a $2½ million request—on the grounds that Congress had denied it twice before. Therefore, they just did not ask for it again. Well, I have had a running colloquy with Secretary of Labor Wirtz on this subject and I went to the trouble—at least from my side of the aisle—

of being sure that the votes were there in the Subcommittee on Appropriations. In my judgment the votes are there. If the administration would push for it they would get it. But the hearings also revealed that the AFL–CIO is opposed to it. They were quite frank in their testimony.

They said if these statistics were available they would be misused, might create an erroneous impression about the problem of unemployment. I am sorry when you point out that funds have not been provided by the Congress to make such surveys operational. But, it should be added that the administration this year did not even request them. I am very, very disturbed about this.

Mr. BOWMAN. Mr. Curtis, you are right. We did not request them this year because we had requested them in two previous years and they had been turned down and it has always been our policy not to include items in budgets that you do not think there is very much chance of your getting. We may have been unwise. I do not know.

Representative CURTIS. Yet, this was after this subcommittee specifically got into this area so that we would have a better understanding of its needs. After this subcommittee did all this work and got the unanimity, the administration withdrew the item. I suspect it was withdrawn not because the administration could not have gotten it through the Congress but because the AFL–CIO did not want it. Until I see a better response to my charge, I will have to conclude that this was the case.

In your statement you refer to standard occupational classifications for statistical purposes. Is this the same thing, that we refer to as the Dictionary of Occupational Titles? Is this a supplement—would you comment on how it relates to the Dictionary of Occupational Titles? Mr. BOWMAN. It is really not related to the Dictionary of Occupational Titles at all. Standard industrial classification is a——

Representative CURTIS. No. This says standard occupational classification.

Mr. BOWMAN. I am sorry. I thought you said standard industrial classification.

Representative CURTIS. No.

Let me read the full sentence:

> At the present time, the Budget Bureau is undertaking a preliminary review to determine whether is should attempt to develop a standard occupational classification for statistical purposes.

Mr. BOWMAN. It does relate to the Dictionary of Occupational Titles but it would not be as detailed as the Dictionary of Occupational Titles and it would organize the data, we hope, so that it would be possible to cross over from one classification to another, but the statistical classification would be designed specifically for statistical uses rather than for placement or other types of uses which is the main reason for the Dictionary of Occupational Titles.

Representative CURTIS. Yes. Is there any attempt to carry this Dictionary of Occupational Titles over to the military establishment's use of skills which have their counterpart in the civilian sector? I think 80 percent of the skills the military needs have their counterpart in the civilian society.

Miss MARTIN. The military does now use the Dictionary of Occupational Titles in classifying the military as well as the civilian people

on their rolls as an additional item of information beyond their military occupatonal specialty. This attempt to develop a standard classification is an attempt to relate the various classifications in use within the Government and outside, like the Dictionary of Occupational Titles and the statistical classification that is used in the census because usually when you want to use this information, you want to use it related to census information about the population as a whole.

We would hope that a standard classification would make it possible to do this easily. The interagency committee which is working on this has representatives on it of each of the military services as well as the Bureau of Employment Security that develops the Dictionary of Occupational Titles, Census, Office of Education, Civil Service Commission and others. In fact, there is a wide interest in this subject in the Federal agencies.

Representative CURTIS. I am pleased to hear this. As I have said before I think the Dictionary of Occupational Titles should be in looseleaf form. With technological advancement the way it is, there are great changes which are not adequately treated by the Dictionary now.

One other item that relates to this is the matter of educational statistics and training. Sylvia Porter in a column just recently, which I put in the Congressional Record, by the way, pointed out that private industry spent $18 billion last year for on-the-job training. I remember seeing some attempts to figure out how much this was 3 or 4 years ago. It was around $14 billion. It is probably very difficult to know what this figure is. Would you comment on whether or not there is any attempt being made to follow the amounts of money spend for training and retraining? Or do these just have to be rough estimates each time?

Miss MARTIN. I am not familiar with those specific estimates. I do know that the Labor Department has been concerned with the lack of information. They have made one attempt to—one trial, a pilot study, in which they found it very difficult to collect this information and they are now looking into other possibilities, possibly not trying to do as thorough a job and being able to collect some of the major elements. But, at the moment I know of no actual data that are available.

Representative CURTIS. These are some specifics. You talk about housing starts. I remember some years ago I asked whether or not we were including mobile homes in our housing statistics. I think that in recent years they have been included. Certainly mobile homes are numerous enought today to be measured. Would you comment on that? How do housing starts relate to the purchase of mobile homes?

Mr. BOWMAN. That is right, but we do count in the statistics the anchored mobile homes, the ones that are being used as homes. I cannot remember in detail now the way in which this estimate is prepared but normally housing starts would not involve that unless, of course, a permit might be issued to build a back porch on the back of a trailer and in that sense——

Representative CURTIS. Yes. I wonder if you could supply it for the record.

Mr. BOWMAN. Yes, sir.

(The material which follows was subsequently furnished by Mr. Bowman:)

TREATMENT OF MOBILE HOMES IN HOUSING AND CONSTRUCTION STATISTICS

Placement of mobile homes on sites are not considered as housing starts in the series compiled by Census. Sometimes local authorities may consider the amount of construction associated with placement of a mobile home, such as preparation of foundation, or building of a porch or another room, as sufficient to require a building permit. However, checks made by Census indicate the number of starts on this basis included in the housing starts series is negligible.

The construction activity series conceptually includes the value of foundation work, installation of utility lines, etc. entailed in the preparation of sites for mobile homes. However, because of inadequacy of data sources for measuring this activity, relatively little actually gets included. The value of the mobile home itself is not included as part of construction activity but is in manufacturing.

A mobile home which is occupied by persons for whom it is the regular place of abode is counted in the inventory in the Housing Census. Those which are vacant and those found occupied by persons who claim another regular place of residence are not counted.

Monthly data on manufacturers' shipments of mobile homes are compiled by the Mobile Homes Manufacturers' Association and are published by BDSA in *Construction Review* and 217,000 were reported for 1966. In addition, there were shipments of approximately 122,700 travel trailers.

Representative CURTIS. Then, one other item. At one time, I got into the monthly corporation profit statistics that are furnished this committee. We have the gross profits and we also have figures on the ratio of profits to sales, dollar values of sales, but I think the most important statistic is the return on equity investment. Granted this gets back into this study of wealth because it is difficult to know what the equity is. The management of a business always looks to the statistics on return on equity investment in determining whether to invest in expansion. What will they get on return? We have some broad figures showing that the return is 14 percent or maybe go on down to 8, around there. Would you comment just briefly on this area?

Mr. BOWMAN. I can comment on this briefly and not so much in terms of the statistics themselves but with the problem. It is true that in ordinary economic analysis we talk about the profit return on investment but the difficulty with equity is that in one industry the capital investment may be largely financed through bonds, in another industry it might be largely financed through equity securities. You compare the two industries on the basis of return on equity, you do not really get a good measure of the comparison between the two industries with regard to their return on the capital investment.

Representative CURTIS. And yet, though, as they pay off on debt, they increase the equity, so that it is an important factor to know. I again emphasize that it relates to whether they will expand, even if it is an expansion financed by debt, they still are looking to what the return is going to be on that dollar. Ultimately they hope to convert the debt to equity as they are paying it off.

Mr. BOWMAN. This is true. You happen to touch me on a spot of major interest because my doctoral dissertation was a statistical study of profits in which I did try to examine the various bases that might be used for comparing profits in different industries.

I still would feel that while equity is a useful way of comparing profits, that it will vary from time to time. depending upon the market for equity shares as contrasted with debt shares. It will not be the same

in different industries because of the way in which different industries can be financed. And all I can do is to give you a personal opinion now and not an official one.

If I were going to make a basic study of profits, not from the standpoint of investment, myself, but as an economic study, I think I would take profits on income originating as contrasted with profits on either sales or any of the other bases. We have some problems there but I think the national income and product acount which allows you to look at different industries on profits in terms of income originating is the most useful one for broad economic analysis.

Maybe my colleague, Mr. Moss, will differ with me on this. He is perfectly able to speak if he wishes to. Do you want to comment?

Mr. Moss. No. I would not argue on this basic question, but I will say that in connection with the wealth project, that looking to the long term, by obtaining national and sector balance sheets, data on debt and on tangible capital and hence on equity, would be provided which would give us a better picture of trends in equity, debt, capital, et cetera, so that these relationships certainly could be studied. But this is looking to the longer term.

Representative CURTIS. Possibly I am more conscious of this because in the Ways and Means Committee, there are so many times when in considering our tax laws we see whether a corporation finances its growth through retained earnings, new equity issues, or from debt. And then, also when we consider renegotiation laws we get into this problem of how you relate the methods of financing economic growth. I think it is most important to this committee. Another place where it comes in is in our financing operations abroad, the interest equalization tax. Here we closed up equities and increase through debt and then we close up debt and it came in holding retained earnings abroad rather than following the pattern of bringing it back to the United States. It becomes important to the Joint Economic Committee in trying to figure out what is going to happen as far as new investment is concerned. There it really is a question of what they think they are going to get in return for a dollar invested. Could they get more by investing it abroad than they could in our own economy?

Well, I have explored this to my satisfaction for the time being.

One other general comment. The Department of Health, Education, and Welfare, I notice, has withdrawn its monthly publication entitled "HEW Indicators." They are keeping the annual one but I have noticed what I thought was a deterioration of this publication over a period of years, at least as a statistical series.

Have you any comment as to why that has been abandoned? I thought they had some very valuable statistics that they published and were helpful on a monthly basis.

Mr. BOWMAN. When it was first inaugurated I took a major interest in it. In fact, I think we encouraged it. I have not maintained my familiarity with the recent changes and the person in my staff who does is not here with me today. I do not believe—do you know what happened in this area. Miss Martin?

Miss MARTIN. No; I do not.

Representative CURTIS. Would you supply for the record a statement as to what is the situation?

Mr. BOWMAN. We will be happy to do that.

(The following was subsequently supplied for the record:)

The Department of Health, Education, and Welfare discontinued the publication of the monthly *HEW Indicators* with the February 1967 issue.

The DHEW is engaged in an attempt to develop a broader and more useful set of social indicators. Many of these series do not show significant changes on a monthly basis. Therefore, the staff resources made available by the discontinuance of the monthly *Indicators* are being used to prepare a considerably expanded edition of the annual *HEW Trends*, which will again this year be published in two parts—part 1 related to national data covering an extended period of years, and part 2, State data and State rankings. We agree that this seems to be an appropriate decision.

Plans for the new edition of *Trends* include a number of additional tables and charts as well as extensive revisions in some of the continuing series.

Representative CURTIS. I do have another question. I was told of someone else's very good suggestion. Professor Stephan of Princeton told the subcommittee he favored better interagency cooperation in the statistical field but that he did not favor the immediate establishment of a national data center. He envisioned a limited and more deliberate program of combining data from diverse sources rather than the sudden establishment of a central center. What would you think of that approach and what would be the best mechanism for carrying it out?

Mr. BOWMAN. Well, Mr. Stephan and I have discussed this on many occasions. I think the best thing I can say now is that I disagree with him.

Representative CURTIS. Thank you.

Representative BOLLING. Mr. Bowman, we are grateful to you and your associates for being with us and for your very helpful contribution to this set of hearings.

The subcommittee will next meet at 10 a.m., tomorrow in room 6226, New Senate Office Building. The subject will be the "Long-Run Possibilities and Problems." On "Statistics for Effective Public Policy" will be Arthur M. Okun, member of the Council of Economic Advisers. On the "Goals and Difficulties of the Government Statistical Program" will be our old friend, Ewan Clague, formerly Commissioner of Labor Statistics.

Representative CURTIS. Mr. Chairman, I presume there will be the usual opportunity to supply additional questions for the record?

Representative BOLLING. That has been made part of the rules. That is certainly so.

The subcommittee will now stand adjourned.

(Whereupon, at 12:05 p.m., the hearing was adjourned, to reconvene at 10 a.m., Thursday, June 8, 1967.)

(Chairman Talmadge's letter to Mr. Bowman and his subsequent reply follow:)

JUNE 15, 1967.

Mr. RAYMOND T. BOWMAN,
Asistant Director for Statistical Standards,
Bureau of the Budget,
Executive Office Building,
Washington, D.C.

DEAR MR. BOWMAN: This is in reference to your very excellent testimony before the Subcommittee on Economic Statistics. I regret that it was not possible for me to attend the hearing on that particular day, although I have given careful attention to your stateemnt and to the discussion which ensued.

Most of the questions which I had planned to ask were covered in your testimony. However, several additional pieces of information would greatly facilitate our deliberations and I wonder if you would please furnish your response for the record.

I. The coordination and integration of statistical series within the Federal Government:

1. In your view, is the machinery provided by the law adequate for the tasks of coordination?

2. What have been the major obstacles in the way of making series compatible—for example, such that productivity and prices could be analyzed accurately?

II. The coordination among users and producers of data:

1. How do the Office of Statistical Standards and the various agencies weigh the needs of users?

2. Do you feel that there is a need for an *index* concerning the availability of Federal data including a description?

3. When changes in definition are being considered, is there a formal hearing of the affected parties?

4. Have there been attempts to measure the burden to respondents in complying with requests for information? Is this burden equitably distributed?

III. The storage and retrieval of data:

1. Are there standards set down by the Office of Statistical Standards to insure that the micro data in various agencies are stored in machine readable form and can be utilized by researchers inside and outside of government, with due consideration for disclosure problems?

2. Are the statistical agencies making full use of modern technology in the processing and storing of data?

IV. The advisability of a "National Data Center" or "Statistical Servicing Center":

1. Can major progress in coordination and integration be made without a National Data Center?

2. Would a pilot project be advisable?

3. If so, what magnitude of investment would be required for a meaningful pilot project?

Thank you for these additional responses and for your very helpful contribution to these hearings.

Sincerely,

HERMAN E. TALMADGE.

EXECUTIVE OFFICE OF THE PRESIDENT,
BUREAU OF THE BUDGET,
Washington, D.C., June 28, 1967.

Hon. HERMAN E. TALMADGE,
Chairman, Subcommittee on Economic Statistics, Joint Economic Committee, U.S. Senate, Washington, D.C.

DEAR SENATOR TALMADGE: Enclosed are my replies to the questions presented in your letter of June 15, 1967. I am very glad to submit this additional information for the record of the hearing.

Sincerely yours,

RAYMOND T. BOWMAN,
Assistant Director for Statistical Standards.

(Reprinted below is the reply of Raymond T. Bowman, Assistant Director for Statistical Standards, Bureau of the Budget, to questions presented in a letter from Senator Herman E. Talmadge, Chairman, Subcommittee on Economic Statistics, Joint Economic Committee.)

I am very glad to supplement my testimony before the Subcommittee on Economic Statistics for the record by answering the questions in your letter to me of June 15, 1967. The questions are repeated before each answer.

"*I. The coordination and integration of statistical series within the Federal Government*

Q. 1. In your view, is the machinery provided by the law adequate for the tasks of coordination?"

A. 1. In my opinion the authority provided by law (Section 103, Budget and Accounting Procedures Act of 1950, and its implementing Executive Order 10253 and also the Federal Reports Act of 1942) is generally adequate for carrying out the statistical coordination function. The machinery established within the Bureau of the Budget has worked well using the principles described in my testimony. The need for well coordinated and appropriately interrelatable statistics, as your Committee recognizes, has not diminished, however, but increased. We recognize, therefore, that some new arrangements and some associated increase in coordination authority in specific areas may be necessary. It is for these reasons that the Bureau of the Budget has been considering a possible proposal for a Federal Statistical Data Center.

Q. 2. "What have been the major obstacles in the way of making series compatible—for example, such that productivity and prices could be analyzed accurately?"

A. 2. It should be recognized at the outset that the compatibility of statistical series for interrelated uses in various aspects of economic and social analysis involves serious conceptual problem. Authority to coordinate is essential but it cannot solve such problems by fiat alone. Since the question seems to be directed more to the issue of obstacles arising from possible inadequacies in coordination authority or procedures I shall not dwell on the conceptual problems although, in my opinion, they are important.

Currently three aspects of present arrangements and authorities make the job of promoting improved compatibility of statistical series more difficult than, in my opinion, is necessary. First, there is no continuing and systematic present way of bringing together and critically examining, in the light of major interrelated uses, the important statistical series of the Federal Government in such a way that the coordinating agency can be close to the actual process. Second, many problems of compatibility center around the consistent and uniform application of industry definitions. Third, there is the inability to secure necessary financial and manpower resources to carry forward integrated programs of work.

The first difficulty mentioned above is a major reason we are interested in the idea of a Federal Statistical Data Center. I covered this in my testimony (page 86) but reemphasize it here. A statistical data center with responsibility for storage and planned retrieval of major series collected by different agencies but required for various interrelated statistical analyses would provide a "feedback" of information on deficiencies in compatibility that could guide the coordinating authority. Need for action to make files compatible would be clearly established on a continuing basis and made specific.

This does not mean that a Federal Statistical Data Center would undertake the coordinating work. It does mean that a strong tie should be established between such a Center and the coordinating authority so as to improve the operations of both. The usefulness of a Federal Statistical Data Center does not mean either that effective actions are now impossible. The Bureau of the Budget has been and is currently

working to bring about improved compatibility of statistical series as I explained in my prepared testimony. We believe our efforts have been successful in the past and that our current projects in the area of production and prices will also be successful. I do believe a Federal Statistical Data Center would significantly improve our capabilities.

The second difficulty mentioned concerns the many problems of comparability that arise in attempts to make interrelated uses of data about a specific industry. Such comparisons were significantly improved by the development of the Standard Industrial Classification (commonly called the SIC) by the Office of Statistical Standards. This provided standard definition of an industry. It is true, neverthless, that agencies may still differ, sometimes on the basis of different information at different times, in their assignment of establishment to a particular industry. Largely because of problems associated with confidentiality of SIC designations, we have been unable as yet to set up for the Federal Government a directory of enterprises and establishments indicating the SIC code assigned by a selected lead agency. Such a director if it were established and maintained in machine readable form and were readily available to all Government agencies could provide a useful checkpoint for establishing the uniformity of industry classifications. It would make it possible to discover the validity of differences over time and would also provide valuable information concerning the effects on industry classification of the use of enterprise units as distinct from establishment units. While large elements of a directory of business units are in existence and are used, no easily accessible directory of the type required has been developed because of problems associated with the interpretation of confidentiality. In my opinion, a directory list of firms and establishments is a must, not only for the reason given but also to establish an economical and consistent frame for selecting industry samples. Here again, we think such a directory could be developed as one function of a Federal Statistical Data Center. If such a Center does not prove feasible then other methods must be found.

Finally, the third difficulty standing in the way of better interrelated statistical series is the matter of financing simultaneously the important tie-in elements of a statistical program. Although the special analysis recommended by your Committee and now a regular part of budget presentations shows the statistical program as a whole, important components of such a program may not receive the necessary appropriations. An example can be taken from one of your questions. Measure of productivity for an industry require, of course, that the output of the industry be comparable with the inputs of labor or capital, or both. But output often has to be measured in current dollars and capital inputs in terms of original costs. In order to measure productivity we must be able to deflate or inflate reported value measures to obtain an approximation to real output. To accomplish these adjustments requires accurate price indexes organized along industry lines. Funds have not been provided by the Congress for sector (industry) price indexes and until they are, the use of inferior price indexes will continue to throw doubt on the validity of industry productivity measures. Measures of capital as an input in real terms also suffer from the deficiencies of price indexes for capital goods.

In brief, I believe the factors I have mentioned outline the principal institutional reasons why the job of coordination is difficult and will at least be something less than perfection. I would be remiss,

however, if I did not point out once more that different types of analyses require somewhat different conceptual frameworks and ideal comparability among series for all purposes is not possible. Comparability for the recognized important uses can, however, be improved and this is what we try to do.

"II. *The coordination among users and producers of data.*

Q. 1. How do the Office of Statistical Standards and the various agencies weigh the needs of users?"

A. 1. The Office of Statistical Standards and the major statistical agencies of the Federal Government attempt to stay alert to the changing needs of users through formal and informal devices. Among these are the establishment of standing advisory groups representing different user interests, the creation of *ad hoc* committees or panels on specific questions, staff participation in professional societies which represent major users of economic or social data produced by the Government, and consultation with experts within and outside the Federal Government. We give great weight to the recommendations of the Joint Economic Committee and the Council of Economic Advisers for improvements needed in the Federal statistical system in order to make it responsive to the needs of economic analysts.

The Office of Statistical Standards has the following standing committees representing non-Federal groups:

Advisory Council on Federal Reports, which provides advice of business enterprises;

Labor Advisory Committee, which reflects needs of labor union research directors for data;

American Statistical Association Advisory Committee on Statistical Policy, which reflects the views of outstanding statisticians on general needs and appropriate methodologies.

Examples of more specialized advisory groups on the needs of users are:

Federal Agency Council on the 1970 Censuses—specifically appointed to collate and coordinate Federal agency needs for information from this source;

Petroleum Statistics Study Group, which reflects the needs of a particular industry;

Review Committee for Balance of Payments Statistics, which assessed needed improvements in a subject field;

Interagency Committee on Foreign Trade Statistics, a standing committee which reviews needs for improved data in this area;

Technical Committee on Standard Industrial Classification, an interagency committee which reflects Federal agency needs in the development, revision and application of the Standard Industrial Classification;

Committee on Labor Supply, Employment and Unemployment Statistics, an interagency committee which provides the Office of Statistical Standards with technical advice on all aspects of employment and unemployment statistics and acts as a clearing house for Federal agencies producing and using such statistics;

Occupational Classification Committee, an interagency committee now working on the development of a standard occupational classification to meet Federal needs;

Interagency Committee on Measurement of Real Output, which has been appointed to develop comparable series in this area;

Consumer Expenditures Survey Committee, a committee specifically assigned the task of assessing Federal needs for such data on a continuing basis.

The Office of Statistical Standards also consults with the Federal Statistics Users Conference and participates in its sessions.

Of course, the advice of different users may conflict—because for different purposes somewhat differing data or concepts are needed; because one user is more interested in historical continuity and another in current measurement only; because one is interested in a quick summary report for a current indicator and another insists on geographic and industry detail, etc.

In addition to the needs of the users, the burden on respondents, the costs to the Government and the technical feasibility of alternative ways of meeting the needs must also be considered. Under the circumstances, there is no easy formula or set of guidelines for weighing needs. In my prepared testimony, I pointed to the use of the national accounts framework as one set of guidelines, which helps us weigh the importance of certain types of data, but for many areas those are not relevant. We try to assess the urgency of expressed requirements, and in particular, to anticipate developing demands for data, at the same time keeping in mind relevant costs and the availability of feasible alternatives for satisfying all or part of the need. In this process, which is obviously an art rather than a science, the hearings and advice of the Subcommittee on Economic Statistics have been particularly helpful in assembling the views of a wide group of users and in focusing on the gaps and deficiencies of the statistical system most in need of improvement.

The problems are particularly difficult when the needs of users conflict. For example, the number of questions which can be asked in a census is limited by considerations of cost, burden on respondents and impact on quality of results. Thus evaluation of such issues requires that we ask questions such as: should this limited space be devoted more to the needs of Federal or local agencies for data required to administer government programs? To the needs of economic, demographic or social theorists for new types of information? or to the needs of business for more detailed geographic data on consumer markets? It is difficult to conceive of any easy system for resolving such conflicting interests.

Q. 2. "Do you feel that there is a need for an *index* concerning the availability of Federal data including a description?"

A. 2. Yes, a really first class comprehensive index would certainly be useful. The preparation of a comprehensive index to Federal statistical data has been proposed from time to time. We have never had the resources to incur the very substantial costs involved in developing and maintaining a really meaningful index on a current basis.

If a Federal Statistical Data Center is established, its responsibilities should include preparation and issuance of appropriate guides or indexes to the data available in the Center; and, by extension, to other bodies of data that could be related to those in the Center. A comprehensive index is a logical part of the work of a Statistical Data Center and its general reference and service function.

It should be noted, however, that the major Federal statistical agencies all publish catalogs, price lists, and other types of guides to their statistical publications. Generally, these are arranged according to subject, and alphabetical indexes are provided where needed. 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.* An up-to-date revision of this bibliography is now in preparation and will be available shortly.

Q. 3. "When changes in definition are being considered, is there a formal hearing of the affected parties?"

A. 3. Changes in statistical definitions are made only after a considerable amount of consultation with both the producers and users of the particular statistics under review. This consultation is a lengthy process. It may be relatively informal or it may be rather formal; in either case it is extensive. However, even the most formal elements in the process—meetings of interested parties—do not have the character of a "formal hearing," but are rather a mutual exploration of a matter in which all participants have a common interest. Views of the affected parties may change over time as proposals are tested, results evaluated and further proposals made before a final decision is taken.

As an example, the recent changes in the definition of employment and unemployment may be cited. The (President's Committee to Appraise Employment and Unemployment Statistics) known as the Gordon Committee recommended changes in the question-wording on the Current Population Survey to: (1) tighten up on the definition of the unemployed; and (2) add information about persons reporting that they neither worked nor looked for work. These recommendations were based on a widespread consultations with users. The hearings of the Subcommittee on Economic Statistics on Employment and Unemployment of December 18–20, 1961 were also available to it. The Gordon Committee specifically invited comments from academic specialists, business, labor, Federal Reserve Bank economists and others known to have an interest. In addition, all members of the Federal Statistics Users Conference were invited to comment.

Once the Gordon Committee report was received by the President, the Subcommittee on Economic Statistics arranged a hearing at which the recommendations were discussed by Gordon Committee members and heads of affected Government bureaus. I also asked the members of the Policy Committee on the Current Population Survey, an interagency committee which I chair, to advise me on a program for implementing the recommendations. As a result of these consultations a small interagency technical group was formed which developed a testing program. The results of the test were sent to the Gordon Committee members for comment, and received general approval. The testing program was described, and early results presented to statisticians at the 1965 meetings of the American Statistical Association. The Bureau of Labor Statistics reported on the results of these tests and described the proposed revisions to both its business and labor advisory groups.

Following a year of trying out the recommended questions in an "overlap" period, the final decision to incorporate the results in the

CPS was made during the summer of 1966. This decision, originating in the BLS, was considered by an interagency technical group, and by the Policy Committee on the Current Population Survey, augmented for the occasion by representatives of all major Federal agency users of the data. This was an important decision and required nearly four years after the publication of the Gordon Committee recommendations. During all this time and with all these formal and informal consultations, we received no comments or requests for hearings from any person or organization other than direct replies to requests for views.

One part of the final change in definitions, that to raise the lower limit of the official labor force count from 14 years to 16 years, had not been a part of the original Gordon Committee recommendations. This question had, however, been thoroughly discussed by the interagency groups mentioned, and had been recommended to the Gordon Committee by many of the persons that Committee consulted, particularly representatives of the business community.

I have described this experience at some length because I believe it is typical of the work that is undertaken when major changes in definition are reviewed.

Q. 4. "Have there been attempts to measure the burden to respondents in complying with requests for information? Is this burden equitably distributed?"

A. 4. Since 1964 the Office of Statistical Standards has routinely required that any proposal for a new or revised report shall include an estimate in manhours of the time required by respondents to complete the report. For repetitive reports this estimate is expressed in manhours per year. During 1965 similar estimates were entered into the records for the entire existing inventory of repetitive reports.

These estimates are used, first, in weighing the costs of a proposed report against its benefits, and also in statistical compilations showing the distribution of the reporting effort required by Federal reports by types of reports, types of respondents, and agencies. They lend thmselves to additional types of analysis which we have not yet had time or resources to develop.

These estimates of reporting time, by their very nature, can be no more than rough approximations and must be used with care. There are difficult conceptual problems in the measurement of reporting effort, and for any given report the time required (however measured) may vary widely from respondent to respondent or for the same respondent according to whether it is a new or a well established report. The total burden entailed by a report is as much a function of the volume of usage as it is of the inherent difficulty of the report. Even a very simple report may, by reason of frequency of filing by a large number of respondents, give rise to a large total "burden"; and total burden may vary widely from time to time (e.g., as in changing participation in a benefit program requiring applications and reports, or the changing volume of foreign trade requiring documentation). We have refrained from compounding the burden imposed by a report by requiring precise measurement of the burden itself. We ask the agency to estimate, with no more than limited consultation with respondents, the time required by a "typical" respondent. This is multiplied by the estimated number of responses to give a total.

Our analysis of these figures to date permits only very broad generalizations. They exclude tax reports since such reports are not subject

to review under the Federal Reports Act. The aggregate totals about 110 million manhours annually. As of the inventory of repetitive reports for December 1966, about half of the measured reporting burden (manhours) falls on individuals or households; somewhat over a fourth on business; about one sixth on State and local government units (including schools); and the remainder on miscellaneous groups including farmers and non-profit organizations. Equally significant is the distribution by nature of report. About 40 percent comes from application procedures whereby the respondent at his own initiative applies for some benefit or privilege. Another 50 percent of the burden arises from other reports incidental to the administration of government regulatory, benefits, procurement, or other action programs. Only about one tenth is due to what we have called "statistical" reports—those carried on for general informational or research purposes. It should be remembered, however, that data collected for administrative purpose often provide statistical byproducts.

These figures obviously do not provide meaningful information on the equity of the incidence of reporting burden on particular persons, companies, or other respondents. This is partly because so much of the burden is associated with voluntary application procedures, but more basically because the burdens associated with the several thousand repetitive or one-time reports are not in general cumulative on particular respondents. Particular reports tend to affect very specific populations which may or may not overlap—veterans applying for insurance coverage; tourists returning from Mexico; households drawn for a sample for a population and employment survey; companies with overseas investments; companies producing organic chemicals; etc. Useful conclusions as to the impact of Federal reporting on typical citizens, on business organizations of various attributes, and on other important respondent groups will depend on further analysis of our data which we plan to pursue.

"III. *The storage and retrieval of data.*

Q. 1. "Are there standards set down by the Office of Statistical Standards to insure that the micro data in various agencies are stored in machine readable form and can be utilized by researchers inside and outside of government, with due consideration for disclosure problems?"

A. 1. The Office of Statistical Standards has not formulated and promulgated specific standards governing the storage for research purposes of micro data. OSS does, however, participate actively in promoting such effort. In May of 1959 it published in *Statistical Reporter* general guidelines prepared by its American Statistical Association Advisory Committee on Statistical Policy after consultation with Federal statistical agencies. I append these guidelines. (Annex I)

Currently, an increasing volume of the major sources of information are being recorded in computer language for storage and use. In significant instances, tapes of micro data have been developed and made available for Government and private research, *under arrangements which prevent disclosure or identification of individual respondents.*

Key examples of tapes or micro data now available to researchers include:

a. The One in One Thousand Sample of returns from the 1960 Census of Population and Housing.

b. The BLS 1960–61 Consumer Expenditure Survey.

c. Statistics of Income Tax Model Tape Based on a Sample of Individual Income Tax Returns.

In addition to micro data, tapes of aggregate information are also available for the specialized uses of research persons inside and outside Government. These computer files include:

a. The tape record for the County and City Data Book.

b. Manpower and employment statistics data at the summary level including:

(1) National, State and area statistics from the current industry programs for hours, earnings, employment and labor turnover statistics, and

(2) Employment, unemployment and demographic information from the Current Population Survey, as well as other small survey programs in the manpower field.

c. Data tapes on the Federal Budget that can be purchased from the Bureau of Standards.

These piecemeal, though significant efforts, are indicative of a transition period toward increasing application of computer technology. It is clear, however, that we are as yet merely scratching the surface. The establishment of a Data Center, if it can be accomplished, would rationalize this beginning and assure better control of confidentiality while at the same time increasing the general availability for statistical uses. The Data Center through its experience in serving the needs of research would provide the logical basis for an effective formulation and implementation of standards for storage and retrieval of statistical information while protecting confidentiality.

Q. 2. "Are the statistical agencies making full use of modern technology in the processing and storing of data?"

A. 2. The application of rapidly developing modern technology by our statistical agencies has been outstanding in the collection and the processing of data. It has, however, lagged somewhat in the storage of data and making them available on a custom basis.

Technological advances in the Federal Government have not been completely dependent upon private developments. It should be noted that the Federal Government has pioneered in the invention and development of technology and methods for data collection and processing. Some of these include:

a. The institution of the first large scale computer—Univac I in the Census Bureau—long since replaced by computers of far greater capacity and speed.

b. The Film Optical Sensing Device For Input To Computers (FOSDIC) which permits the direct input of precoded data to computer tapes.

c. Procedures have been developed for using the computer in generating and addressing report forms "custom-tailored" to fit individual respondents.

d. Extensive development of probability sampling techniques particularly with regard to the Current Population Surveys and other current surveys. There is a danger in underestimating such

pioneering by the Federal Government if too much emphasis is given only to advances in hardware technology.

e. Major advances have been made in editing by use of computers.

It must be admitted in candor there have been gaps among agencies in this development and in particular a lag in the application of machine technology to storage and specialized access to our statistics in machine usable form. This includes macro as well as micro data.

In looking toward the establishment of a Statistical Data Center, and well before its initiation, we are in process of organizing work designed to develop file management procedures which would insure the storage and accessibility of the major elements of Federal statistics. The establishment of such a Center will provide the concrete setting for a more effective and dynamic application within and outside the Center of modern technology, including hardware and software, and applied to all phases of data development: its collection, storage and processing, by the agencies, and its creative analysis by users.

I am presently asking the major statistical agencies to prepare an up-to-date statement on their uses of electronic data processing equipment and will provide your Committee with copies of these statements when they become available. I do not believe it would be possible to provide them for this record.

"IV. *The advisability of a "National Data Center" or Statistical Servicing Center*".

Q. 1. "Can major progress in coordination and integration be made without a National Data Center?"

A. 1. While I would agree that there is always more than one way to achieve given purposes it is my considered judgment that at this juncture in the development of the Federal statistical program a Statistical Data Center organized along the lines I outlined in my prepared testimony is the most efficient and promising way of achieving valid statistics for comprehensive interrelated analysis which can be made economically available without sacrificing confidentiality or permitting the invasion of privacy directly or indirectly.

Q. 2. and 3. "Would a pilot project be advisable? If so, what magnitude of investment would be required for a meaningful pilot project?"

A. 2. and 3. I don't believe a meaningful pilot project is feasible—at least in any sense in which I understand the term "pilot." As I explained in my prepared testimony the scope of the Center can be restricted by the various devices I outlined. Its effective operation will, however, require time, continuity and general work with all types of data for statistical uses—at least in samples. An initial budget of about 2 million dollars per year—for the first two years would be my present minimum estimate to start a Statistical Data Center. Some costs, after the first two years, would be covered by charges for services rendered which should cover the direct costs of such services. Exclusive of hardware costs appropriation required after the first two years should be about $5 million per year for the next three years.

ANNEX I

STATEMENT OF PRINCIPLES ON AVAILABILITY OF STATISTICS

"Availability of Federal Statistical Materials to Nongovernmental Research Workers: A Statement of Principles" has been transmitted to the Office of Statistical Standards, Bureau of the Budget, by the Advisory Committee on Statistical Policy. This committee was established by the American Statistical Association in 1951, at the request of the Office of Statistical Standards, to advise on broad matters of public policy in the statistical area.

The new statement of principles was adopted by the committee in its meeting on March 19. It had been developed following consideration of the problems involved over a long period of time, and circulation of preliminary drafts for comments from Federal agencies concerned. In transmitting the statement, the committee recommended "that the Office of Statistical Standards continue its study of the problems of making Federal statistical materials available to nongovernmental workers and keep itself informed about the operations of the statistical agencies affecting these problems, in accordance with the provisions of the Statement of Principles, so that they may be reviewed from time to time and the Principles revised if necessary."

The text of the statement follows:

"AVAILIBILITY OF FEDERAL STATISTICAL METERIALS TO NONGOVERNMENTAL RESEARCH WORKERS: A STATEMENT OF PRINCIPLES

"I. Federal statistical and administrative data-collection programs often produce records capable of valuable statistical analysis beyond that which can or should be carried out by the collecting agency or any other agency of the Federal Government. In order that the optimum benefit may be obtained from Federal data-collection activities it should be the policy of the Federal Government to allow and to encourage the further analysis of these materials, under appropriate arrangements, by or on behalf of nongovernmental research workers.

"II. Although some general principles may be offered for the guidance of agencies with such statistical materials in making them available for further processing, each such agency must determine and assume responsibility for its policies and procedures in the light of the nature of its program and data and the demands for its data.

"III. No agency should enter into any arrangement for the supplementary processing of statistical materials, regardless of whether reimbursement is provided, which will interfere with the regular statistical program of the agency.

"IV. In general, requests for further analysis of Government data should be met as fully as possible by making special tabulations to the specifications of outside users.

"1. While a priority should be given to bona fide research uses in the general public interest, special tabulations should be permissible for all legitimate uses, both public and private, including, for example, marketing studies.

"2. The same rules for protecting confidentiality of individual responses must apply to special tabulations as are applied to the regular tabulation program.

"3. The agency should make only such special tabulations as appear to it to be justified in the light of the limitations of the data when the tabulations are to be available for general use or possible publication. Less exacting standards are permissible only when the data are not to be published but used for special analysis by competent analysts fully aware of the limitations.

"4. To the extent that special tabulations are deemed to serve a special, as distinguished from the general public, interest, the full costs shall be charged to the sponsors of the tabulation.

"5. Special tabulations should be in the public domain, available for publication by both the Federal agency and the outside sponsor, except as provided in IV, 3 above.

"V. Research needs which cannot be adequately served by special tabulations can under proper circumstances be met by allowing nongovernmental workers to work with the raw materials, work sheets and other intermediate materials within the agency.

"1. Such an arrangement is appropriate for research projects in the general public interest which would not be carried out without private sponsorship.

"2. The agency should take appropriate steps to insure that candidates for this privilege meet appropriate standards of competence and integrity.

"3. The agency should expect to instruct such workers as to the source, characteristics and limitations of the data and to cooperate with them, but it may properly set reasonable limits on the extent to which its own staff and facilities are committed to the project.

"4. While the agency should give to the results of the research such technical review, from the point of view of presentation and use of agency data, as it deems feasible and appropriate, the agency assumes no responsibility for these results. Any publication based on them should include a clear disclaimer to that effect.

"5. The agency should take whatever steps are necessary to protect the confidentiality of the data supplied by individual respondents, subject to the usual penalties for disclosure and other requirements of the agency law.

"VI. Under extraordinary circumstances an agency may make available to outside research workers copies of original data or intermediate materials which involve no disclosure of confidential data for further processing outside the agency. This arrangement is appropriate only for studies clearly in the public interest, too complex to be carried on under other arrangements, to be carried out by workers of known competence to make valid use of the materials, working in close cooperation with the agency staff.

"VII. Appropriate advance planning by agencies will promote maximum exploitation of data collected by the Federal Government by one or another of the above methods. Anticipation of demands for further tabulations beyond those planned for publication may be reflected in questionnaire design, card design and tabulation procedures. The relatively few surveys which lend themselves to duplication of original materials for use outside the agency can probably be identified in advance, so that planning may take this use into account. In general, any steps to make the survey procedure a matter of systematic record, intelligible to other competent research workers, will aid users to make valid use of the data."

Present members of the Advisory Committee on Statistical Policy are: Ralph J. Watkins (chairman), William G. Cochran, Gertrude Cox, E. Dana Durand, Walter Hoadley, Jr., Howard L. Jones, William R. Leonard, Rensis Likert, Isador Lubin, Frederick F. Stephan, William L. Thorp, and Samuel S. Wilks, William J. Carson of the National Bureau of Economic Research serves as secretary of the committee. (Raymond T. Bowman, Assistant Director for Statistical Standards, Bureau of the Budget)

THE COORDINATION AND INTEGRATION OF GOVERNMENT STATISTICAL PROGRAMS

THURSDAY JUNE 8, 1967

CONGRESS OF THE UNITED STATES,
SUBCOMMITTEE ON ECONOMIC STATISTICS OF THE
JOINT ECONOMIC COMMITTEE,
Washington, D.C.

The subcommittee met, pursuant to recess, at 10:07 a.m., in room 6226, New Senate Office Building, Hon. Herman R. Talmadge (chairman of the subcommittee) presiding.

Present: Senator Talmadge.

Also present: James W. Knowles, director of research, and George R. Iden, staff economist.

Chairman TALMADGE. The subcommittee will please come to order.

This morning the subcommittee begins its fourth and last scheduled day of hearings on the subject of the coordination and integration of Government statistical programs. The topic of this morning's hearing concerns the long-range possibilities and problems of our statistical system.

We are fortunate to have as witnesses, Mr. Arthur M. Okun, member of the Council of Economic Advisers; and Mr. Ewan Clague, formerly Commissioner of Labor Statistics.

Gentlemen, we appreciate your coming to contribute to our hearings.

Mr. Okun will discuss the statistical improvements for more effective public policies, and Mr. Clague the goals and difficulties of the Government statistical programs.

Mr. Okun, you may proceed as you wish.

STATEMENT OF HON. ARTHUR M. OKUN, MEMBER, COUNCIL OF ECONOMIC ADVISERS

Mr. OKUN. Thank you, Senator Talmadge.

As a member of the President's Council of Economic Advisers, I am an avid consumer of economic statistics. As you know, the Council is not a data-gathering agency. Rather it is our job to interpret and analyze economic data collected by other agencies and to advise the President on the implications and significance of the figures for public policy. The data producing agencies treat us like good customers. They recognize our wants and needs; they are responsive to our requests; they even cater to our idiosyncrasies. We get good service and we appreciate it. Yet, in the nature of the case, we never get all the numbers we want as accurately or as rapidly as we would like.

1. The Record of Progress

The Federal statistics program is good by any relative standard, and it is getting better all the time. Tremendous forward strides have been made in the 21 years since the Employment Act first set up the Joint Economic Committee and the Council of Economic Advisers. Over that period, there have been striking changes in the scope, quality, and availability of economic statistics. This is graphically illustrated in any comparison of the contents of today's *Economic Indicators* with the early issues. For example, quarterly national product data were not available on a deflated basis—that is, adjusted for price changes—until late 1958. Even many of our key annual series in the national accounts were developed only in recent years. The seasonally adjusted monthly unemployment rate was, in effect, first available in 1955; and the comprehensiveness and reliability of this key information has been increased several times—most recently at the start of this year.

Much of our progress stems from the decision, made cooperatively with the Joint Economic Committee, to use the national accounts as the conceptual framework within which to consider and design programs to improve our economic statistics.

Improved statistics have aided us immensely in the job of continuously assessing where the economy is and where it is going. It is hard now to imagine how my early predecessors on the Council performed their job without these essential pieces of information. Our knowledge about where the economy is today—based on prompt and reasonably accurate statistics—supplies a solid launching pad for our forecasts of future developments. In this respect, we are ahead of every other nation; in many, it is still necessary to "hindcast" the last few months before one can look ahead.

In addition to our improved knowledge about the recent performance of the economy, progress in statistical programs has also provided anticipatory survey data which bear directly on the future, reporting the expectations or plans of businessmen and households. Outstanding is the Commerce-SEC plant and equipment survey, which has removed much of the mystery from forecasting business investment expenditures over the near term and which has repeatedly and directly influenced policy.

2. Room for Further Improvement

Despite our record of progress, there remain many areas where policy decisions could be aided if our data were more comprehensive, more reliable, or more promptly available. I will not try to catalog the shortcomings of our statistics. But I should like to offer a few examples. They are particularly easy to find in the area of construction, which has been so crucial in the fluctuation of the economy over the past 2 years. We have only scraps of direct information on the rate of expenditures on additions to and alterations of residential structures—currently running at $4½ billion a year. We have no way of linking effectively the dollar volume of construction activity with employment in the industry or with unemployment rates of construction workers, since these series come from different sources with different coverage. We have no comprehensive Federal statistical series

on construction costs. We do not have meaningful data on the financing of construction projects. Finally, homebuilding is the one major private investment area which is not covered by a Federal anticipatory survey.

The shortcomings are equally abundant in the international area. Two gaps in our statistical knowledge I might mention, are up-to-date balance sheet information on U.S. investments abroad and adequate measures of the prices of our exports and imports.

Sometimes we run into problems with preliminary statistics designed to give quick results on the basis of partial samples. The most notable example of late was the marked overestimate in the advance report of retail sales for March.

Where we have problems of a serious nature, one generally finds there are no easy and costless solutions; if they were, they would have been applied. But I am convinced that we must keep working for solutions even when they are difficult and costly. Of course, outlays on our Federal statistics programs must pass the tests applied in careful and prudent management of Federal funds; the benefits of additional information must justify the costs. But there are potentially tremendous benefits at stake. Literally, billions of dollars can hang on Federal economic policy decisions. And there are times when the expenditure of some thousands of dollars for added statistical information could greatly improve the efficiency of such decisions.

Quite apart from the needs to fill gaps and to firm up weak spots, our statistical programs need financial support for development and experimentation. Our hard-pressed data collectors need the time and the resources to take stock of the opportunities and to explore the more promising ones. Since the great potentialities of anticipatory data have been demonstrated in the case of the plant and equipment survey, there have been numerous suggestions for added surveys of anticipations. Some which have not yet been implemented include plans for new contractual saving, plans for homebuilding, travel and leisure plans of consumers, production and employment plans of manufacturing industries, and expenditure plans of State and local governments. In some of these cases, the feasibility and value of such surveys can be determined only by giving them a try on a small scale. Such experiments of this sort deserve particular support and encouragement. To mention one successful case in point, Census now conducts a valuable quarterly survey of consumer intentions to buy durable goods; the survey was initially developed in an experimental program financed by the Federal Reserve System.

3. Our Rising Aspirations

Developments in economic policy and improvements in the performance of the American economy have enlarged the demands for statistical information. At one time, the economic policymaker was essentially a fireman, standing by much of the time until the alarm sounded the onset of recession or inflationary boom. Now, however, policymaking is clearly a continuous matter, aimed to help promote steady growth and noninflationary prosperity all the time. An information system could be adequate in sounding the alarm to herald major disruptions and still fall far short of meeting the needs of our current policy strategy.

Moreover, a fully prosperous economy generates needs for more detailed information to guide policy. In a slack economy, the major assignments of policy are clear. Such indicators as the overall unemployment rate and the rate of capacity utilization are good guides to the fulfillment of those tasks of expanding the economy. However, with the unemployment rate below 4 percent today, the desirable pace of growth and expansion must be gaged more carefuly. We need to know the distribution of unemployment and job vacancies among regions, industries, and occupational groups; we need to know the operating rates of capital in many industries. For these reasons, the development of adequate data on capacity (in both manufacturing and nonmanufacturing industry) and on job vacancies should stand high on our priority list of needed guides for stabilization policy.

A full employment economy also brings to the fore the interrelationship between monetary and fiscal policy. It increases the need for detailed information on the relation between financial flows and income-expenditure flows. This puts special emphasis on accurate and prompt flow-of-funds information that is integrated with the national income and product accounts.

My examples are drawn from data bearing on stabilization policy and overall economic developments, because these are the areas where I personally most often wrestle with the data. But I do not wish to ignore the importance of statistical information in other economic policy decisions of the Federal Government. For example, the growing concern with the plight of the poor in recent years has created a need for more detailed information on income distribution, the social characteristics of the hard-core unemployed, the mobility of labor, the economic benefits of general and vocational education, the relation between medical care and income, and so forth.

4. The Partnership of Data and Research

There is a virtuous circle between the development of economic statistics and improvement in analytical economic knowledge. The statistics producer gives the empirical analyst new opportunities; his results aid the policymaker; and the policymaker's needs stimulates new efforts by the data producer. For example, the development of the national income accounts permitted empirical testing of many economic hypotheses concerning consumption and investment behavior. These findings, in turn, suggested ways in which the accounts could be improved and supplemented with additional information.

Often, even generally, the economic analyst in this sequence has been outside the Government. It is highly desirable that the bulk of our basic economic research should be conducted in our universities and private research foundations, rather than in Federal agencies. Federal economic statistics should be grist for the research mill of the private sector; and the output of this mill should continue to enhance the understanding and capability of the policymaker.

In this way, improvements in policymaking over the longer run depend heavily on the volume of Federal statistical information that is made available to private researchers and on the efficiency with which that information is disseminated. There are great opportunities for constructive partnerships whereby we in the Government can help academic researchers by making data readily available, and whereby

they can help us by providing new insights into economic behavior gleaned from these data.

The availability of Federal statistical information may also encourage the academic community to focus its efforts more effectively on the pressing problems of the real world. If parts of the academic community sometimes seem surprisingly detached from the key issues of modern society, one of the reasons may be that some academicians despair of their ability to get the information needed to grapple effectively with these problems. It has been said that a social scientist is a man who searches under a bright street lamp for a wallet that he knows he lost in a dark alley. Whether or not there is any validity to this charge, we should be working to erect bright lights in dark alleys.

The establishment of a Federal statistical center is one promising way to improve the accessibility and availability of data to interested private researchers. It can increase the efficiency of information storage and retrieval and can take full advantage of the latest technological advances involved in high-speed data processing equipment.

As the Kaysen committee has indicated, no expansion of published data in printed form can provide the flexibility or comprehensiveness that can be obtained by proper storage and retrieval of the detailed original data collected in surveys and statistical programs. Many key questions about economic behavior can be answered only by exhaustive statistical investigation of samples drawn from large bodies of economic data reporting on the characteristics and behavior of individual units.

Let me go into one example. We know that the volume of business investment is related to levels of utilization of capital equipment. Obviously, a firm with much excess capacity has less incentive to invest than one that is making full use of its capacity. We also know that business investment is encouraged by high levels of corporate profits, which raise the prospective yields on capital projects and which ease the problem of financing investment.

When we look at aggregate statistics or even industry statistics, however, we find that high utilization and high profit rates go together very closely. Thus, it is hard to sort out the relative influence of utilization rates and profit rates on investment. Yet, some of our key policy choices in income taxation can depend on the relative role of profits and utilization as determinants of investment. We can hope to get answers only by examining large numbers of firms and relating their investment, utilization, and profits by sophisticated statistical techniques that can sort out the various influences. In a sufficiently large sample, some firms with identical utilization rates will have markedly different profit rates, and some firms with identical profit rates will experience different utilization rates. Considerable research of this kind has already been carried out. But more remains to be done. And it could be done more promptly and more definitively if private economic researchers had greater access to the data of our Government agencies.

This is only one of the many conflicts in hypotheses about economic relationships that can never be adequately settled on the battlefield of aggregative data. Detailed cross-section analysis is the necessary route to truth in many cases. And if economists are to have detailed cross section data on individuals or firms, the problem of confidentiality arises. This is a real and urgent problem. Judging from recent discussions of it, I am confident that it is not an intractable problem

and that it can be resolved without compromising the fundamental right of privacy.

Improved efficiency in the storage and retrieval of data is important to consumers of statistics within the Government as well as to academic users. When the Council needs special tabulations or disaggregations or statistical analyses, the data agencies are happy to cooperate. But we know the requests of the Council can impose a burden, diverting time and effort from other tasks. In particular, the statistical agencies are often not geared up to meet particular needs without significant costs and disruption. If there was a center whose specific function was to meet such needs, I think we at the Council would request additional information more often, and I believe that our national arsenal of economic knowledge would be the better as a result.

I also believe that the comprehensive effort to arrange and codify available Federal statistics for the purpose of efficient storage and retrieval would uncover some duplication and some gaps of which we are not now fully aware. It might bring some data problems to the forefront where they would attract more attention and encourage more coordinated joint effort. For example, our key figures on employment, sales, profits, and investment in manufacturing come from four different sources relying on different samples. This is not a basic defect, but neither is it ideal.

In assembling all our information on the manufacturing sector, we might well strike upon promising opportunities for consolidating data collection in this area.

The possibilities for strengthening the partnership between our existing data and economic research represents one important way in which Federal statistics programs can continue to contribute to the development of more effective public policy. Both our statistics programs and our policymaking machinery are dynamic systems which can continue to improve together and to reinforce each other in the years ahead.

Thank you.

Chairman TALMADGE. Thank you very much for a very fine statement, Mr. Okun.

Would you favor the initiation of a pilot servicing data center?

Mr. OKUN. I believe, in principle, that the Federal Statistical Center has much to offer, and the initiation of it on a pilot basis could be a good way to get started. It could give us more information about how the center could be operated, what the demands for its services were, what we learn when we put together our data in a comprehensive fashion for improved storage and retrieval. I think it is important, at the same time, to recognize that doing this on a piecemeal basis—such as a pilot study suggests—would not yield the same benefits that we would get by adopting it on a comprehensive basis right at the outset. To some extent, the notion of comprehensiveness—of trying to put all the data together and assembling it in systematic fashion—is an essential feature of the Kaysen Committee's suggestion. If we do launch the center on a pilot basis, we really should bear in mind that we are working toward comprehensiveness, unification, a single system.

Chairman TALMADGE. Would a national data center, as in the Kaysen report, be of material help to the Council?

Mr. OKUN. In the short run, it would establish someone whose job was to offer statistical services. This is a limitation that we find today

in making requests to other agencies for the kind of data we want. We know we can get what we want from them, we know they will try to help us, but we know we are asking them to try to do something which is a diversion from their main task and a new and added burden on them. If we had a statistical service center or a national data center, we would have a group which is enthusiastically ready, able, and willing to meet our needs. That would be their job, and they would welcome our requests as part of their main activity. In the longer run, the important contribution that the data center could make to the Council's efforts is by strengthening this partnership between the private researchers and the needs of policy, as I indicated in my statement. A great deal of the economic knowledge that is relevant to policymaking has to come from universities and such private foundations as the National Bureau of Economic Research and The Brookings Institution. The more data they have and the more we can interest them in doing research on matters that are relevant to economic policy, the better our economic policy is going to be.

Chairman TALMADGE. Would it help to develop an evaluation of specific Government programs?

Mr. OKUN. Would the national data center help?

Chairman TALMADGE. Yes.

Mr. OKUN. The evaluation of Federal expenditure programs is, as you know, largely the work of the individual agencies that operate those programs and of the Budget Bureau. I do not really feel fully qualified to say what data are needed for the evaluation of Federal expenditure programs. Obviously, in most cases, the programs themselves need to generate the data—and their own operating management data—which tell them how they are doing and how well they are operating. Particularly in guiding the development of new programs, however, we would benefit from information that could be available in the Federal data center.

Chairman TALMADGE. Which statistical improvements should have the highest priority?

Mr. OKUN. It is very difficult to award the top prizes for high priority in statistical needs. For one thing, the costs of various possible improvements may vary greatly. It might be highly desirable to have a particular set of statistics, and yet, if getting it accomplished were so expensive that it foreclosed many other opportunities, we would have to move it down on the priority list. I spelled out several areas of need in a letter that was published by the Joint Economic Committee in "Improved Statistics for Economic Growth," March 1966 (pp. 79–81).

I noted earlier the continuing statistical gaps in the construction field. In the area of labor markets, we sorely need data on fringe benefits and on unfilled job vacancies. Balance sheet information on investments abroad and new statistics on export and import prices deserve high priority in the international field. I also consider it essential to expand and improve our surveys of business and consumer plans. Finally, I might mention that the need for better information on inventories has been underlined by the critical role of that sector in the current economic situation.

Chairman TALMADGE. Thank you very much. We appreciate your appearing with us this morning, and your testimony will be very valuable in our deliberations.

Mr. OKUN. Thank you. It was my pleasure.

Chairman TALMADGE. Mr. Clague, you may proceed as you see fit, please.

TESTIMONY OF EWAN CLAGUE, FORMER COMMISSIONER OF LABOR STATISTICS

Mr. CLAGUE. Mr. Chairman, I have a written statement which I would like to submit for the record and then, if you please, I would like to talk extemporaneously in a brief summary of it.

Chairman TALMADGE. Without objection, the statement will be inserted in the record, and you may proceed as you see fit, sir.

(Mr. Clague's prepared statement follows:)

PREPARED STATEMENT OF EWAN CLAGUE

THE COORDINATION AND INTEGRATION OF GOVERNMENT STATISTICAL PROGRAMS

Mr. Chairman, I am happy to respond to the invitation to discuss the questions you and your Subcommittee have posed on the general subject of Federal Government statistics. At the outset, I want to emphasize that I am speaking as an individual and not as a representative of any government agency. I retired from the Federal service in December 1965 after nearly 35 years of government service.

I. COORDINATION OF STATISTICAL PROGRAMS

There is in existence a system of coordination of statistics for programs within the Federal Government, a system which has been operating more than 30 years. This coordinating system resulted from the work of the Committee on Government Statistics and Information Services in 1933–34. That Committee was established by the American Statistical Association at the request of the Government, which asked the Committee to conduct a comprehensive survey of Federal statistics and to make recommendations for their improvement. The Committee recommended a decentralized system, with the basic statistical agencies being located in the respective departments—Agriculture, Commerce, Labor, etc. However, recognizing the need for coordination, the Committee further recommended the establishment of a Central Statistical Board, with a budget and a staff, whose function was to establish and maintain coordination among the agencies. Some years later this Central Statistical Board became the Office of Statistical Standards of the Bureau of the Budget, where it is still functioning today.

This system of cooperation has worked well in a number of cases where good cooperative relationships have been established among Federal agencies. In 1959, by agreement of the Secretaries of Commerce and Labor, a cooperative arrangement was worked out between the Bureau of Labor Statistics and the Bureau of the Census for the production of the overall statistics of labor force, employment and unemployment. The Bureau of Labor Statistics obtains the funds, while the field surveys and the tabulations are conducted by the Bureau of the Census on a contractual arrangement. The Bureau of Labor Statistics is responsible for the analysis and publication of the results. This is a highly efficient and economical operation for the production of the monthly labor force statistics.

Furthermore, by using the Census household survey samples, it is possible to develop special labor force and manpower reports linked to the continuing monthly labor force series. Those special reports are on such subjects as income of the unemployed, work experience of the labor force in a given year, women's participation in the labor force, unemployment of youth, etc.

This cooperative arrangement has proved itself to be so economical and efficient that other agencies adopted it. The Public Health Service has developed a partnership arrangement with the Census for sample surveys of the health of the population. More recently, the Office of Economic Opportunity has tied in to the BLS-Census program for some of their research and statistical studies.

I cite this as one example of the effectiveness of coordination where cooperation and good will among the agencies can be established and maintained. I believe that this pattern could be more widely applied in the Federal service.

When the statistics involve State and local agencies, the problem becomes much more difficult. Many operating programs conducted by State and local agencies are certain to have diversities and variations which make the collection of nationwide uniform statistics quite complicated. Furthermore, State and local statistical departments tend to devote more attention to their own needs for data and less attention to the national coordinator.

Yet even with these handicaps, I can cite an example of effective coordination in statistics. I refer to the Federal-State program of employment, hours and earnings statistics operated by the Bureau of Labor Statistics, the Bureau of Employment Security, the State Employment Security agencies, and some State Departments of Labor.

To several State Departments of Labor goes the credit for initiating, prior to World War I, a program of collecting from employers monthly reports on employment and payrolls. During World War I the Bureau of Labor Statistics started a nationwide reporting system along the same line. Before long a cooperative arrangement was worked out whereby the States became the collecting agencies for the Bureau of Labor Statistics in their respective jurisdictions, thereby avoiding duplicate collections of the same data.

When the Employment Security program began operating in the late 1930's, it became possible to link up this sampling system with the comprehensive reports on employment required by the Bureau of Employment Security. Under the present arrangement, the Bureau of Labor Statistics makes contractual arrangements with the cooperating State agencies to transmit to Washington the sample monthly reports which constitute the national data issued by BLS every month. Meantime, the State agencies use the same reports for State and local area data within each State.

This joint statistical program has derived its strength from two significant advantages. One is that the original Social Security legislation contained a strong section on requirements and standards of reporting by the States. When the Employment Security program was transferred to the Department of Labor, the Bureau of Employment Security continued to have authority to establish nationwide uniform requirements for State and local office reporting. A second advantage is the fact that the Bureau of Employment Security is responsible for distributing one hundred per cent grants for the administration of the State agencies, which means that the Bureau can provide to the States the funds necessary to obtain the statistics and to report them to Washington.

This statistical system is a truly remarkable example of economy and efficiency in the face of great administrative diversity. A single establishment report by an employer provides monthly data on employment, hours of work and earnings which, in combination with others, yields statistics for the economy as a whole, for 51 States and for perhaps 200 local areas, with detail for several hundred industries.

Where such favorable administrative and budgetary arrangements do not exist, it is my impression that Federal-State-local statistical programs do not work so well. Yet, this is a problem which must be solved. So long as States and localities are partners with the Federal government in the administration of joint operating programs, it will be necessary to establish joint statistical programs covering the same agencies. A continuing effort must be made to improve Federal-State statistical programs.

There is a third set of organizations with which Federal statistical agencies conduct research and statistical work. Those are the private contractors, such as foundations, universities, or private firms, which contract for the performance of certain research services.

There are many advantages to the use of such private research agencies for the making of special one-time studies or the pursuit of basic reesarch in a specific field. Such agencies have better facilities for recruiting local staff; they can tap the experience of experts who are not in the government; and the project can be ended when the report is completed. On the other hand, my general impression is that such private contracts are often more costly than if they were done by the government agency itself.

However, I am convinced that continuing statistical series for public use must be conducted by government agencies. In any basic series, such as the Consumer Price Index, it is absolutely essential to maintain continuity and stability. This can best be assured in an established and reputable government agency.

Also, I believe that more care is needed in order to distinguish between these projects which can appropriately and effectively be done by private organizations and those which should be conducted by the government agencies themselves.

In conclusion, on the subject of coordination, I believe that more support and more resources should be provided to the Office of Statistical Standards of the Bureau of the Budget. We have the machinery for coordination; what is needed is to strengthen it.

II. NATIONAL DATA CENTER

When the Committee on Government Statistics and Information Services reviewed the Federal programs in 1933–34 (as I mentioned above), they arrived at the conclusion that there should not be a single, government-wide statistical agency in the Federal Government. They voted for coordination instead. In my judgment that conclusion was sound.

The proposal for a National Data Center is not, as I understand it, a proposal for a single statistical agency, but rather an extension of the present coordinating system. As such, I would support that proposal as an ultimate goal. However, I want to point out that there are many complications to be overcome before such a center can achieve what seems to be expected of it.

The issue of privacy has already been raised. Many of the statistics collected by Federal agencies are obtained on a voluntary basis, accompanied by a pledge of confidentiality. A business firm or individual may willingly report to a statistical agency for public purposes when he would refuse to report to a law-enforcement agency, which might use the information against him. There are many degrees of concern with respect to confidentiality. Many people are sensitive about public knowledge of their incomes. We have all noted recently in the papers the objection of an individual to having his name and address turned over to mailing lists. Of course, many individuals would object to having their names and addresses made available to bill-collecting agencies. I recall in the Bureau of Labor Statistics that sometimes business firms were insistent on making sure that the fact of their participation in a BLS survey was not made known to the industry or to the public generally. That is one of the reasons the BLS has always been most careful to maintain the confidentiality of the names of its sample reporting firms.

I must also express a note of caution against too much optimism as to the usefulness of raw data to the prospective users. The Bureau of Labor Statistics has had experience which has a bearing on this question. When we tabulated the results of the 1960–61 Family Expenditure Surveys, which the Bureau used in the revision of the Consumer Price Index, we put the results on computer tapes and offered to make these available to both government and private agencies. What we found was that the full time of one or more BLS staff members was absolutely essential to service outside agencies in their utilization of the data. Any researcher attempting to use data collected and processed by another agency is bound to have scores of questions arise in the course of his studies. The only available sources are those who conducted the original study. It just isn't possible for an outsider without help to make the most effective use of raw data unsupported by experienced and informed interpretation.

I would recommend that, if and when steps are taken to develop any National Data Center, preliminary action must be taken first, to establish safeguards which insure the confidentiality of the individual records; and second, to provide for the financing of a servicing system by the statistical agencies contributing data to the Center.

With respect to the initial character of the Center, I like the suggestion made by Professor Fred Stephan that the experiment begin on a limited basis, perhaps with the setting up of "A national statistical index and library to serve users as the indexes of medical and legal literature serve their users."

III. ADMINISTRATION OF STATISTICAL PROGRAMS

One of the most trying experiences of a director of research and statistics in a Federal agency is the obtaining of the necessary funds for (a) the maintenance of adequate safeguards for continuing statistical series; and (b) the progressive improvement of those series, that is to say, improvement corresponding to the expanding uses. Any continuing series will deteriorate with time, unless new samples are established, new problems analyzed, and new techniques explored. Tests have to be made from time to time to discover weaknesses.

Furthermore, if the statistics acquire policy and operating uses, the users have a tendency to stretch the applications almost to the breaking point. It is hard to keep users fully aware of the fact that a statistical figure which is adequate for the labor force as a whole is subject to a wide margin of error when it is applied to a relatively small group.

When public criticism becomes widespread, as in the case of the unemployment figures five or six years ago, then there is plenty of action. President Kennedy appointed a Committee to Appraise Employment and Unemployment Statistics. Your Joint Economic Committee held a series of hearings at which the Bureau of Labor Statistics was able to make public its answer to the criticisms. The President's Committee studied the problem for a year and came out with a comprehensive series of recommendations for revision and improvement of the data. Congress then acted upon the recommendations, one of which resulted in the establishment of a special sample of households, which was used for testing and checking over a period of several years. The revised and improved labor force, employment and unemployment statistics of 1967 are the product of that work.

What I wish to emphasize is that the statistical agencies themselves are fully aware of the shortcomings of their various statistical series. They can propose tests designed to insure that quality standards are maintained and that needed improvements are made from time to time. I hope that your Subcommittee will continue to provide, as you have done in the past, a forum for the directors of statistical agencies to present their developing problems, so that positive remedial action can be taken in advance of any general public criticism.

A second problem of administration concerns analysis and interpretation of data by the producing agency. In my experience, it has been difficult to obtain and retain funds for such analysis. The importance of this function has not always been fully appreciated, although effective analysis and interpretation is often the best way to insure public understanding of the figures as well as to guard against misinterpretation. Even when funds and staff are obtained for such a unit, it often is weakened by pressure of current operating needs. In case of staff shortage, the requirements of the current statistics for accuracy and timeliness will always get priority.

A third example is the difficulty of maintaining a staff engaged in exploratory research in statistical methods. It is hard to measure the accomplishments of such a unit, since its findings may not show up in immediate visible and tangible results. In addition, it too is subject to the pressure of current operating needs. I know by experience the difficult choice an administrator must make when the longer future comes up against the urgent present.

It is difficult for your Committee to deal directly with such problems, since they arise in part through day-to-day administration of statistics. However, it is important that adequate funds for analysis and research be provided, and that these functions should be performed in the first instance by the statistical agency.

IV. LONG-RANGE PLANNING

With respect to long-range statistical programs, I want to mention one perennial problem to which your Committee has given attention in the past, but which deserves reiteration. I refer to those general purpose statistics which are basic to the Federal statistical system but which have no generally recognized specific public uses.

As an example, I can cite the contrast between the Wholesale and the Consumer Price Indexes of the Bureau of Labor Statistics. The Consumer Price Index is known throughout the country as a measure of the cost of living and as a vital tool for collective bargaining and wage escalation. There are labor and management users ready to support funds for that index.

Now, as a matter of fact, the Wholesale Price Index, is also widely used in escalation contracts between business firms, or between Government agencies and business firms, but there is very little visibility to this use. However, the Wholesale Price Index (or rather, its constituent sub-indexes) is absolutely essential to the Gross National Product and other National Accounts. Adequate price indexes for individual industries are required in order to convert dollar product into real product, and thus determine the rate of economic growth. But this is a use which cannot be widely understood by the general public. It requires an expert committee such as yours to understand fully the contributions of certain statistical series to the Federal system as a whole.

Let me emphasize also that the productivity indexes for the economy as a whole and for the major sectors are derived from output data adjusted for

price changes. So the productivity indexes which are being so widely used today are dependent upon adequate price statistics.

I have selected this example of price statistics because your Committee is aware of the problem. Five years ago, and again last year, your Committee held hearings on price statistics, in the first case with specific reference to their significance for the National Accounts.

A look at the future also requires us to take notice of new programs, such as manpower and poverty, which will require statistics for administrative, program and policy purposes. The operating agencies will inevitably center their attention upon the short-range research and statistics which they need to enable them to operate successfully. Your Committee will need to be concerned about the basic continuing statistics which will be needed to measure progress in the achievement of long-range results.

Once more, let me cite an example from my own experience. Back in the days of World War II the House Appropriations Committee instructed the Bureau of Labor Statistics to prepare a standard family budget which would show in dollars (rather than indexes) how much it costs to live. In response, the Bureau developed and published in 1946 a four-person family budget designed to provide a modest but adequate level of living in the larger cities of the United States. For lack of adequate upkeep, this budget got out of date and was dropped from the Bureau's program after 1951, except for an interim revision in 1959. Now, from the data obtained in the family expenditure surveys of 1960–61, the BLS is soon going to be producing a series of family budgets—for several different types of families and for three levels of living conditions.

It requires no great insight to see that such data should be absolutely basic to policies for social welfare and economic well-being of the American people. Yet, changing conditions will make these budgets obsolete soon after they are published, unless we find some way to keep them up to date—as we have failed to do heretofore.

On this point, I want to make a specific suggestion. Back in the middle 1950's the BLS advanced a proposal for an annual survey of family expenditures in a small sample of families so that the results could be used (a) to check when the Consumer Price Index would be in need of revision, and (b) to keep the family budgets up to date. The proposal never caught on.

There are a number of purposes which such annual surveys could serve. They would furnish a measure of consumer behavior, and consumers constitute the largest economic factor in our economy—they purchase nearly two-thirds of the Gross National Product. Such surveys would also furnish valuable data for the household sector of the National Accounts and would make an important contribution to the analysis of economic growth. They might also make it possible to devise new measures of quality adjustments in the Consumer Price Index.

There is one last area to which I wish to call attention—international statistics. International trade, international finance and international agreements are expanding all over the globe. A hundred nations are entering international markets, many of them for the first time on any significant scale.

Yet, in the United States this is one of the weakest areas in our statistical program. As has been pointed out many times, we do not have adequate statistics of import prices, export prices, wages and fringe benefits, productivity, unit labor costs and many other series which are necessary for an apraisal of our international position. Of course, this is an especially difficult field, since some of the data require the active cooperation of other nations. Furthermore, the technical problems of comparability are formidable.

But these problems are not insoluble. Much information is available and more can be obtained. A good deal can be done in this country by ourselves. Is there any way in which we can stimulate the development of an adequate program of international statistics?

Mr. CLAGUE. First of all, I am speaking as an individual; I do not represent any particular Government agency. I retired from the Federal service in December 1965, after nearly 35 years in the service.

Mr. Okun has just spoken to you as a user of statistics and I represent, in a sense, the producers of statistics, the statisticians who produce the data. In my paper I have used a good many examples, practically all drawn from the Bureau of Labor Statistics. I want to emphasize to you that I regard these as being representative of the problems of

other agencies. I am talking about things I know about, but I believe that in every instance I am citing examples that could have much wider application in the Federal statistical service.

I have grouped my comments under four major headings or topics. One is, "Problems of Coordination," in which I describe the machinery for coordination, another is "The National Data Center." Then I want to say a few words on the "Administration of Statistics in the Federal Service," and finally, my fourth topic is "Long-range Planning," looking into the future.

With respect to the topic of coordination, we do have an operating system in this Government which was established by the Committee on Government Statistics and Information Services back in 1933–34. That is when the Government asked the American Statistical Association to establish a committee, an overall commitee, which would make a thoroughgoing review of Federal statistical work. I think that the revolution in statistics which has occurred since that time stems from the work of that committee.

They favored, not a central statistical agency, but a federal system of statistics, so to speak, a coordinated system, and they set in motion the machinery to bring that about; namely, a Central Statistical Board which existed for a number of years. It later became the Office of Statistical Standards in the Bureau of the Budget, where it still operates. You heard from Mr. Bowman yesterday, the Director of that office. I want to emphasize that that system can work very well, and I have cited in my paper some illustrations of it.

Example No. 1 is the joint labor force statistics produced by the Bureau of Labor Statistics and the Census which has been a very valuable policymaking tool for the last 15 to 20 years. It has been described by Mr. Bowman, so I shall not go into it in detail, but I do want to emphasize that one of the advantages of that kind of a continuing statistical system is that you can tie in special studies. The Bureau of Labor Statistics issues special labor force reports, based on special samples, which can be related to the overall system of reporting in such a way that we get maximum utilization of the data. Here are some examples: "Work Experience of the Population in 1965, Poverty Areas of Our Major Cities; Adult Men not in the Labor Force." These numbers have already reached about number 80. This is an efficient way of combining special studies with regular reports.

Similar arrangements have been made by the Public Health Service and the Census. The Bureau of Labor Statistics published recently a summary of "Work Limitations and Chronic Health Problems," which was a joint Public Health Service and Census report. More recently, the Office of Economic Opportunity has joined in putting some of its studies into this general statistical system.

The coordination of State and local data is a much more complex problem. That is because State and local agencies have special needs of their own. It means that there is a tough problem of trying to get the nationwide coordination needed for national statistics in the midst of State and local diversity. Nevertheless, I can cite another example of success in this field. It was over 50 years ago, prior to World War I, that some of the State labor departments began collecting information on employment and payrolls for employers. During World War I the Bureau of Labor Statistics started collecting these data on a national basis. These two groups joined together and we had the Bureau

of Labor Statistics and the State labor departments cooperating in collecting employment and payroll statistics. When the social security program came into operation, particularly employment security, so had widespread unemployment insurance coverage of employers, so that program was brought into the partnership by providing the benchmarks, the basic counts of employees which we could use to adjust the current monthly reports of employment and payrolls. At the present time we have the Bureau of Labor Statistics, the Bureau of Employment Security, the State employment security agencies and some half dozen State departments of labor that cooperate in Federal, State, and local statistics.

My impression is that in a number of other programs there has not been such a successful operation of Federal, State, and local statistics. I think this example might be more widely applied in other programs.

The third group is the private agencies—foundations, universities, business firms—who cooperate from time to time under contract or by grants in research and statistics. There are both advantages and disadvantages to the use of private agencies, as I have mentioned in my paper. I want to make two points in that connection. I firmly believe that the continuing statistical series should be produced by a Government agency. I do not see, for instance, how the Consumer Price Index could be produced by a private agency and have the standing which it does. Therefore, I want to urge that adequate funds be provided for Government agencies to produce the continuing statistics. My second point would be that there needs to be careful selection of projects, the special projects or special surveys, which should be done by the Government agencies themselves and which by private contracts.

In conclusion on this whole subject of coordination I want to say that I think we do have the necessary machinery. I think the Office of Statistics Standards in the Bureau of the Budget should be strengthened by support and resources to enable them to do the job that they are responsible for doing.

With respect to a National Data Center I understand that this is not a proposal for a single central statistical agency. The Committee on Government Statistics over 30 years ago recommended against that and I think their recommendation was a sound one. However, I do support the National Data Center as an ultimate goal. I want only to present certain cautions and conditions.

No. 1, is the maintenance of confidentiality. The Bureau of Labor Statistics through the years has operated on a strictly voluntary basis of cooperation from employers and respondents of all kinds. I cannot, in the course of my two decades or so as head of the Bureau of Labor Statistics, recall more than three or four or five studies that failed for lack of adequate voluntary cooperation. This voluntary cooperation is dependent upon confidentially, so my first caution is to make sure that any National Data Center has adequate protection for the confidentiality of the individual records of persons as well as of business firms.

The second point I would like to make is that in the operation of any center of this kind the supplying agency—that is, the statistical agency supplying the information—will need the funds and staff to provide adequate answers to the questions that arise. We had an experience of this sort in the Bureau of Labor Statistics in connection with our family expenditure surveys in 1960–61 when we collected the

family data for about 15,000 families. Many Government agencies wanted the use of these data, many private agencies also. We produced tapes in the Bureau of Labor Statistics which we made available for a price to these other agencies. What we found is that we needed to supply also both staff and machine time in order to answer the questions that inevitably arose from the users trying to use data with which they were not directly familiar.

So, I would emphasize to your committee that the setting up of a National Data Center should include the supplying of either funds or staff to the contributing agencies who will be supplying their data so that the users can be properly serviced by the Center.

My third topic relates to the administration of statistics. In my long experience in trying to get funds for statistics in the Federal Government, I have run up against three perennial problems.

No. 1 is the maintenance of standards in the continuing series, such as the Consumer Price Index, the employment statistics, and many others. Over a period of time, months, and years, a series will inevitably decay, accumulate weaknesses as it goes along, unless it is continually reviewed, tested, and improved. One of the difficulties is that this deterioration may proceed for some time until it breaks out in the form of public criticism. Then, indeed, we get action, but we are picking up the pieces from what might have been prevented.

You may recall that 5 years ago there was a vigorous attack on the unemployment statistics. Your Joint Economic Committee promptly held hearings, at which the Bureau of Labor Statistics was able to answer these criticisms in public. President Kennedy appointed a Committee to Appraise Employment and Unemployment Statistics. The report of this Committee was mentioned by Mr. Bowman yesterday. Congress acted upon the Committee's recommendations, at least to the extent of providing the Bureau of Labor Statistics and the Census with a special test sample, so that we could test alternative methods of getting employment and unemployment statistics. And now, the new revised employment and unemployment statistics of 1967 are the fruit of that activity.

However, my main point is that the statisticians in the Federal Service can document for your committee their weaknesses and limitations and their growing difficulties. I hope that you will continue to provide a forum in this committee for the statisticians to express their judgments on this point, so that positive action can be taken before something breaks open in public criticism.

That concerns the maintenance of standards. My second point under the administration of statistics concerns analysis and interpretation. And here I believe strongly that, in the first instance, adequate funds for analysis and interpretation should be in the hands of the producing agency itself. They must have the necessary analytical staff, so that they can deal with the problem of public presentation and public understanding of the figures and, therefore, be able to take care of developing criticism on the outside which arises from misunderstanding. This does not mean shutting out other agencies in the Government or outside the Government from doing further analytical work on any continuing series. I simply stress the importance of getting adequate funds for the producing agency itself. I know from experience the way in which operating priorities can push aside these analytical jobs. We had to get our Consumer Price Index on time. That is so important that

80-826 O—67——10

other work is going to take second place when we are up against any limitation of funds or staff.

The third area of administration which I will mention briefly is each statistical agency ought to have a small unit for methodological research on statistical methods. We ought to be doing more exploratory work in the field of statistics production. And again, this is another which often succumbs to pressures to get the figures out.

My last general topic relates to long-range plans. Here I would like to call attention to several fundamental problems. A basic problem which is widely recognized by your Joint Economic Committee and on which you have held hearings over the years, relates to the general purpose statistics which do not seem to serve any specific, immediate, short run use. I think that this point warrants reiteration. So, in spite of your knowledge of the subject, I am going to mention it again.

An illustration from my own experience is the Consumer Price Index as compared to the Wholesale Price Index. The Consumer Price Index is used for wage escalation and in all sorts of contracts throughout our economy. Literally millions of people feel every month that they are affected by what happens to that index. Consequently, the Bureau of Labor Statistics can get support by the users of that index. On the other hand, the Wholesale Price Index, which also happens to be widely used by business firms and the Government in many escalation contracts, is not widely known to the general public and, therefore, there is no constituency for that index among the general public. Yet the price indexes for individual industries constitute the foundation of the real product statistics of the national accounts of the gross national product, of the measurement of economic growth, and of the productivity or output per man-hour statistics of the BLS. Consequently, wholesale prices are vitally important statistics right at the heart of an integrated system of national accounts. It is that kind of statistic which your committee will need to support, because there may not be an outside constituency pushing for it.

Looking to new programs, such as manpower and poverty, let me cite another illustration. Back in World War II the House Appropriations Committee, trying to deal with the cost of living index as it was called then, asked why the Bureau of Labor Statistics could not prepare a family budget that showed in dollars how much it cost a family to live. In response to that request the Bureau did prepare a four-person standard family budget which was issued in 1946. In my first year as Commissioner, one of my duties was to present that budget to the public. It was the Joint Economic Committee of that day which gave me that opportunity. At any rate, what I want to emphasize now is that we produced that budget for a while but lost it in the 1950's. Now the Bureau of Labor Statistics is about to come out with a new set of budgets of this sort, based upon the 1960–61 family expenditure surveys which were conducted for the revision of the Consumer Price Index.

This time there are going to be several levels of budgets, three different levels, I believe. First, it will apply to several different types of families, including an elderly couple and eventually a one-person family. These budgets will bring out an important point, namely, that there is a wide differential in the cost of an equivalent standard of living in different sections of the country.

For example, using the latest figure in the Bureau of Labor Statistics budgets in 1959, I would like to point out there was a differential of as much as 15 percent above the average and 15 percent below the average in the cost of a four-person family budget in different cities. In such cities as Seattle, New York, Chicago, and Washington, D.C., the indexes of cost would be about 115, using 100 as the national average cost of that budget. On the other hand, in such southern cities as Houston, New Orleans, and Atlanta, the cost would be about 85 percent of the national average. So these figures of $3,000, $3,600, or $4,000 for a family of four represent a national average, but they do not reflct the real costs in any particular city. This differential is very important for policymaking purposes.

Now, the proposal which I wish to advance in that connection is that we will need for these new programs annual studies of family expenditures so that we may, from year to year, be measuring, testing, and recording the changing conditions of family living in this country. The Bureau of Labor Statistics did propose such annual surveys back in the 1950's, but it never caught on. I want to say to your committee now that I believe the time may be more ripe now, and I would urge that we consider the possibility of establishing such annual surveys of family expenditures. I think they would have many uses, not only for keeping a check on the Consumer Price Index itself, but also for the determination of consumer behavior. And, as we all know, consumers take nearly two-thirds of the gross national product. They are the largest single economic factor in our economy. We ought to know more about them.

Finally, in conclusion, Mr. Chairman, the last subject that I want to mention is international statistics. This is really the weakest of our fields, I think, in the statistical work of the Federal Government. At present, we have a world situation in which more than a hundred nations are entering the international markets. Some of them were scarcely in these markets at all only a year ago. Now they are all appearing in one form or another. International trade is expanding, international agreements are being made, and yet we really don't have adequate statistics of import prices, export prices, wages and fringe benefits, productivity or output per man hour, unit labor costs, the whole range of statistics that measure the position of the United States in international trade.

So, that is another field which warrants additional attention by your committee.

Thank you, Mr. Chairman.

Senator TALMADGE. Thank you very much for your broad and very perceptive statement.

Much progress has been made in the area of statistics during recent years and no one has worked harder for that progress and contributed more to it, than you, sir.

I would like to ask you a question or two concerning the matter of the coordination and integration of the existing program. In your view, is the machinery provided by law fully adequate for the task of coordination?

Mr. CLAGUE. I believe, Mr. Chairman, as far as the law itself is concerned, it is adequate. The Office of Statistical Standards does have authority under the 1942 act, I believe, to regulate the statistical agencies of the Government. I know in my own experience that they have

turned me down on studies that I wanted to conduct. They do have disciplinary powers and, of course, they sit right in the Budget Bureau. They can pass the word along, in the Bureau of the Budget itself, to the Estimate Division as to what they favor or disapprove.

On the other hand, I do find that in practice they have a tough job. They come up against the urgent need of administrative statistics and program statistics for an operating agency which has a job to do. While I have never sat in their chairs, I can understand, being a member of a major department and earlier, of the Social Security Board, that when a program agency decides it needs some statistics, it can exert a great deal of pressure on the Bureau of the Budget. That is why I presented some of my ideas to your committee here today. Your committee is an overall agency looking at these matters in the large. I think a good deal can be done if efforts are made to support and give confidence to the Office of Statistical Standards. I would recommend a larger budget and more staff. I think they need to follow more closely the statistical proposals of many agencies.

Chairman TALMADGE. Do you have any other recommendations besides a larger budget?

Mr. CLAGUE. No, I think not.

Chairman TALMADGE. You point out that the present system has worked well where good cooperative relationships have been established among Federal agencies. Why has progress not been greater in the making of series compatible, for example, such as unit costs, and pricing by industry, could be analyzed more easily and with greater accuracy?

Mr. CLAGUE. I think, Mr. Chairman, if I interpret that question correctly, this relates to the point I was making about the Wholesale Price Index and its constituent parts. This is a series about which the Bureau of Labor Statistics is quite aware of the limitations that exist. For example, many times we advanced proposals for collecting buyer's prices, as well as getting prices from the sellers. That would give a two-way shot at the true price. Was there a special discount? Was there an extra tied into the sale in order to conceal the price change of the original article? Well, the Bureau of Labor Statistics has been aware of this problem. The solution, I think, is more understanding in the Congress of the reasons for improving those particular statistics.

Chairman TALMADGE. In regard to cooperation among Federal, State and local problems, do you have any specific suggestions which might further the cooperation?

Mr. CLAGUE. Yes. I wish that there were some way in which the Federal agencies were able to provide the funds necessary for research, statistics and reporting that would be useful to them. I say this because, in a good many instances, the grant and formula programs that the Federal Government works out with State and local agencies—that is, the operating programs—frequently do not make special provision for statistics. The result is that the State and local administrators may not establish competent research and statistics units in their agencies. I think it would help if the Federal agency administering a program had some free funds which it could use specifically for research and statistical reporting from State and local agencies. In the success story I painted with respect to the Employment Security Program, that is the employer payroll reporting program, an important point is that the Bureau of Employment Security has the right to make 100 percent

administrative grants to the States. Therefore, they are able to provide statistical money to the States to perform special projects or maintain current reporting.

In my experience in the Bureau of Labor Statistics I never found much trouble in getting the States and localities to produce statistics when we could give them the funds to do it. When funds are provided on a formula basis, State and local agencies may not feel interested in putting more of their own funds into a national report.

Chairman TALMADGE. Next I would like to discuss the proposal for a National Statistical Servicing Center. I understood Mr. Stephen was in favor of a pilot project. Would you favor such a pilot project?

Mr. CLAGUE. Yes; I would. In spite of the cautions I expressed in my paper, I believe we should make a start. There is now no central point at which people anxious to use Federal statistics can find even what exists. I like Professor Stephan's suggestion that at least we start out trying to assemble information on the data that already exist. Later we could experiment with various kinds of interrelationships.

Chairman TALMADGE. Administratively, where should such a center be located?

Mr. CLAGUE. Well, I am sorry, Mr. Chairman, I have not thought that one out. Obviously, it would be independent in the sense that it should have its own budget. I hope it would. Now, in what agency to link it up, I do not know. It would be in some central place, perhaps the Bureau of the Census, if this a separate budget were maintained. Perhaps it could be attached to the Office of Statistical Standards, but on second thought, I believe it would better if that agency were not wrapped up with an operation of their own.

Chairman TALMADGE. Turning to another area of long-range planning, you point out that we need more analysis and interpretation of data and I could not agree more. What measures need to be taken so that there will be more analysis, particularly integrated analysis of series by several agencies?

Mr. CLAGUE. Yes. That is a separate point that I mentioned in my paper. I think that it is important for the producing statistical agencies to have adequate budgets, adequate funds, so that they can do a good job of interpreting and publicizing their own data, and by publicizing, I mean making available to the research fraternity as well as to the general public. In addition, we have in the Office of Business Economics a good illustration of an agency which is a secondary user, so to speak, of data collected elsewhere. A strengthening of the work of a general agency like that would result in better integration and interrelationships. There are a number of agencies like that which should be strengthened. Furthermore, some producing agencies may also be doing analytical work and should be supported in that function. For instance, in the analysis of economic growth, Mr. Bowman yesterday mentioned the fact that the results are obtained by the collaboration of many agencies and by the integration of data from a variety of sources. Perhaps my answer should be that any agency which has an integrating function to perform should be strengthened in that respect.

Chairman TALMADGE. Should the proposed major changes in statistical series be subject to some form of notice and hearings in the agency

similar to that required of regulatory agencies under the Administrative Practices Code?

Mr. CLAGUE. Mr. Chairman, I did not get the first part of that question.

Chairman TALMADGE. Should the proposed major changes in the statistical series be subject to notices and hearings in the agency similar to that required in the regulatory agencies under the Administrative Practices Code?

Mr. CLAGUE. I think, Mr. Chairman, I would not favor that. My feeling is that, as long as reporting is voluntary, a statistical agency should be able to try its hand at collecting information that might be useful to it. My impression is that formalized hearings patterned on those of the regulatory agencies, would be somewhat restrictive on the development of statistics. The procedural rigidities might inhibit the agency from developing new series and even handicap them in maintaining old ones. I think I would vote for some more flexible mechanism.

Let me cite one other alternative. In the Bureau of Labor Statistics we have two advisory councils, one from labor and one from management. We do check with them and they check with us, on all the series the BLS produces. In addition, they are our major respondents. I think this advisory system has worked very well. I do not recall any basic criticism of the work of the Bureau of Labor Statistics from either of those groups. I think I would vote for a more flexible mechanism rather than establishing a system of hearings patterned on the regulatory agencies.

Chairman TALMADGE. You raised an interesting question about lack of international statistics. Is there any way that we could stimulate the development of adequate programs of international statistics?

Mr. CLAGUE. Yes. Certainly, one way is to provide some funds to gather that information and develop the necessary statistics. I tried to develop those statistics when I was Commissioner, and I think that Commissioner Ross has continued the effort. The BLS has been eager to develop statistical series of import prices and export prices. There is no reason why we should not have an index of that kind, in addition to our indexes of wholesale and consumer prices.

Also, for a number of years the Bureau of Labor Statistics has had a small staff working on the problem of international productivity comparisons. This means comparing wages, productivity, and unit labor costs in different countries. The Bureau is coming out shortly with a special study on the steel industry comparing several Western European countries with the United States. I think, Mr. Chairman, that more interest should be expressed in this general field, and secondly, that funds ought to be provided because I can assure you the funds at present are really quite small in relation to the urgent need for facts of this sort.

Chairman TALMADGE. Mr. Clague, we appreciate very much your appearing before us this morning. You speak with a vast experience and knowledge in this field and have contributed very much to the committee. Thank you, sir.

The committee will stand adjourned until further call.

(Whereupon, at 11:05 a.m. the subcommittee adjourned, subject to the call of the Chair.)

APPENDIXES

APPENDIX I

"PURPOSES AND USES OF FEDERAL STATISTICS" [1]

CONFERENCE OF THE WASHINGTON CHAPTERS OF THE AMERICAN STATISTICS ASSOCIATION AND THE AMERICAN MARKETING ASSOCIATION

CONTENTS

PREFACE

The conference on "Purposes and Uses of Federal Statistics" was organized by the Washington Chapters of the American Statistical Association and the American Marketing Association. Briefly stated, its purpose was to explore the general framework of the federal statistical system to develop guidelines for collection of new data and to improve the uses of existing data.

This preface is intended to describe briefly the background of this report and to acknowledge the assistance of many persons who have contributed to it. While this report is the final product of this conference, it is only a first step towards the exploration of the subject. It may take many reports of this nature before the whole arena of federal statistical system is investigated.

With the passage of a heavy volume of new legislation by the 89th Congress the need for program orientation of federal statistics was clear. The American Marketing Association, whose members are major users of federal statistical data, wanted to develop a dialogue on purposes and uses of existing data and to focus on the needs for new data. The Association asked the Editor of this report to organize the conference. The American Statistical Association joined hands and thus a joint project was created. Invitations were issued to three groups: (1) Congressman Thomas B. Curtis of Missouri was invited to review the entire federal statistical system, (2) Those who were primarily responsible for producing at least part of the federal data—Mr. Robert Voight, Bureau of the Census; Dr. Faith Clark, Department of Agriculture; Mrs. Helen Lamale, Bureau of Labor Statistics; were in this category and (3) those who used the federal data—Miss Mollie Orshansky, Social Security Administration; and the discussants, Mr. Roye Lowry, Bureau of the Budget; Mr. Al Mindlin, District of Columbia Government; Dr. Daniel A. Swope, National Canners Association; Mrs. Catherine Martini, National Boards of Real Estate and Mr. Norman Frumkin, National Planning Association, filled this role. Thus, the conference could be called an exchange of thoughts between the producers, users and policy reviewers of the federal data. The conference was held at the Museum of History and Technology Auditorium in Washington, D.C., on April 11, 1967.

Sincerely yours,

PAUL AHMED, *Editor.*

[1] Edited by Mr. Paul Ahmed, National Center for Health Statistics.

ADDRESS BY MR. PAUL I. AHMED, NATIONAL CENTER FOR HEALTH STATISTICS, CHAIRMAN OF THE CONFERENCE ON "PURPOSES AND USES OF FEDERAL STATISTICS"

It is indeed a pleasure for me to be the Chairman of a conference which has such a distinguished list of participants. Our hope is that this conference will develop the much needed dialogue about the needs for new data and some information about its uses.

What are some of the issues for us to discuss today? For the record let me state these issues.[1]

(1) To me the most important issue in the data collection field is to develop data to evaluate program effectiveness of recent legislation. The 89th Congress, in which you, Mr. Curtis, participated with such distinction, created a variety of historic legislations. The Nation now needs to know whether these laws are fulfilling its purposes. For this we need to refine our demographic data, as well as to develop quality data. For example, now one year after the medicare legislation, one needs to know not only the hospital utilization patterns and the number of physician visits per person, but also the quality of such care received. Statistical planners should involve themselves in value judgments and provide answers to questions like these—What is an adequate medical care standard for a family? What are the adequate standards of preventive medical treatment? What percentage price increase in wholesale prices will demonstrate inflationary pressure etc.? The time is here when we need further orientation in this direction. An important step has been taken by the Bureau of the Budget which ordered all Federal agencies to develop across the board planning, programming and a budgeting system. This is useful as it will measure performance against objectives. Without this we may make progress without direction. The federal statisticians have made a tremendous improvement in the art of sample design, training of interviewers, Interviewer's Manual, transcription and coding manuals, etc. The sampling problems, however, are a means to an end and we should never lose sight of the end. We as researchers will all benefit by learning whether we are getting our dollar's worth and fulfilling our objectives. The cost effectiveness and cost benefit analysis, perhaps will tell researchers just that.

Perhaps establishment of an office of program planning and evaluation as an arm of Congress, an idea I discussed with you, Mr. Curtis, will demonstrate the usefulness of this approach.

(2) The second important area that is commanding attention now and requires even more attention deals with the administration of federal programs as they affect state and local governments. Information is required for two kinds of impact: Firstly, what are the resources of local political jurisdiction and localized labor markets. The implementation of legislation such as the Economic Development Act would be made more meaningful with these kinds of information. For example, a major concept imbedded in this Act is the concept of the economic growth center, that is to put emphasis on economic assistance in areas which have economic growth potential and which in turn can draw upon the resources of and feed income out to the depressed areas around its periphery. To implement this concept data on available resources must be developed. The other information which should be available is the impact of such programs as space and defense or medical research on local communities. When these large programs move rapidly up or down it is important for the state and local communities to have available the kind of information which will allow them to make reasonable forecasts of the impact of these changes on their local economy.

Another area in which local communities, states and the federal government need to join hands is to develop data for decisions as well as for deliberation. Unfortunately, at the present time data collection at all levels suffers from this weakness. No business will survive today or in the future by developing data which will provide no options for decisions of the management. All levels of government need to orient themselves towards obtaining data that will help them provide options and trade-offs. For example, trade-offs between urban freeways and urban mass transit will require information specifically highlighting the area of "options."

[1] The views presented here are that of the author and not of the National Center for Health Statistics.

(3) The third important objective of federal data collection relates to the formulation and evaluation of aggregate national policy. In the economic areas we have developed perhaps one of the most sophisticated accounting systems in the world. In a full employment economy, the need for *better, more sensitive,* and *more consistent* data are most critical because we are operating on a much narrower margin. The need for consistency is demonstrated by the fact that our price data are based on one set of categories and our cost data on another set of categories, while we are trying to determine the relationships among productivity, prices and wages. Of course I don't need to say that an accurate estimate of the impact of fiscal and monetary policy on states and local communities is far from being available.

(4) The fourth area in which the need for statistics exist relates to the whole array of programs of federal assistance to local and state governments. A number of these programs are based on formulas reflecting need —generally through a population criterion—and financial ability—generally through a per capita income criterion. Timely data, perhaps from tax returns, to indicate levels of personal income as of a certain date in the Census tracts and countries is needed. Such data will contribute effectively to the dialogue now developing on the Tax Sharing Plan.

To accomplish the above objectives there is a need to improve statistical series so that they may be interrelated. Statistical information concerning social conditions, political actions and economic results and potentials must be so designed as to be usable in interrelated ways for selected issues. This means that more emphasis must be placed on the development of the micro as well as macro analysis. This may mean emphasis on longitudinal studies which tell us how individuals, families, and group change their status over time and not merely how many are in a different status group at a different point of time. This may mean emphasis on family data, which tells us how many children in the family are educated or have health insurance coverage, and whether the family is with children or is "other" type of family.

This brings me to the central problem: statistical planners need to concentrate more on the "whys" of the situation. Granted, they are hard to obtain, but not impossible. The sources of economic depravity can only be alleviated if they are known. Why do some move out of a given status, while others do not and why do some move into the same status others have left? Answers to questions like these can be found by concentrating on the "whys." This necessary step will make our data more meaningful.

Let me conclude by saying that there is a need to make better use of available information by analyzing and highlighting the operational uses of the data. Also more data is needed to evaluate the operations of federal, state and local governments. For example, government procurement agencies need to provide valid data on the distribution of their purchases by industry and geographic location, and measure its impact on the local economy. Comparabiliy of data among states and localities at a meaningful level of detail is necessary and perhaps may require some coordination. A national data bank proposal is already in the offing. Well organized centers for assembling, collecting and retrieving data for various users will be an important step to provide the users something they need. More important however is to put something worthwhile in it so that users can take it out. This means in one sentence—produce decision oriented data.

These are some of the areas this conference will deliberate on. Now I know you are anxious to hear Congressman Curtis and the other distinguished guests. Congressman Curtis needs no lengthy introduction. He is a ranking Republican on the Joint Economic Committee and the House Ways and Means Committee. He is a life trustee of Dartmouth College and as President Eisenhower described him, he is "an exceptional member of Congress." He received the distinguished congressional service award from the American Political Science Association and LL.D., Honros Causa from Westminster College at Fulton, Missouri. Ladies and gentlemen I have the great pleasure of presenting the Honorable Thomas B. Curtis of Missouri.

Strengthening the Tools of Economic Policy

(Remarks of Hon. Thomas B. Curtis)

I think that some of the most rewarding and challenging—and, I might add, frustrating—jobs in government are held by those dedicated and highly professional individuals who collect and interpret the numbers we all live by today.

The civil servants who man our statistical agencies can never rest on their oars. No matter how much our economic statistics have been improved and refined, there is always a demand for more and better economic intelligence.

In a sense, this is a measure of their success. The gross national product and the balance of payments have almost become household words. The fact that over the past decade GNP and balance of payments information has moved from the financial page to the front page is clear testimony to the growing importance of statistical information for sound government and private decision-making.

In a small way, I have had the privilege of participating in this effort. I consider my service on the Economic Statistics Subcommittee of the Joint Economic Committee as one of the most interesting and important assignments of my Congressional career. We don't often make the headlines, but we do have the satisfaction of knowing that our work has contributed importantly to sharpening up the tools of economic policy.

There are a number of reasons why the demand for improved economic statistics is greater today than ever before.

First, the so-called "new economists" are attempting to "finely tune" fiscal and monetary policy in order to keep the economy at high employment without inflation at all times. One of the most critical obstacles to the successful use of push-button economic policy is the weakness in current statistical data on which policy decisions and forecasts of economic activity must rely.

Second, a host of new social and welfare programs have been enacted in recent years which depend for their success on statistical information which is now unavailable or available only in rudimentary form.

Third, changing conditions in an economy marked by a rapid increase in new technology, by a shift from manufacturing to services and distribution, and by a continuous exodus of workers from the farms create new policy concerns and new uses for statistical information.

I want to discuss each of these new demands for statistics in somewhat greater detail and along the way make some suggestions on how I believe we can and must sharpen up our tools of economic policy.

There was considerable discussion at the Joint Economic Committee's annual hearing on the President's Economic Report this year on the ability of the "new economists" to "finely tune" their economic policies to the needs of the economy. The Committee was repeatedly told by private witnesses that during the postwar period and particularly in the past year and a half, monetary and fiscal policies have tended to destabilize rather than stabilize the economy.

The policy problem is particularly difficult at high employment. Weaknesses in economic forecasts and analysis and in policy execution sharply limit the government's ability to shape appropriate policies. At high employment it is not enough to know whether a particular economic series is going up or down. We must know by how much the series is moving up or down. This is a more difficult problem and, for the most part, our present statstics do not provide the answers soon enough or with enough precision.

With nearly full utilization of resources, there is very little margin for policy error. Frequent changes in the degree of fiscal and monetary stimulus or restraint becomes especially dangerous in such a period.

There is another area where the government's attempts to influence the private economy run into difficulty because of gaps in our statistical knowledge. Until this year, the wage-price guideposts specified a single trend productivity figure, which the administration said was the proper guide by which to evaluate individual wage and price decisions. Our statistics on prices and productivity measures have been improved in the past several years, but their accuracy and reliability still leaves much to be desired.

Aside from many objections that can be made to the desirability or equity of the guideposts policy, it should be kept in mind that the measuring sticks now in use in the wage-price field provide a shaky and unreliable basis for a sound guideposts policy.

The point of these observations is that we need improved economic statistics and new and more reliable measures of economic activity. In addition to better price and productivity data, the minority members of the Joint Economic Committee this year made several recommendations which would result in better quantitative economic projections.

In our minority views in the Committee's Annual Report we suggested that there be quarterly revisions in the original gross national product forecasts for the year made by the Council of Economic Advisers. Along with the majority of the Committee, we also called for an improvement in the federal budget infor-

mation system, including quarterly estimates on budgetary receipts and expenditures and the presentation of the budget each year in the context of a long-run set of budgetary projections.

We also believe that a statistical series should be developed which measures wealth in the economy. This was recommended last year by the Subcommittee on Economic Statistics as a supplement to the gross national product series. Gross national product measures economic activity. This may or may not increase wealth. Certainly no one would say that the increase in gross national product that occurs during a war represents an increase in wealth or gives an accurate picture of true and meaningful economic growth.

A statistical series on wealth also would help us to evaluate federal spending programs by distinguishing between those which contribute to our human and material wealth and those which merely stir up economic activity.

One step towards the goal of improved federal expenditure policy would be the development of a capital budget for the Federal Government in order to separate out and identify wealth-creating expenditures which merit public support.

The second reason why we need an improved system of economic intelligence relates to the requirement for statistical information written into important legislation over the past several years. This includes such programs as the Federal Aid Highway Act of 1962, the Economic Opportunity Act, the Appalachian Redevelopment Act, and the Elementary and Secondary Education Act. As the Federal Statistical Users Conference has repeatedly pointed out, the requirements for information embodied in this legislation poses both opportunities and problems.

Much of the legislation requires the development of data on a wide variety of subjects, including population, employment by industry, per capita income, and income-consumption patterns in urban areas. If the data are collected effectively, we may develop an improved body of basic information relating to small geographic areas. At the same time, there is the danger of duplication of effort, waste of scarce resources, and an oppressive growth in the paperwork burden on respondents.

The requirements for the collection of more local, regional, and state statistics clearly calls for a coordinated approach by the agencies involved.

It is especially important that the data be developed in such a way that one area may be compared to another. Without comparable data, there will be no common measure to evaluate the success or failure of specific programs or to determine whether particular programs should be expanded or curtailed.

Another important area for a coordinated approach to statistics gathering is in the manpower training and retraining field. Training carried out under the Manpower Development and Training Act, the Economic Opportunity Act, in the military services and in the vocational education and apprenticeship programs has mushroomed in recent years. Yet we will lack an adequate tool for anticipating future needs for trained workers of different kinds in different areas of the country.

There is also a desperate need for more information on the training carried on by private employers. I was pleased to see that this has now been recognized by the Department of Labor. In the 1968 budget, $500,000 has been requested for surveys and research to be used as a basis for developing a sound policy for assessing the role which the Federal Government should play in training and retraining manpower. An important part of this study will be the gathering of information on the amount and kind of vocational training now provided by private employers.

The greatest gap in our manpower policy is a statistical series on job vacancies. Job vacancy statistics have been endorsed by the Joint Economic Committee, the National Industrial Conference Board, Dr. Walter Heller, Dr. Arthur Burns, and many other experts. The feasibility of collecting such statistics has been demonstrated by pilot studies conducted by both the National Industrial Conference Board and by the Bureau of Labor Statistics. Yet the Labor Department now appears to be stalled in its efforts to make further progress.

There are numerous important objectives which vacancy data would serve. Perhaps the most important is as a guide to public and private training and retraining programs. The key requirement of the Manpower Training and Development Act is that training be for a job vacancy that actually exists. Until we develop a series on the number, type, and location of job vacancies, we are really in the dark when it comes to developing sound training programs.

Data on job openings would also give a better picture of current opportunities in the labor market and where they exist. Even when unemployment is high, many jobs go begging. A survey taken in 1966 by the Manpower Research Council indicated that roughly 4 percent of all jobs in the country were vacant. The 3 million estimated vacant jobs was about the same as the number of people unemployed at that time. In Rochester, New York, the National Industrial Conference Board determined in its study in 1965 that there were actually more vacancies than unemployed persons.

Vacancy statistics could also serve as a leading indicator of the level of general economic activity. It would provide an indication of the ability of the economy to undergo the stress of structural change that might occur, for example, in a rapid defense build-up or in layoffs in employment stemming from shifts in demand or technological developments. They would also he helpful in determining the extent to which demand in the economy could be increased without running into wage and price inflation.

The precise cost of a reasonably satisfactory job vacancy program has been estimated as between $5 and $8 million a year. To undertake the collection on a quartely basis for approximately 80 major labor areas would cost about $2½ million a year. In terms of the more efficient use that would result from the billions now appropriated for manpower development, the investment would pay handsomely. I hope that all of you here today will lend your support to this important project with the objective of getting the Administration to move forward with the collection of these vital job vacancy statistics.

More information is also required on new skills that are developing in our rapidly changing economy. One of the most important economic questions today is whether automation creates more jobs than it destroys. I think it does. But these jobs are frequently geographically apart from where the jobs destroyed existed. And they are frequently in different skills.

As rapid technological change continues, skills change and become obsolete. No longer can a skill learned in the formative years assure lifetime employment. Training and retraining on or off the job are increasingly part of the work pattern.

The hard realties of training the unemployables are that they will not be capable of learning the higher skills demanded in the jobs newly created by automation. Those with jobs must be trained and willing to do so. They must take the new jobs, thus leaving their old jobs available to those bolow them in the ladder of skills, if they too will train. The unskilled and semi-skilled with training will fit into the jobs left vacant by those upgrading their skills.

A job destroyed is easier to identify than the new jobs which are created. A job destroyed has nomenclature; it has a human being attached to it. The newly-created job frequently does not have nomenclature and does not have an individual human being attached to it.

There are those who argue that automation destroys more jobs than it creates. I think our disagreement lies in the fact that they use a narrower definition and possibly a more correct one than mine. I use the term in its broadest sense. But whatever the definitional differences, we have got to do a great deal more in developing an early warning system on the new jobs that are being created, apply nomenclature to them, find out where they are located, and train men and women to fill them.

A related problem is the relationship of our military establishment to the civilian sector. I have seen articles over a period of years in Labor Department publications which point out that about 80 percent of the skills needed by the military have their counterparts in the civilan sector.

I have been distressed to find that in the military itself there seems to have been very little development of nomenclature for these skills or coordination between the military and agencies which are in the training field. Yet we are spending over a billion dollars a year, at least, in the military sector training people in skills which exist in the civilian sector.

These observations point up the crucial importance of identifying and applying nomenclature to newly developing skills and occupational categories. The Dictionary of Occupational Titles must be constantly kept up-to-date and, in my opinion, this could best be accomplished by putting it into looseleaf form. In this way, additions to or changes in occupational nomenclature could be periodically added to the Dictionary without waiting years for the publication of a new edition.

The final reason for the need for better economic statistics is the many changes occurring within the economy, such as one shift to services and the decline in

agricultural employment. As Commissioner Ross has pointed out, in the early years of the Bureau of Labor Statistics, emphasis was on such matters as industrial injuries, labor turnover, labor management disputes, and mass unemployment. Many of today's critical issues in economic policy involve manpower and human resource development, equal employment opportunity, elimination of poverty, regional economic development, and the problem of hard core unemployment.

Let me elaborate on one example. Changes in the labor force in recent years have had a significant bearing on the significance of our employment statistics. The rapid growth of welfare programs, including unemployment insurance, help maintain a flow of income during periods of unemployment. Today even a man of modest means can practice some discrimination in job selection as a result of these programs. He can better afford to shop around for a job suited to his needs. and interests. Our statistical measures, however, do not take into account this voluntary aspect of unemployment nor the fact that it is probably increasing in our society today.

Our labor force data are also affected by the rapid increase of working women and teenagers. Many of these workers are parttime or intermittent workers, a fact which creates unavoidable intervals of unemployment. The existence of more working women and teenagers also reduces labor force mobility. Neither group is as able or willing as male employees to terminate their employment and to take a job in another city or often in another area of the same city.

Other important changes have been taking place in our economy. Today we are beginning to look at a person's full life—his tender years, his years of education, his productive years, and his years of retirement. We have been developing the mechanisms and the programs for spreading a person's lifetime income from his productive and earning years to the non-productive years.

The first mechanisms developed were in the nature of savings from the productive years to provide for retirement pensions, annuities and retirement systems. At the same time we are developing the mechanisms whereby people can pool their common risks against an untimely imminution of earning capacity from (a) death, (b) disability through accident or sickness, (c) interrupted earnings resulting from, e.g., military service and economic downturn, (d) and now, obsolescence of skills.

Since World War II, we have been developing the mechanisms to spread income forward in anticipation of earnings from the more productive years to the less productive years. We have developed new forms of consumer credit to encourage home ownership, purchase of consumer durables, and most recently, to provide the capital investment for education. A great deal of today's consumer credit constitutes real savings inasmuch as the expenditures relate to increased wealth and increased earning capacity, not to mention increased standard of living of the debtor. It is indicative of this understanding of lifetime income that income averaging techniques, crude as they are, were introduced into the federal personal income tax laws in 1964.

The emphasis needed for further development lies in phasing individuals into the labor market and phasing them out again on retirement. One does not abruptly—or should not abruptly—enter the labor market or retire from it. The better retirement systems we are developing permit a phasing out, utilizing in different ways the talent perfected by experience of the older citizens. The better educational systems use a variety of phasing-out mechanisms.

Above all, we are beginning to understand that people are not committed full time in the labor market. The eight hour day and the forty hour week attest to this. Hopefully we will begin to move more broadly into the eleventh month year and possibly to the concept of the fallow seventh year—the sabatical leave. However, the women in our society are increasingly entering the labor market before marriage only to retire for the period of raising children, and then to reenter later on a planned, part-time basis, which frequently later develops into full-time employment.

All of this brings me back to my opening point—the need for new and improved statistics to meet changing conditions in the economy.

These numerous suggestions for new statistical information will probably discourage even our most enthusiastic and dedicated statisticians. As usual, they are being called upon to perform Herculean tasks and then given inadequate budgets and staff to do the job.

The budget treatment of our statistical agencies is one of the best examples I know of being penny wise and pound foolish. In the fiscal 1968 budget, less

than nine one-hundredths of one percent of total new obligational authority is earmarked for statistical programs.

I often wonder how much we could save in federal expenditures if we had available improved statistics to provide better guidance to policy-makers in developing new programs or operating old ones. I would guess that billions could be saved compared to the relatively small outlays that would be necessary for additional progress on our statistical programs.

Barring any such breakthrough, we shall have to satisfy ourselves with what we have available. Our progress will be slow, but I hope it will be steady. The interest on groups such as your own is certain to have an impact, and I wish you well in your efforts to speed up progress in this important work.

Census Tools for Marketing[1]

Robert B. Voight, Special Assistant, Office of the Assistant Director for Research and Development, Bureau of the Census

INTRODUCTION

The Bureau of the Census has enjoyed a long and fruitful relationship with the American Marketing Association. We have benefited greatly from the advice and counsel of the AMA Census Advisory Committee established in 1946. As one of the four major Federal statistical agencies, the Census is a basic producer of data, the use of which can lead to more intelligent marketing judgments and the better functioning of the economic system. We are interested in making Census statistics better working tools for the marketing profession both from the practical level of helping the marketer increase or keep his present market share to improving and enriching his marketing research efforts and to aiding and strengthening the decisions and planning of marketing directors and corporate planning officers.

Before discussing in detail the various types of Census data and services useful for marketing and presenting some examples of the application of Census information to marketing problems, let me mention briefly the subjects of the Census Bureau covers. We take censuses of population and housing every 10 years; every five years we take census of agriculture, business, manufactures, mineral industries, transportation, and State and local governments. We produce a great many monthly, quarterly, annual and special reports covering these subjects and a substantial amount of foreign trade statistics showing values and quantities of exports and imports by commodity detail and by country. Many of you, I am sure, are familiar with some of these. In addition we assemble and summarize statistics from a variety of sources—Government and private, and publish them in a convenient source books such as the recently introduced Pocket Data Book and the U.S. Statistical Abstract. The 88th annual edition of this best seller will be issued this summer. It will contain more than 500 pages of information bearing on some aspect of marketing.

Unpublished Data and Computer Tapes

Although many of you may be familiar with the Census Bureau's published reports, it may be of interest to know that these represent only a small portion of the total data resources in the Bureau. We have on file a variety of data on computer tapes and punch cards which represent almost limitless possibilities of subject cross-classifications and selections of geographic areas. Only the most essential and most widely useful data are presented in the published reports. In some instances, the tabulation programs of the major censuses and surveys have provided, as a by-product or adjunct to the preparation of the published figures, statistics in unpublished form showing additional subject or geographic detail. These tabulated but unpublished materials are available at nominal prices which cover the cost of making copies. More and more of the unpublished statistical aggregates are now being made available on magnetic tape compatible with the customers computing equipment or on punch cards. Another sizable data source is to be found in the special compilations and tabulations which have been prepared by the Bureau at the request of various Federal agencies and private organizations. In order to keep the public informed of special tabulations

[1] Given at a conference on "Purpose and Uses of Federal Statistics," sponsored by the American Marketing Association, Washington, D.C., April 11, 1967.

that are prepared, the Bureau regularly lists them in the *Bureau of the Census Catalog*. Some examples of these special tabulations that might be interest to marketers are:

Establishment and sales data for grocery stores by State, county, city, and standard metropolitan statistical areas, provided on one reel of tape, based on the 1963 Census of Business.

Establishment and sales data for passenger car dealers and automotive repair shops by State, county and city.

Number of retail stores by census tracts and kinds of business.

New gasoline service stations and repair garages authorized by building permits by States and SMSAs by month for 1965.

Construction plans of State and local governments concerning construction projects not yet begun as of June 1, 1965, and providing nationwide data on number and dollar volume of construction projects. by type of government, size of project, function, estimated date of construction start, interval for planning, interval from start of planning to start of construction, and interval from completion of planning to start of construction.

These are only a few of the several hundred special tabulations that become available each year.

Source Books

In talking about Census Bureau statistics for marketing, I think it would be helpful to describe the major source and data guides that provide a summary of the various types of data available and how you may obtain them.

At the outset let me suggest that the best way to keep informed of what is being produced by the Census Bureau is to subscribe to the *Bureau of Census Catalog*. The subscription price is $1.75 for four consecutive quarterly issues—each quarterly issue cumulates to an annual volume in the fourth quarter. In addition you receive 12 monthly supplements to keep you up-to-date. Subscriptions are available from the Superintendent of Documents at the Government Printing Office. In addition I would suggest that you write to the Bureau of Census and ask to have your name placed on the mailing list for all publication announcements and order forms. This may flood you with paper for a bit, but after you have become familiar with them you can begin to discriminate and ask to be sent only those covering subjects which you feel will be of particular interest and value to you.

Turning to Census Bureau data source books, I should point out that the *U.S. Statistical Abstract* is up-dated annually to reflect the changes and developments in the social, political, and economic structure of the United States. For example, the latest issue contains 66 entirely new tables and significant inclusions in many of the other 1,200 odd summaries. To name just a few new items of value in the field of marketing—salaries in private industry, income of the aged population, reasons for retirement, income of white and negro families, sales of capital assets by individuals, mergers and acquisitions in manufacturing and mining.

Shortly after the close of the year, the Bureau will issue the 1967 edition of the *County and City Data Book*. This source book is compiled about three times a decade. Its primary purpose is to provide a selection of recent statistical information for counties, cities, and other relatively small geographic areas. More than 160 statistical items are presented for each county and city of 25,000 or more. They are also shown for regions, divisions, States and the standard metropolitan statistical areas. Descriptive text and source notes are included to help the user better interpret the figures shown.

To cite just a few possibilities for marketing purposes, this source book provides data on the population, education, employment, aggregate and median income, housing, bank deposits, time and demand; savings capital of savings and loan associations; local government finances; employment and payrolls in manufacturing; retail sales, employment and payrolls; agriculture-size of farms, value of products sold, etc. In summary, it provides a detailed picture of the situation social and economic in each county and city of the country and represents an informative book for marketing needs. This source book is available also on magnetic tape and punch cards for those users who wish to take advantage of further analytical manipulations of the data in combination with other information they may have in machine-readable form.

Beginning with the 1964 issue, the Bureau now publishes *County Business Patterns* annually. In addition to a U.S. Summary report, a separate report for each State is published. These reports, show first quarter employment and taxable payrolls and the number of reporting units (i.e., establishments) by employment size class by major industrial classification. Here is a frequent and timely pulse

feeler of the economic situation in each county and for the standard metropolitan statistical areas in each State. It is especially useful for analyzing market potentials, establishing areas in which to assign customer cultivation and advertising efforts, comparing past sales volume with present volume by area to detect potential areas for expansion of sales efforts, and to analyze the economic mix of various areas for new possibilities for services and sales expansion. This information is also available at cost on punch cards and computer tapes.

CENSUS DATA GUIDES

In trying to provide better service to and communication with potential users of Census data and at the same time reduce the amount of time required to determine whether certain statistics are available, we are providing various data guides. They should be of particular interest to the marketing profession.

We have just recently published a *Directory of Federal Statistics for Local Areas*. This guide contains no statistics, rather it is intended to serve as a comprehensive finding guide to current sources of Federally published statistics for governmental and socio-economic areas below the State level. It provides information useful to city, county or State planning agencies; marketers interested in firm location and marketing possibilities; and organizations concerned with local or urban problems. It includes a description of sources for data on population characteristics, health and vital statistics, construction and housing, labor and employment, income and earnings, prices, banking, commerce and trade, manufacturing, transportation and communications, agriculture and fisheries, mining, governments, law enforcement and other data.

Leading Census Programs is a tabular summary of the major Census programs. In 30 pages we have shown when the program was initiated; the subject matter included; the kind of coverage—complete census, sample surveys, or estimates; how the data are collected; the frequency and reference period, the timing of publications, the content; the special tabulations and unpublished data available; the use and users; and qualifications, if any. We believe it is an exceedingly useful document to have on any marketer's desk. It is available from the Bureau without cost.

We have found the *Guide to Census Bureau Statistics—Subjects and Areas* to be another type of source information popular with marketers. Interestingly, it evolved out of discussions and conferences such as this one in response to questions about what the Census collected and published. It also describes the various types of geographic areas, some 20 in all, observed by the Census, and contains a section on how to obtain Census materials and services. It's free.

USES OF CENSUS DATA FOR PINPOINTING MARKETS

One of the major values of Census data for marketing purposes is the wealth of small area detail available which allows the marketer to pinpoint and compare sales potentials by geographic units ranging from city blocks or combinations thereof to tracts, cities, and counties.

Census Tracts are now regular divisions of all of the standard metropolitan statistical areas. Many smaller cities also have been divided into Census Tracts. In the 1970 Census data for some 45,000 of these will be published covering all of the SMSAs. These tracts provide a useful framework for studying of land-use data, for marketing surveys, for the study of neighborhood housing conditions and the distribution of various types of housing equipment, and for determining facility locations and other business decisions. The Census publishes more limited statistics by city blocks within Census Tracts.

This wide range of data published by Census Tracts permits the marketing researcher or planner to select those areas where it is evident that his product or service or resource will be readily sought. Another important use of Census Tract data is the study of various combinations of contiguous tracts made by the marketer to determine the most suitable location for a branch unit in terms of the neighborhood he seeks to serve.

Census tracts are a unique tool in that they permit historical comparability for small areas within the urban complex so that the amount of change within small areas of the complex can be measured. The fact that they are defined in the same manner using the same general rules in all SMSAs permits a measure of statistical stability in comparing the small area characteristics of several different SMSAs in determining sales potentials, investigating site locations and changes in market outlets.

Another advantage of the Census tract system is the fact that a great deal of other local data are available for these areas for analyses, such as business licenses, land use, construction and demolition, sales, clients, and subscribers, to mention only a few. They also serve as reasonably comparable building blocks in establishing larger areas for marketing purposes and they are a handy device for manipulating and analyzing various potential spatial aggregations for determining sales quotas, potential "walk-in" customer populations, the variety of possible users for a particular site, etc.

Many products are sold mainly to certain groups in the population. The kind or amount of clothing purchased is related to age, sex, occupation and education as well as income. Buying books is related to educational attainment. Certain products are purchased chiefly by house owners. In building a branch store or seeking a new rental quarters a merchandiser needs information on the number of people in the area and their characteristics. Census tract data provide the answer.

Tract data are used by real estate organizations, banking, savings and loan and finance companies, insurance companies, mail advertising firms, and newspapers in their planning decisions. These uses provide a second level of helpful information to the marketer who makes the effort to investigate them.

CONSTRUCTION STATISTICS

The Census Bureau collects and tabulates data on building permits from 12,000 permit-issuing jurisdictions in the United States in which 80 to 90 percent of all new residential housing units starts occur as well as 85 percent of all new non-residential building construction. The Census Bureau provides data on the 3,900 most active places each month and all 12,000 places annually. The reports give data on the number of one, two, 3–4 family and 5 or more family buildings authorized. The non-residential authorizations are shown in 15 categories including industrial, schools, hospitals, stores, office buildings, etc. In addition the additions and alterations to existing residential and non-residential buildings and their permit valuations are provided. We plan in the future to have building permits identified by Census tract so that you can tie income, age groups and a host of other Census and non-Census data to new construction statistics.

The Bureau is planning to take the first national Census of Construction this year since 1939. This will cover all the large employer construction establishments and a scientifically selected sample of the smaller ones. It will include contract construction, subdividing and developing, and operative or merchant builders including all types of subcontractors and specialty contractors. The data will also show the types of construction work undertaken, whether the project is publicly or privately owned, the location and the amount of subcontract work done for other contractors. This information represents a new tool for assessing the marketing potential in the construction industry.

BUSINESS STATISTICS

In addition to *County Business Patterns* mentioned in the discussion of Census source books, one of the most useful series of publications you could use in analyzing business activity within cities and standard metropolitan statistical areas are the *Major Retail Center* reports from the Business Census. The present series covers 132 Central Business Districts located in the largest 116 SMSAs and more than 1,000 major retail centers in these SMSAs. The Central Business District is that downtown area of very high land valuation characterized by a high concentration of retail businesses, offices, theaters, hotels, and service businesses and an area of high traffic flow. Major retail centers outside the Central Business District are those concentrations of retail stores which include a major general merchandise store—usually a department store. They include not only the planned suburban shopping centers but also the older "string" street and neighborhood developments which meet the above prerequisites. Frequenly a major retail center includes the planned center plus stores in adjacent blocks that contain at least one store in the general merchandise, apparel, or furniture-appliance categories. Number of stores and sales are shown for convenience goods stores, shopping goods stores, and all other retail stores. The reports contain maps showing the CBD in street detail, the MRC locations, and a physical description of the area of each MRC. These reports provide a ready index of business growth, comparisons between shopping areas, and over time, portray the financial volume and physical growth of business activity in a city and its surrounding suburbs.

80–826 O—67——11

It is anticipated that the 1967 Census of Business will expand this particular data series to all the standard metropolitan statistical areas, thus providing information for more than 100 additional Central Business Districts and an estimated 200 to 300 additional major retail centers.

MANUFACTURING STATISTICS

The Censuses of Manufactures and Mineral Industries provide key measures of activity in manufacturing and in the various mineral industries. The area bulletins are used primarily to answer questions regarding the range and magnitude of industrial activities in specific geographic areas. Industry bulletins show nationwide totals for each industry and indicate its geographic spread. Finally, there are subject reports dealing with such topics as inventories, capital expenditures, and size of establishment. These meet many specialized needs for information regarding trends and developments in our industrial system.

Two special series of reports on the *Location of Manufacturing Plants* of the 1963 Census of Manufactures have been issued recently by the Census Bureau. These reports provide information on the number of manufacturing establishments in each of the approximately 430 manufacturing industries classified according to their employment size in the State and county within which they are located. The information will be available on computer tape also.

The Annual Survey of Manufactures carries forward the key measures of manufacturing activity that are covered in more detail every five years in the Census. It provides, during intercensal periods, basic statistics which serve as benchmarks for current business comparisons and as measures of industrial production and productivity of interest to marketers.

NEW MARKETING TOOLS IN FUTURE CENSUSES

In the next series of Population and Housing Censuses we undertake, we expect to provide much greater geographic flexibility. We expect to identify data down to the individual side of a city block in urban areas so that it can be grouped into any geographic configuration you call for. Also we expect to identify these block sides by coordinates of latitude and longitude. This will open up a whole new area of potential values to marketers. For example, you could ask us to give the characteristics of the population and housing in one mile circles from one of your branches or from a prospective future branch location you are considering. We will be in a position to map cities for you in square mile grids or other configurations on the computer to graphically display the characteristics of such areas to provide an income and housing value chart, commuting patterns, new housing, old housing, deteriorated housing, etc. Some of the future possibilities are quite exciting since you will be in a position to insert your own data into such matrices by using a Census address coding guide and derive certain ratios and percentages, densities, etc., based on the Census aggregates.

For assigning geographic codes in the major urban areas we are preparing computerized address coding guides which will include street names, block face identification, intersecting streets, the range of address numbers for each block face according to census tract and zip code area, and the area identification codes required for Census Bureau tabulations. These block face coding guides will provide great flexibility in making it possible to tabulate Census and other information for any desired areas. Both computer tapes and printouts of these address coding guides will be made available along with Census maps, as tools for small area identification.

Through the use of address coding guides the Census will be able, for the first time, to record information for geographic units ranging in size from one side of a city block to an entire urbanized area. Tabulations will be possible, for example, for both sides of a street or streets through a city or area. The limit to the flexibility of the information available for various areas will be disclosure rules and the cost of tabulation. We do not plan to provide this capability for all city delivery areas but hope to accomplish this for entire urbanized areas and for cities of 25,000 or more inhabitants (not all of which may be possible), or for the bulk of such areas. It may be necessary to cut back on some of the smaller cities and urbanized areas. A further limitation is the extent of city delivery postal service; beyond these areas we do not plan to code on the block face level, although reporting by block is expected to be feasible to the boundaries of urbanized areas. In any event, Census data will be available in far more geographic detail than ever before.

With little added effort a copy of the "Census" address coding guide for an area can be modified locally for broader use by the addition of identification codes for areas such as sales territories, delivery routes, and so forth. With this accomplished, local flexibility for marketing purposes is virtually unlimited.

We anticipate that this development in data flexibility, which will provide a standard set of small geographic bits as building blocks in assembling data in virtually unlimited types of areas, will be one of the major contributions of the 1970 Censuses to planners in all fields.

The second possibility of considerable importance in certain fields of interest stems from our proposal to identify the locations of blocks or block faces by coordinates. Although this is not a certainty, it is definitely in our plans. In any event, the system will be so designed that coordinates can be introduced later if resources are not available to introduce them into the system prior to the Census.

Within the areas covered by address coding guides, we expect to have coordinates for block faces; for other parts of urbanized areas, and for rural areas, coordinates probably will be established for "standard locations" consisting primarily of Census tracts or minor civil divisions in the rural areas. The coordinates will be recorded in degrees of latitude and longitude to four decimal places, that is to 36-feet at most, but those who wish to employ state plane or other standard coordinates, rather than latitudes and longitudes, will be able to convert them.

This program opens up a whole area of data availability and analysis heretofore not attainable. Spatial relationships of social and economic data can be examined, density and distance correlates established, statistical aggregates established in terms of distances from a given point, in equal squares of certain size, or other configurations. The characteristics of people and housing within a certain distance of a proposed shopping center can be examined in considerable detail to determine the potential shopping volume. Many other important uses will come to light only after considerable exposure of this facility after the Censuses are taken.

OTHER FUTURE DEVELOPMENTS OF INTEREST TO MARKETERS

We have heard from many places smaller than 50,000 asking that they be provided with Census statistics on a tract basis. The Bureau has announced that it will recognize tracts which are established in these cities through local initiative. Recognizing the tracts means that we will tabulate statistics by tract, but it does not mean that we will be able to publish the tabulations for these smaller areas in the regular Census reports. Nevertheless, the unpublished tabulations will be available at the cost of reproduction.

Housing census results have regularly been published for city blocks. Limited statistics were issued for city blocks in all cities which had a population of 50,000 or over in the 1960 Census. In addition, the Bureau had announced that other communities which wished to have these statistics by city blocks could arrange to have them if they would prepare the necessary block identification materials and reimburse the Bureau for their added costs. Block statistics, including a limited number of housing items and the total population, were published for nearly 750,000 city blocks. In 1970 we hope to extend the block reports to the closely build-up areas surrounding cities of 50,000 and over, i.e., we hope to include the entire urbanized area. An attempt will also be made to provide block statistics for cities with a population of 25,000 to 50,000. If these additions can be effected, the total number of blocks is likely to be on the order of 1,600,000; roughly twice the number for which reports were issued for 1960. It should be stressed at this point that these expansions of the block data are hoped for; however, it cannot yet be stated with assurance that resources will be available to do this.

There is a clear call for greater detail on place of work from the 1970 Census. If it is possible to secure reasonably accurate identification of places of work by street and number, as in the case of residence, the coding of work place to block faces to be aggregated by small areas as desired will be technically feasible. This, coupled with information on methods of transportation used to go to work, will provide information of considerale interest to many market research and planning people.

Of specific interest to this audience is the fact that the Bureau expects to provide from the Housing Census two more categories in the value of property to identify dwelling units in the $35,000 to $50,000 group and those $50,000 or more. Information will be available on families living in high-rise apartments since the respondents will be asked to indicate whether they live in building

having up to 13 or more floors. This should provide another dimension to the study of potential customer densities in small geographic areas.

Family income will be tabulated in $1,000 intervals up to $12,000, $12,000 to $15,000, $15,000 to $20,000, $20,000 to $25,000 and $25,000 and over. The 1970 questionnaire will ask for considerably greater detail on income other than earnings. This expansion in income data will be of considerable value in market research planning.

POPULATION ESTIMATES FOR SMALL AREAS

Although the Bureau of the Census has a well-established State and national population estimates program, it was not until 1963 that the Bureau undertook the development of population estimates for standard metropolitan statistical areas. The initial report in the standard metropolitan statistical area series was released in April 1964 presenting estimates for the 15 largest standard metropolitan statistical areas and their 68 constituent counties. The program has expanded each year since then and currently estimates of population are published for each of the counties in the 55 largest standard metropolitan areas of over 500,000 population in 1960. These 55 metropolitan areas include 190 counties, with a 1965 combined estimated population of about 90 million.

By next year the reports will hopefully include population estimates for the largest 75 metropolitan areas of the country, including about 230 counties. Each of these standard metropolitan statistical areas had more than 300,000 population in 1960. A major target of the program is to provide estimates for the 100 largest metropolitan areas and their constituent counties by the end of the decade.

The Bureau of the Census does not, as part of its regular program, prepare population projections for areas below the State level. Recently, however, the Bureau has started to explore the possibility of cooperating with the States to develop comparability among small-area estimates. The ultimate objective is a program in which each State would prepare "official" estimates of the population for counties and cities, using methods and sources of data which meet certain standards that are mutually accepted. These figures would be recommended for use in relation to all Federal and State programs and the Census Bureau would publish these so that they would be available for general distribution. The Bureau would continue the regular estimates of population for the nation and the states. Meetings are currently being held with representatives from the State governments to see if such a program can be established.

The national population projections for 1985 range from approximately 239 million to 273 million as compared to a projection of 197 to 198 million this year. By 1975 it is expected that the age group 25–29 will nearly double from 11.3 million to 19.3 million while the number in the age group 45 to 54 will actually decline slightly. Consider the changes in your marketing strategy that this will bring about!

SUMMARY

This discussion has focused on sources of Census data of interest to the field of marketing. The availability of data in unpublished form, on computer tapes, and available from special tabulations has been described. Suggestions have been given concerning data guides that should be of value to marketers. We have described some of the Census tools for pinpointing markets. New marketing tools in future Censuses and other future developments we believe will be helpful have been outlined.

We welcome your reactions and suggestions for further developing Census data as tools for marketing.

Poverty Statistics—What They Say and What They Don't Say

Mollie Orshansky, Social Security Administration, Office of Research and Statistics

It's an exciting time for the social sciences. We live in a self-conscious and yet trusting society: A society more acutely aware than ever before of its inequities and its iniquities, but yet convinced they can be corrected. A visiting Englishman once remarked that in the U.S. we believe all we have to do to solve any problem is to pass a law about it. I should like to amend that; today we think we need but toss it into a computer and the one and only solution comes out complete with seven carbons.

The Social Scientist, along with the other eggheads, has been invited out of his classroom, the library, the professional meeting, the government cafeteria, and the other places he customarily flees far from the madding crowd, to apply his special art to issues in the public domain. It is a challenge to respond, to bid for the personal satisfaction that goes with knowing something you do can affect public policy and it offers an opportunity to participate in major decisions. But it does not come without its price.

It will strain one's humility; it will strip away the defenses our scholarly colleagues normally allow us—the technical terms unintelligible to the uninitiated; the refuge and subterfuge of qualifying footnotes; the escape hatches of the plea for more and bigger samples; and that ultimate safety valve, the interim report. After all, if further research on the subject is not going on, it should be! To be sure one can cling to these demurrers but time and Congress will not wait and the unresolved questions will get bigger and harder or even different. Worst of all we may amass more data only to find as have others before us that many a beautiful theory will be inhibited by the facts.

So what else is new? And anyway what does all that have to do with the subject of the morning? Well, there is always the unwritten rule guaranteeing a speaker the right to a bit of high-flown rhetoric. But there is a less tenuous connection. I contend first that research should be purposive; that the techniques selected must be suited to these ends; and above all that we must deliberately and knowingly distinguish between the fact and the fancy—the data used and the arbitrary interpretations assigned to them. If you recall I did say we were going to apply our art to public issues. Now let us begin.

My assigned topic is poverty. Unlike some other statistics, those relating to poverty are not primary but secondary. That is, they are derived from statistics obtained for other reasons. No one yet has proposed determining who is poor by the simple expedient of going out and asking the direct question, though it might well turn out to be a chastening exercise. Instead we derive our counts by appliyng analytical techniques, seasoned in liberal measure with statistical inference, and diluted with a healthy helping of divination. To quote the Commissioner of Labor Statistics, Dr. Ross, "for some questions prayer is more relevant than calculation." Perhaps more than any other socio-economic indicator poverty, like beauty, lies in the eyes of the beholder. But is this really a bad thing.

Over the years economists and statisticians, particularly those in government employ, have given much thought to the appropriate role of research. Is it and should it be basic or applied, pure or so to speak, impure. Does the social scientist overstep his mark when he makes judgments; should he even if requested make recommendations; or does he serve best when he merely arrays the facts? (The word is "array" rather than "present" because not even for the sake of discussion do I concede data really can be untainted by the values of the investigator.)

Today, to be sure, in an era of problem-solving and model-building, with every investigator entitled to his own computer, such soul-searching could seem a little passe. But whatever the possibilities for socio-economic research in general, with poverty numbers one can only be more subjective or less subjective; one cannot be nonsubjective. Ultimately in making his decisions the analyst will, in choosing standards, have to resort to the unalterable dictim my father used when I questioned his pronouncements: "Why? Because I say so, that's why!"

Like any other level-of-living index poverty can exist only in a frame of reference. It must be localized in place and in time. Indeed, beyond the physiologic exigencies of food, water, and warmth—and here in the U.S. we are well beyond them—it can have little meaning without contrast, without someone to notice. To be poor implies to be without those things or those opportunities considered essential in a given society, or more bluntly to be deprived of those goods and services and pleasures which others around you take for granted as a right. Incidentally, in today's one-world those others may be citizens of other countries as well as one's own. These days when it is in fashion to talk about the invisible poor, let us remember that if they be invisible to us, we and what we enjoy are very visible to them.

Does this mean then that inequality per se is equated with poverty? Obviously not. The Lorenz curve with its isometric norm is a little outmoded. In a nonegalitarian money economy built on a work ethic, with wages set not by need but by ability and application, there will always be some who have less than others. This in itself need not disturb us. It is when having less turns into having far too little and when, moreover, the burden of such privation con-

tinually falls heavier on some groups than on others, that the public conscience rests uneasy. And so we come to a working definition of poverty as a policy issue—the number and kinds of people to whom we wish to direct public concern. It is in this context and only this that the current statistics on poverty have relevance.

In a sense we can establish the level of living—or the level of concern—to embrace as many or as few as we please; we can choose the criterion so that everyone it counts will be among the undoubted poor but others equally deprived will escape our notice; or we can ensure that all the deprived will be on the poverty roster but only at the price of identifying as poor many who ought not in fact be so designated. There is in short no perfect scheme, and none that is value-free. Having chosen a level, we can only try to give every kind of family or concumer unit its fair chance to be numbered among those who claim our attention. If we must foreswear infallibiilty, we can at least cling to the rule of reason.

The real question at a conference on purposes and uses of Federal statistics may be not why one derived the particular measure used but indeed why we did anything at all. Suffice it to say that the Social Security Administration like nature abhors a vacuum. The poverty line originally set by the Council of Economic Advisers was a tentative $3,000 cutoff for a family of two or more and $1,500 for a person living alone. Family size differentials were acknowledged to be important, but were not readily at hand. Accordingly a couple with $2,900 could be considered poor while a family of six with $3,100 would not be. Inevitably this led to an understatement of the number of children in poverty relative to aged persons. And it was really this inequity SSA tried to resolve. Inasmuch as the $3,000 figure coincided with an amount earlier set by SSA as a rough minimum for a family of four, it was not too surprising that for the year 1963 the number estimated as poor with new criteria tapered for family size differed little from the number derived with a single poverty line undifferentiated for size. The real difference was in the composition of the group identified. The new cluster of poverty lines lowered the number of poor families with an aged head from 3.1 million to 1.5 million, while it raised the count of children being reared in poverty from 10.8 million to 15 million.

As of 1965, were we to revert to the Council's original measure, the poverty roll would include 3 million fewer persons than our current estimate, as the figures below indicate:

	Millions poor in 1965	
	Original CEA definition	SSA poverty index
Total	29.8	32.7
Unrelated individuals	4.7	4.8
Under age 65	2.0	2.1
Age 65 or more	2.7	2.7
Members of families	25.1	27.8
Children under 18	9.3	14.3
All other	15.8	13.4

The difference in the profile of the poor calculated by the two methods is perhaps greater today than before because of the fact that large families, particularly those with a woman as the head are becoming a larger proportion of the total households in poverty, and it is the large households that the two schemes treat differently.

In 1965 then, on the basis of the data collected by the Bureau of the Census in its March 1966 *Current Population Survey*, the tally of the poor stood at 32.7 million or about a sixth of the noninstitutional population, compared with a fourth so designated in 1959, as the tables in the attached note suggest. The tables indicate also that by an alternative measure, higher than the poverty line but still far from what one could designate as gracious living, there were all told a total of 47 and a half million who lived within the bleak circle of poverty or at best hovered about its edge. Although this represented considerable improvement over the situation in 1959 (using the same criterion adjusted only for changes in consumer prices) the improvement was not shared by all in equal measure.

Indeed, those groups with income farthest away from their estimated need in 1959 were the groups that showed the least improvement in economic status.

POVERTY IN THE UNITED STATES, 1959-65

A brief summary of what we know about the poor as we define them may be in order. In 1959, just under 40 million persons, representing 22 percent of the noninstitutional population of the United States, were living in families (including single-person units) with annual incomes below the poverty line. By 1965, the number of persons living in poverty by this county was 32.7 million—6 million fewer—and 17 percent of the noninstitutional population.[1]

The drop in the number of poor was largely a result of the increased job opportunities and higher earnings levels resulting from the favorable economic conditions of these years. As a result a larger proportion of the poor in 1965 were persons with limited earning capacity or those whom age, disability, or other factors kept out of the labor market entirely.

In 1959 of all households counted poor, 8 million were headed by a man and 5½ million by a woman; by 1965 the number of poor households headed by a man had dropped to 6 million, but those headed by a woman remained almost unchanged. And although there were now only five households in poverty for every six in 1959, the number of one or two person families with an aged head remained as it had been, close to 4 million. Indeed, despite improvement, accounted for in large measures by the increasing number of aged drawing OASDI, persons aged 65 and over were still the most poverty-stricken group in the Nation.

In 1959, 37 percent of persons 65 and over were living in poverty, compared with 21 percent of all other age groups. Six years later, the poverty rates were 30 percent for the aged and 16 percent for all others.

A majority of the aged live alone or with just one other person. In 1965 two out of five households consisting of one aged person or an elderly couple fell below the poverty line, compared with but one in seven of all other households. The families of the aged generally have lower incomes than younger households of the same size because they are less likely to include a steady earner, and because the public programs which help many of the aged generally pay less than the earnings they are intended to replace. On the average aged couples or persons living alone must get along on less than half the money income available to a young couple or single person—a difference greater than any possible differential in living requirements.

The fact that aged men and women are less likely to work regularly than younger persons is the main reason why poverty is so much more prevalent among the aged. When families are matched by work experience and sex of the head, aged families are not so much worse off than others. For example although the poverty rate for all aged men's families is twice that of younger ones, when the head works full time the year round the rate of poverty among the aged is only 50 percent greater than among others. And indeed when the head does not work at all the average aged family will do better than a corresponding younger family because of the social security and other public support programs more readily available to older people. Among the families of men who did not work at all in 1965, 25 percent of the aged were in poverty compared with 35 percent when the head was 55–64, and 42 percent if he was under 55.

The role of social security and other public programs in ameliorating poverty is quite evident also in the situation of families headed by a women. Because a woman responsible for a family cannot work as readily as a man and will earn less when she does, the families of women are generally much poorer than men's families. But by age 65 when most men heading a family are not working regularly either, the economic gap between the man's and woman's family lessons. With a head under 55 a woman's family is five times likely to be poor as a man's; between 55 and 64, the woman's family is two and a fourths as likely to be poor as the man's by age 65 or older, the risk of poverty for the woman's family is not quite twice that of man's, and if both are not working at all, the woman's family is no more than one and a fourth times as likely to be poor as the man's.

While the aged, the disabled, and families headed by a woman with children make up the hard-core poor, there is a substantial amount of poverty among

[1] These estimates are based on special tabulations from the Current Population Survey made by the Bureau of the Census for the SSA. The data have been published in a series of articles by Mollie Orshansky in the *Social Security Bulletin:* (See the *Social Security Bulletin* for January and July 1965, and April, May, and December 1966) and summary figures used in the Economic Report of the President and Annual Report of the Council of Economic Advisors for January 1966 and January 1967.

families headed by a man who works full-time but at low wages. In most cases, these are large families. In 1965, for example, 17 percent of the households headed by a man who worked 50 weeks or more and who had four or more children were poor as compared with 4 percent of fully employed male family heads with three or fewer children. When one counts children rather than families, the seriousness of the problem becomes evident. In spite of improvements in the last few years, there were 14.3 million children living in poverty in 1965, nearly half in a family with five or more children, about a third in families headed by a woman, but also nearly a third in families headed by a man who worked full time all year.

Just above the poverty line is a group with incomes that are still lower than what one would like to think of as an American standard of living. This near-poor group included 14.6 million persons in 1965, so that the total living below a low-income level was 47.3 million.

These counts of poverty and low income are based on the sample Current Population Survey. As such they exclude persons in institutions, many of whom are among the poorest. They also measure poverty on the basis of the total income of all related persons living together. Thus, for example, a widow who lives with her son or daughter because she does not have sufficient income to live alone will nevertheless not be counted as poor unless the total group is poor. Similarly some mother-child families who share quarters with relatives do not appear in the count of those living in poverty even thought they could not get along on their own. The number of such "hidden poor" is significant, particularly among the aged. Because people in our society value highly the opportunity for independent living, it is useful to measure poverty or low income also on the basis of the income of the immediate family—an individual, a couple, or a couple and their children. Taking cognizance of the number of such persons whose own resources are insufficient but who escape poverty by living with relatives whose combining income is adequate for all would raise the poverty roster by another 2.8 million persons, of whom 1.7 million are at least 65 years old.

DRAWING THE POVERTY LINE

Poverty has many facets, not all reducible to money. Even in such terms alone, it will not be possible to obtain unanimous consent to a list of goods and services that make up the *sine qua non* and the dollars it takes to buy them. The difficulty is compounded in a country such as ours, which has long since passed the stage of struggle for sheer survival.

In many parts of the world, the overriding concern for a majority of the populace every day is still "Can I live?" For the United States as a society, it is no longer whether but how. Although by the levels of living prevailing elsewhere, some of the poor in this country might be well-to-do, no one here today would settle for mere subsistence as the just due for himself or his neighbor, and even the poorest may claim more than bread. Yet as yesterday's luxuries become tomorrow's necessities, who can define for today how much is enough? And in a society that equates economic well-being with earnings, what is the floor for those whose earning capacity is limited or absent altogether, as it is for aged persons and children?

In one sense, the difficulty of defining poverty is the price of our success story, the symbol of a society in which life has long since ceased to be a struggle just to stay alive. It signalizes the conviction that for each of us, as for all of us, this can now be so. It also means that few will be satisfied with only a minimum when so many others have so much.

It is perhaps more difficult to set a standard for poverty as a public issue than for other purposes, because in the final analysis such a procedure implies how much of our public funds and energies we wish to commit. There is not and there cannot be a uniform standard by which it can be stated unequivocally who is poor and who is not. Moreover, the means by which one may arrive at a rigorous determination of need for a specific family in a particular situation are not available for assessing economic well-being in the aggregate. Almost inevitably, lacking a case-by-case approach, a criterion applied across the board will fail to identify as poor some who are, or count as poor others who, all things considered, are, in fact, not needy. But if it is not possible to say without reservation how much is enough, it ought to be possible to say how much, on an average, is almost surely too little. Even more important than the level at which we peg the concept of too little—particularly if the findings are to relate to public action—is the proviso that the measure used depict at least roughly an equiva-

lent degree of need or relative adequacy for households of different size and composition.

Available standards for food adequacy

Despite the Nation's technological and social advance or, perhaps because of it, there is no generally accepted standard of adequacy for essentials of living except food. Even for food, social conscience and custom dictate that there be not only sufficient quantity but sufficient variety to meet recommended nutritional goals and conform to customry eating patterns. Calories alone will not be enough.

The food plans priced by the U.S. Department of Agriculture for nonfarm families today include both the low-cost one well known to welfare agencies and one at a newer economy level which costs about one-fourth less and is designed for short-term use when funds are extremely low. Most families spend considerably more. In 1955, the latest year for which we have details, only one-tenth of all nonfarm families spent less than the plan calls for. Today, 12 years later, the number with such meager food outlays is not doubt even fewer. If a family follows this plan exactly, adequate nutrition is attainable, but in practice nearly half the families that spent so little fell far short: of families spending at this rate in 1955, over 40 percent had diets that provided less than two-thirds the minimum requirements of one or more nutrients.

The SSA poverty index: the rationale

The Social Security Administration poverty index is an attempt to specify the minimum amount required to support an average family of given composition at the lowest level consistent with standards of living prevailing in this country. At best, it can stipulate only the income at which an acceptable level of consumption may on the average be possible, not necessarily plausible. Useful as a broad gauge, it cannot be applied automatically to each individual family without additional study. Not even as a screening device ought it be considered to yield an exact count of the poor in absolute numbers; but it can delineate the relative extent of privation among defined population groups and serve to outline a target for action programs.

To ensure this it is essential that the incomes selected represent equally well large families and small, children as well as grownups, and insofar as possible, families living on farms as well as in cities. The latter relationship, indeed, is but an extreme variant of the need to approximate the differences in cost of living from one place to another, a problem perennially difficult to resolve.

Failing a market basket to demarcate the line below which deprivation is almost inevitable and above which a limited measure of adequacy is at least possible, an adaptation was made of a principle most of us learn by heart: as income increases, families spend more dollars for food, but this larger amount takes a smaller share of income, leaving proportionately more money for other things. Accordingly, a low percentage of income going for food can be equated with prosperity and a high percentage with privation. Economists looking for a quick way to assess relative well-being of dissimilar groups have long resorted to this device.

This procedure was followed but with an important modification. It was assumed that equivalent levels of adequacy were reached only when the proportion of income required to purchase an adequate diet was identical. The fact that in practice large families often seem to spend more of their income on food turns out, on analysis, to come about only because, on the average, large families, particularly those with several children, have lower incomes than small families.

The index level

The procedure had the important merit that for food, a measure of adequacy is available in the Agriculture Department food plans, whereas adequacy standards for other categories of family living are not.

The starting point for the SSA poverty index is the amount of money needed to purchase the food for a minimum adequate diet as determined by the Department of Agriculture. The food budget is the lowest that could be devised to supply all essential nutrients using foods readily purchasable in the U.S. market (with customary regional variations). The poverty line is then calculated at three times the food budget (slightly smaller proportions for one- and two-person families) on the assumption—derived from studies of consumers—that a family that has to spend a larger proportion of its income on food will be living at a very inadequate level. The food budgets and the derivative poverty income cutoff points are estimated in detail for families of differing size and composition

(62 separate family types) with a farm-nonfarm differential for each type. This variation of the poverty measure in relation to family size and age of members is its most important distinguishing characteristic.

Because the level of living implied by the poverty index is lower than we think most people would regard as an appropriate measure of adequacy of income for retired persons or disabled workers and their families or widows and children, we have also developed a slightly higher index. We call this the low-income index and it is definitely low income.

The revised BLS minimum but adequate budget, when it is completed in the next few months, will almost certainly be significantly higher. For food, for example, it uses the moderate-cost food plan which costs about a third more than the low-cost plan and nearly three-fourths more than the economy plan which is the core of the SSA poverty index.

Varying the index for family type

The SSA poverty threshold is set separately for 124 different types of families according to the sex of the head, the total number of adults and children under 18, and whether or not they live on a farm. The poverty criteria have been computed at two levels, one related to the price of the familiar Low-Cost Food Plan of the U.S. Department of Agriculture, and another to a more restricted Economy Plan.

On the basis of observed household food practices, it was decided that for families of three or more, the total cost for food (as suggested by the economy plan) should take no more than one-third of family funds. Two-person families would be expected to spend no more than 27 percent of income for their food, and one-person units were assumed to need 80 percent the income of a two-person family.

The actual cost of the food plan per person differs with family size—a little more in small families than in large—as well as with the age and sex of the members. The overall economy plan cost per person, assuming four to a household, was estimated at $4.60 a week for January 1964. For an average four-person family, using the household composition prototypes we worked out, the food costs at the poverty line in 1963 came to 70 cents a person a day, or 23 cents a meal. All other items were supposed to be furnished for twice this amount, or $1.40 a day. These, it will be remembered, are the amounts that families above the poverty line were assumed to be able to spend. No allowance was made for any meals away from home, for between-meal snacks, or food for guests. All such extras must come out of the same food money, or out of the limited funds available for other things. In large families, which generally include more children, the amounts allowed per person were less. For all types of four-person nonfarm families averaged together the poverty criteria or income cutoff points averaged $3,130 for 1963 and $3,200 for 1965.

A study in 1960–61 revealed that nonfarm families by then spent, on the average, 23.5 percent of aggregate income for food. Actually, however, it was only families with incomes of $6,000 or more whose average food costs were in this range. With incomes of $2,000–3,000, families of two or more were devoting a third of income to food—the ratio we assumed for our index. Families in this income class, averaging just over three persons, reported an outlay for all food almost identical with the cost of the economy plan in 1963 or 1964 assuming four to a family. At this rate, the critical income for such a family would be $3,150, compared with the $3,130 derived *a priori*. At current prices, incomes of this magnitude hardly provide for riotous living.

A budget for farm families

The food-plan quantities are priced only for nonfarm families. In setting the poverty line for farm families it was necessary to determine for them how much on an average would be purchased and how much homegrown. In the absence of information to the contrary, the food-income relationship was given the same significance for farm as for nonfarm families in connoting income adequacy. Indeed, farm families in 1955 spent a third of net money income for purchased food, the same as other families, but their purchases represented only 60 percent of the retail value of all food they used. Home production obviously had declined since 1955, but the magnitude of the change was not yet known. With no more recent information on the level of home production—an important cost element for the farm household, it was assumed that the average farm family in 1963 would still obtain 40 percent of its food requirement from the home farm, and therefore the poverty line was set at 60 percent that for a nonfarm family.

Subsequently, information from a 1961 food consumption study showed that by 1961 home food production had dropped to no more than 31 percent the total value of food used by farm families. It would seem more appropriate, then, to peg the income required by a farm family at the poverty line at about 70 percent of the equivalent nonfarm figure rather than 60 percent and the procedure was amended to correspond. Among the farm families studied in 1961, average expenditure for food represented 20 percent of money income. Families with $1,000–2,000 averaged 35 percent, and those in the next higher income class 28 percent. Food purchases by families spending 33 percent of income were estimated by interpolation at $3.62 per person per week, with $3.13 going for food at home. This figure represents 69 percent of the amount spent by the nonfarm families devoting the same proportion of income to food.

For farm families spending this way, the average family size was the same as for the parallel nonfarm families (3.1), and family income averaged $1,838, or 71 percent that of the nonfarm families. For the year 1963 the incidence of poverty among farm households increased by about 733,000 persons when the higher income cutoff point is used and raised from 1 in 11 to 1 in 9 the proportion of the poor who lived on a farm.

By way of caution about the general applicability of the farm-nonfarm ratio developed, it must be recognized that the manner in which the Bureau of the Census obtains its income data tends to understate farm income and therefore to overstate poverty to a greater degree for farm families than for nonfarm families. The farm family, asked for a quick estimate of its income (including operating expenses), is likely to assign al utilities, transportation, and shelter costs to the farm side of the account rather than prorate a share as the cost of family living. In approximating farm-nonfarm equivalence on the basis of Census income distributions—which must provide the basis for the poverty index, one may therefore postulate a lower ratio of farm to nonfarm money income than would apply if the income data were obtained by methods similar to those of the Department of Agriculture household expenditure studies.

HOW ADEQUATE IS THE STANDARD?

The measure of poverty thus developed is arbitrary. To be sure, it applies only in America. Few could call it too high. Many might find it too low. Assuming the homemaker is a good manager and has the time and skill to shop wisely, she must prepare nutritious, palatable meals on a budget that for herself, a husband, and two young children—an average family—would even today come to about 70 cents a day per person.

For a meal all four of them ate together, she could spend on the average only 95 cents, and to stay within her budget she must allow no more a day than a pound of meat, poultry, or fish altogether, barely enough for one small serving for each family member at one of the three meals. Eggs could fill out her family fare only to a limited degree because the plan allows less than two dozen a week for all uses in cooking and at the table, not even one to a person a day. And any food extras, such as milk at school for the children, or the coffee her husband might buy to supplement the lunch he carries to work, have to come out of the same food money or compete with the limited funds available for rent, clothing, medical care, and all other expenses. Studies indicate that, on the average, family members eating a meal away from home spend twice as much as the homemaker would spend for preparing one for them at home. The 20–25 cents allowed for a meal at home in the economy plan would not buy much even in the way of supplementation.

There is some evidence that families with very low income, particularly large families, cut their food bills below the economy plan level—a level at which a nutritionally good diet, though possible, is hard to achieve. Indeed, a study of beneficiaries of old-age, survivors, and disability insurance—limited to 1- or 2-person families—found that only about 10 percent of those spending less than the low-cost plan (priced about a third higher than the economy plan) had meals furnishing the full recommended amounts of essential nutrients. Not more than 40 percent had even as much as two-thirds the amounts recommended. Only when food expenditures were as high as those in the low-cost plan, or better, did 90 percent of the diets include two-thirds of the recommended allowances of the nutrients, and 60 percent meet them in full. Few housewives with greater resources—income and other—than most poor families have at their disposal could do better. Many might not do as well.

WHAT'S WRONG WITH THE POVERTY INDEX?

Much has been said and more will be about the limitations of the poverty index, what might be called the poverty of the poverty line. It refers only to current income—to some this implies a weakness in that it ignores assets and other money receipts. It makes no adjustment for income in kind except for income from farming, and to be sure there are those who don't like that one either. To some it signifies a failure to allow for the temporary component of income because it ignores year-to-year change in income. And to others it is suspect because it ignores life's nonmonetary satisfactions and the multiple ills afflicting the poor in addition to income insufficiency.

All these criticisms have merit, but let us make our bow to the last first. If money alone will not solve the problems of poverty it is still true that without money nothing else will avail much either. Mathematically it falls in the category of necessary though not sufficient conditions. Pragmatically it is undoubtedly true that the persons who declaim loudest that "money isn't everything" are those who already have some.

Ignoring assets is a more serious defect, yet in the only income data available on a regular basis, namely those collected by the Census Bureau from the Current Population Survey Samples, assets do not appear. On the other hand save among the aged, we find few poor households with generally substantial assets. The data collected for the OEO in 1966 should help us know better where asset holdings are concentrated, but it will still be difficult to devise a satisfactory method of handling them. It is not easy to see how to take due account of assets in a poverty criterion without discouraging savings, when as may well happen the poverty threshold developed as a statistical tool becomes a program eligibility criterion. Moreover, some forms of assets are not regarded as negotiable assets by their holders. Life insurance, for example, represents to many aged persons not savings but provision for their funeral costs.

With respect to the temporary income thesis we know little about income flow for cohorts of families and how it affects conusmption. Farmers and other entrepreneurial families, perhaps more than others, are subject to the hazards of paper poverty because they may use income in one year to enhance their business position and improve income prospects for the future. And the voluntary poverty assumed by the graduate student while completing his education is a familiar phenomenon. But while we speculate on those who are poor only temporarily, we might give a moment to those who are only momentarily not poor: in many households the interruption of income because of unemployment or other reasons may prevent adequate planning of spending and inhibit needed consumption even though on a recap basis total income for the year comes above the poverty line. In other words, irregularity of income and uncertainty as to its amount may be as much of a hazard to economic and social well-being as low income.

As for nonmoney income, the bulk of all nonmoney transfers—if one includes not only free medical care and food stamps but also fringe benefits to workers, health insurance premiums, expense accounts, vacation allowances, stock options, free or reduced tuition, commodity discounts, and the like—may well, like many of the income tax benefits, go to the nonpoor rather than the poor. The full effect of incorporating these into the income distribution might be to skew it even more than now with a resultant upping of the poverty line.

At first blush the value of consumption suggests itself as superior to income for a measure of poverty status. The point is, of course, that income standards are presumed to be measures of consumption potential. Shifting to consumption rather than income as the reference unit does not eliminate the problems of measurement and definition, but merely exchanges new ones for old. Currently data on consumer expenditures supply estimates of total purchase commitments rather than cash outlays during the year. Our household expenditures survey schedules have not yet accepted the buy now-pay later dictum as a way of life. What is more, we shall still face the need for assessing the value of goods received without direct outlay—will they be given a retail value (assuming the family knows it) although often they do not replace any item the family must buy? Should they in the case of a farm or other business family be valued at the income foregone which is the price they could be sold for? Should the value of homeownership be set at a fair return for investment or more realistically at the saving, if any, over what a neighbor in similar circumstances pays for rent. And how shall we determine the appropriate replacement (or purchase) rates in the absence of existing inventories which can portray the real consumption potential?

There remains a whole host of problems of arbitrary selection—as for example the appropriate food pattern to price. This in itself largely determines the level

of living defined for the rest of the budget, be it developed category by category, or in the absence of objective measures derived in one operation from an assumed income-food relationship. In any case, the presumed relationship between food and other family living items or between food and income is critical. It is worth mentioning again that the selection will not be value-free. With a fixed set of numbers relating to food and income, and with the same basic assumption about the utility of food expenditures as a thumbnail guide to adequacy, Mrs. Rose Friedman derived a standard and consequently a poverty tally considerably lower than that of the SSA, whereas Dr. Alan Haber's assumptions based on the same data resulted in a higher one. (The dollar criteria for a four-person family ranged from $2,200 according to Friedman to $3,400 according to Haber, for the year when the SSA level averaged out at $3,130.) Yet with basically different assumptions about how to determine equivalent need of households of different sizes, the BLS, SSA, IDA, Mrs. Friendman, and Dr. Eleanor Snyder arrived at much the same scale of relatives. The method used by the SSA in effect applies to the adult male nutrition unit as expressed in the food plans, but assumes probably incorrectly that the sum total for all other categories of family living can be related to age and sex of the members in the same ratio as food needs.

WHAT THE POVERTY STATISTICS DO

Because family size scales arrived at in such varied ways are pretty much of a constant, any poverty index which incorporates them probably could be counted on to give us a reasonable ordering of groups in the population by their degree of vulnerability. Thus the direction of differences between them and their order of magnitude could be useful even if the absolute numbers admittedly are weak. They could help us pinpoint action arenas and evaluate progress. They could illumine the special problems of the South, the slums versus the suburbs, and the difficulties faced by minority groups. They would be improved or at least achieve more stature with some time spent studying cost-of-living differentials, assuming there are any, from place to place and time to time. There must be a framework for adjusting the poverty line, however determined, for change over time in productivity and the general level of economic activity, and even for price change. Our knowledge of how consumers, poor or otherwise, adjust to rising prices and how they trade off one category of family living for another is still by gosh and by gum.

Perhaps the worst use of the poverty numbers is their most needed one. The present poverty lines, developed for gross measurement and admittedly imperfect, are being used as antipoverty guidelines pending results of further and better research. It is necessary to have an operational procedure, but to have it applied to families on an individual basis for program eligibility is at once the simplest and the least defensible extension. The range of individual need cannot be encompassed at a single stroke. All of the limitations discussed with reference to the poverty standard for group assessment are increased beyond measure when applied unaltered to a specific case. Even with separate poverty criteria for each of 124 different family types it was necessary or at least practical to assume only one set of age-sex prototypes for any given family group. Thus the poverty line assumes a family of five with three children always will need more than a family of four with two. Yet everyone knows that three little tots under 6 will not require as much food and clothing as two husky teenagers. In practice, the poverty lines used for program eligibility are abbreviated even further, because only one average income threshold is used for any family size rather than varying it by number of children and number of adults. On the one hand this seems inadequate; on the other hand a constant complaint about Public Assistance procedures is that too much effort is expended by social workers to determine eligibility tailored-to-measure.

As time goes on and our expertise improves along with our exposure, some of the difficulties now confronting us may disappear or at least become so familiar that they no longer disturb us. We can hope for the poverty criteria as for other social problems now confronting us that technology and methods of analysis will grow to meet the need. Over the years we have become sophisticated enough to aggregate data so as to be able to generalize. But now at least in economics and related aspects of the social scene we are become so much more sophisticated as to see the need to disaggregate. It is now not so much the central tendency we seek to isolate but the deviants from it. And our techniques and procedures will have to adjust accordingly.

WHAT THE POVERTY STATISTICS DON'T DO

Our numbers cannot tell us yet some of the most important things we need to learn—about the legacy of poverty—the chain of despair we suspect is transmitted hand-to-mouth so that indeed the sins of the fathers are visited on the children, even unto the third generation. They do not tell us whether it is best to change the world in which our poor live or help them conform to the one they're in. They give us little but faith to lean on in predicting lasting or side-effects of one anti-poverty program versus another. Nor can they help us know in advance whether it is best to concentrate on the adult or the child, and perhaps to find later that what advances one may hinder the other. Our impulse as technicians and as citizens is to call for more and better data. Yet the real difficulty is not that we don't have enough numbers but what we do with those we have. Much of the controversy over the statistics of poverty turns not on the poverty statistics but over the different meanings attributed to them.

It is science we cry for but in reality it is philosophy we crave, with the numbers serving only to rationalize or justify what has already been decided. This is probably as it should be, provided we are humble enough to acknowledge it and don't drape ourselves in the mantle of virtue as we proffer each our own particular brand of eternal truth. In these days of consumer awareness, however, we must all as producers label our packages complete with a list of ingredients and with instructions for use. But in our role as consumers we should follow the instructions. And if as often happens with a convenience product we prefer ways of our own to use it we ought not to cry "foul" if it then does not turn out like the picture.

We wait for the answer and work while we wait. In the meantime, our poverty statistics, weak though they be, show us where problems are even if they cannot always reveal exact dimensions. We have yet no way to define poverty to everyone's satisfaction, not even our own. But we have not yet reached the point where we cannot act for lack of the right numbers. Unlike some other calculations those relating to poverty have no intrinsic essence of their own. They exist only in order to help us make them disappear from the scene. And so if I may plagiarize from myself, remedial actions need not wait on statistical refinements. If as individuals and as a Nation we can think bold solutions and dream big dreams we may be able to wipe out the scourge of poverty even before we can agree on how to measure it.

RESEARCH AND STATISTICS NOTE

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE,
SOCIAL SECURITY ADMINISTRATION,
OFFICE OF RESEARCH AND STATISTICS,
February 16, 1967.

(Note No. 5—1967)

THE POOR IN 1965 AND TRENDS, 1959–65

The five tables presented in this Note continue the series of estimates of trends in the poverty and low-income status of the population developed by the SSA. Earlier data and descriptions of the indexes can be found in a series of articles by Mollie Orshansky in the *Social Security Bulletin* for January and July 1965 and April and May 1966.

As described in detail in those articles, the SSA variable poverty and low-income indexes are based on the economy and low-cost food budgets developed by the Department of Agriculture and the general assumption that families are poor if these amounts represent more than one-third of their total income (more than 27 percent for couples and somewhat less for single persons for whom economies of scale in housing are not applicable). The poverty and low-income cutoff points are adjusted each year to reflect changes in food prices. Table 1 shows the weighted averages of income criteria for families of different size and composition that were used in estimating the incidence of poverty and low-income status in 1965.

Tables 2–5 present 1965 estimates of poverty and low-income status for households and the individuals in those households. Additional data for 1965 will be published in subsequent R&S Notes. These basic tables are made available now for others who wish to use them.

TABLE 1.—*Weighted average of poverty and low-income criteria* [1] *for families of different composition by household size, sex of head, and farm or nonfarm residence, March 1966*

Number of family members	Weighted average of incomes at poverty level						Weighted average of incomes at low-income level					
	Nonfarm			Farm			Nonfarm			Farm		
	Total	Male head	Female head	Total	Male head	Female head	Total	Male head	Female head	Total	Male head	Female head
1 member	$1,570	$1,635	$1,530	$1,110	$1,145	$1,070	$1,890	$1,980	$1,840	$1,340	$1,385	$1,290
Head under age 65	1,615	1,685	1,560	1,140	1,180	1,090	1,950	2,040	1,880	1,380	1,425	1,315
Head aged 65 or over	1,500	1,515	1,495	1,055	1,060	1,045	1,085	1,835	1,790	1,265	1,285	1,255
2 members	2,030	2,040	1,975	1,415	1,420	1,365	2,725	2,745	2,610	1,905	1,910	1,800
Head under age 65	2,100	2,110	2,025	1,475	1,480	1,410	2,810	2,835	2,665	1,980	1,985	1,860
Head aged 65 or over	1,890	1,895	1,880	1,325	1,325	1,325	2,545	2,550	2,500	1,785	1,785	1,760
3 members	2,495	2,505	2,405	1,740	1,745	1,660	3,265	3,275	3,175	2,280	2,285	2,210
4 members	3,200	3,200	3,180	2,250	2,255	2,205	4,145	4,150	4,050	2,920	2,920	2,825
5 members	3,765	3,770	3,730	2,640	2,640	2,640	4,835	4,845	4,730	3,395	3,395	3,370
6 members	4,235	4,235	4,220	2,970	2,970	3,055	5,440	5,445	5,345	3,820	3,820	3,860
7 or more members	5,205	5,215	5,090	3,630	3,635	3,560	6,615	6,630	6,455	4,610	4,615	4,515

[1] Required income in 1965 according to Social Security Administration poverty or low-income index for a family of given size and composition. Family income criteria weighted together in accordance with percentage distribution of total units by number of related children and sex of head, as of Current Population Survey, March 1966.

For detailed description of the Social Security Administration measures of poverty and low income and their rationale, see the Social Security Bulletin for January 1965 (pp. 5–11) and July 1965 (pp. 3–10).

TABLE 2.—*Trends in poverty and low-income status, 1959–65: Number and percent of noninstitutional population who were poor or near poor*

Type of unit and income level	The poverty and low-income roster							The incidence of poverty and low-income status						
	1959	1960	1961	1962	1963	1964	1965	1959	1960	1961	1962	1963	1964	1965
	Persons poor or near poor [1] (in millions)							Percent of persons poor or near poor [2]						
Total with low income	54.7	54.8	52.9	52.8	51.0	49.8	47.5	31.0	30.5	29.3	28.6	27.3	26.3	24.8
Poor	38.9	40.1	38.1	37.0	35.3	34.1	32.7	22.1	22.3	21.1	20.1	18.9	18.0	17.1
Near poor	15.8	14.7	14.8	15.8	15.7	15.7	14.8	9.0	8.2	8.2	8.5	8.4	8.3	7.7
Unrelated individuals	5.6	5.5	5.6	5.6	5.6	5.8	5.6	52.5	50.5	49.7	50.3	49.8	47.8	46.5
Poor	5.1	5.1	5.0	4.9	4.9	5.1	4.8	47.4	46.5	44.7	44.2	43.9	42.0	39.8
Near poor	.5	.4	.6	.7	.7	.7	.8	5.1	4.0	5.0	6.1	5.9	5.8	6.7
Under age 65	2.9	2.8	2.7	2.7	2.6	2.6	2.5	41.4	39.2	38.0	38.6	37.6	35.5	33.7
Poor	2.6	2.6	2.4	2.4	2.4	2.3	2.1	36.8	36.1	33.9	34.5	34.2	31.2	28.7
Near poor	.3	.2	.3	.3	.2	.3	.4	4.6	3.1	4.1	4.1	3.4	4.3	5.0
Aged 65 or over	2.7	2.6	2.8	2.9	2.9	3.2	3.1	74.2	73.0	71.0	69.4	69.6	67.4	66.9
Poor	2.5	2.4	2.5	2.5	2.5	2.8	2.7	68.1	67.1	64.4	60.0	59.7	59.3	57.5
Near poor	.2	.2	.3	.4	.4	.4	.4	6.1	5.9	6.6	9.4	9.9	8.1	9.4
Persons in families	49.1	49.3	47.4	47.3	45.5	44.0	41.9	29.7	29.3	27.8	27.2	25.8	24.8	23.4
Poor	33.9	35.0	33.1	32.2	30.4	28.9	27.9	20.4	20.8	19.4	18.5	17.3	16.3	15.6
Near poor	15.2	14.3	14.3	15.1	15.1	15.1	14.0	9.2	8.5	8.4	8.7	8.5	8.5	7.8
With no children under age 18	9.3	9.3	8.8	8.4	8.6	8.0	8.0	22.3	21.9	20.4	19.5	19.6	17.9	17.3
Poor	6.5	6.5	6.0	5.4	5.5	4.9	4.8	15.5	15.3	13.8	12.6	12.5	11.1	10.4
Near poor	2.8	2.8	2.8	3.0	3.1	3.1	3.2	6.8	6.6	6.6	6.9	7.1	6.8	6.8
With children under age 18	39.8	40.1	38.7	38.8	36.8	36.2	33.9	32.1	31.7	30.3	29.8	27.9	27.3	25.5
Poor	27.4	28.6	27.2	26.7	24.9	24.0	23.1	22.1	22.6	21.3	20.5	18.9	18.1	17.3
Near poor	12.4	11.5	11.5	12.1	11.9	12.2	10.8	10.0	9.1	9.0	9.3	9.0	9.2	8.1
Adults	16.1	16.5	16.2	16.0	14.7	14.5	21.4	26.9	26.8	26.3	25.5	23.5	23.0	19.6
Poor	10.7	11.4	11.2	10.8	9.6	9.2	13.6	17.9	18.5	18.2	17.2	15.4	14.6	12.4
Near poor	5.4	5.1	5.0	5.2	5.1	5.3	7.8	9.0	8.3	8.1	8.3	8.1	8.4	7.2
Children under age 18	23.7	23.6	22.5	22.8	22.1	21.7	20.4	37.1	36.2	34.1	33.8	32.1	31.3	29.3
Poor	16.7	17.2	16.0	15.9	15.3	14.8	14.3	26.1	26.4	24.2	23.6	22.2	21.3	20.5
Near poor	7.0	6.4	6.5	6.9	6.8	6.9	6.1	11.0	9.8	9.8	10.2	9.9	9.9	8.8

	Households poor or near poor [1] (in millions)							Percent of households poor or near poor [2]						
Total with low income [3]	17.7	17.3	17.1	17.0	16.7	16.4	15.8	31.7	30.7	29.7	29.3	28.5	27.4	26.1
Poor	13.4	13.4	13.0	12.6	12.3	11.9	11.5	24.0	23.7	22.6	21.8	21.0	19.8	19.0
Near poor	4.3	3.9	4.1	4.4	4.4	4.5	4.3	7.7	7.0	7.1	7.5	7.5	7.6	7.2
Families of 2 or more	12.0	11.8	11.5	11.5	11.1	10.6	10.2	26.7	26.0	24.9	24.4	23.5	22.2	21.0
Poor	8.3	8.3	8.0	7.8	7.4	6.8	6.6	18.4	18.3	17.3	16.5	15.5	14.2	13.7
Near poor	3.7	3.5	3.5	3.7	3.7	3.8	3.5	8.3	7.7	7.6	7.9	8.0	8.0	7.3
With no children under age 18	4.3	4.3	4.0	3.9	3.9	3.7	3.7	23.6	23.2	21.5	20.7	20.7	19.1	18.5
Poor	3.0	3.0	2.7	2.5	2.5	2.3	2.2	16.4	16.3	14.5	13.4	13.2	11.7	11.2
Near poor	1.3	1.3	1.3	1.4	1.4	1.4	1.5	7.3	6.9	7.0	7.3	7.5	7.4	7.3
With children under age 18	7.7	7.5	7.5	7.5	7.2	6.9	6.4	28.8	27.7	28.0	26.9	25.3	24.5	22.9
Poor	5.3	5.3	5.3	5.2	4.9	4.5	4.4	19.7	19.5	19.2	18.6	17.1	16.0	15.6
Near poor	2.4	2.2	2.2	2.3	2.3	2.4	2.0	9.1	8.2	8.8	8.3	8?2	8.5	7.3
1 to 2 children	3.7	3.4	3.4	3.3	3.1	3.0	2.8	21.1	20.3	20.3	18.9	18.0	17.7	16.5
Poor	2.5	2.3	2.4	2.3	2.1	1.8	1.9	14.3	13.6	14.2	13.2	12.0	10.8	10.9
Near poor	1.2	1.1	1.0	1.0	1.0	1.2	1.0	6.9	6.7	6.1	5.7	6.0	6.9	5.7
3 to 4 children	2.7	2.6	2.6	2.6	2.5	2.3	2.2	35.3	33.0	31.7	31.3	29.8	28.0	26.7
Poor	1.8	1.8	1.8	1.7	1.6	1.5	1.5	23.3	22.9	21.9	20.2	19.5	18.8	18.2
Near poor	.9	.8	.8	.9	.9	.8	.7	12.0	10.1	9.8	11.1	10.3	9.2	8.5
5 or more children	1.5	1.5	1.5	1.7	1.5	1.6	1.4	67.9	68.2	63.0	65.3	5.88	56.6	52.0
Poor	1.1	1.2	1.1	1.3	1.1	1.1	1.1	51.2	54.2	47.1	48.7	43.5	41.0	38.5
Near poor	.4	.3	.4	.4	.4	.4	.4	16.7	14.0	15.9	16.6	15.3	15.6	13.6

[1] Income for the specified year, of family unit or unrelated individual below the Social Security Administration index at the poverty level by family size and sex of head or, alternatively, at the somewhat higher low-income level (see the Social Security Bulletin, April 1966 issue, pp. 20-21). The SSA index has been adjusted for price changes during the period.

[2] The percent that poor or near poor persons (or families) are of total number of persons (or families) in each category in the noninstitutional population. All persons in institutions and children under age 14 who live with a family of nonrelatives are not represented in the low-income roster because income data are not collected for inmates of institutions or unrelated individuals under age 14. As of March 1966, there were about 346,000 such children and 2,133,000 persons of all ages in institutions.

[3] Includes unrelated individuals shown separately above.

Source: Derived from special tabulations by the Bureau of the Census from the Current Population Survey for March 1960 to 1966.

TABLE 3.—*Poverty and low-income status of noninstitutional population in 1965: Number and percent of persons in households with 1965 income below SSA poverty or low-income levels, by age, race, and family status*

[Number in thousands]

Age and family status	Total persons					White					Nonwhite				
	Total [1]	In households with low income				Total [1]	In households with low income				Total [1]	In households with low income			
		Poor		Poor and near poor [2]			Poor		Poor and near poor [2]			Poor		Poor and near poor [2]	
		Number	Percent	Number	Percent		Number	Percent	Number	Percent		Number	Percent	Number	Percent
All persons	191, 535	32, 669	17. 1	47, 256	24. 7	168, 853	22, 362	13. 2	34, 250	20. 3	22, 681	10, 308	45. 4	13, 007	57. 3
Unrelated individuals	12, 132	4, 830	39. 8	5, 643	46. 5	10, 476	4, 007	38. 2	4, 708	44. 9	1, 655	824	49. 8	936	56. 6
Members of family units	179, 403	27, 839	15. 5	41, 613	23. 2	158, 377	18, 355	11. 6	29, 542	18. 7	21, 026	9, 484	45. 1	12, 071	57. 4
Children under age 18 [3]	69, 639	14, 282	20. 5	20, 373	29. 3	59, 522	8, 652	14. 5	13, 477	22. 6	10, 117	5, 630	55. 6	6, 896	68. 2
Under 6	24, 163	5, 252	21. 7	7, 652	31. 7	20, 366	3, 121	15. 3	4, 995	24. 5	3, 798	2, 132	56. 1	2, 658	70. 0
6 to 13	31, 749	6, 630	20. 9	9, 325	29. 4	27, 276	4, 038	14. 8	6, 202	22. 7	4, 472	2, 591	57. 9	3, 121	69. 8
14 to 17	13, 728	2, 400	17. 5	3, 396	24. 7	11. 880	1, 492	12. 6	2, 278	19. 2	1, 848	908	49. 1	1, 118	60. 5
Persons aged 18 to 24 [4]	19, 669	2, 769	14. 1	4, 087	20. 8	17, 305	1, 901	11. 0	2, 922	16. 9	2, 364	869	36. 8	1, 164	49. 2
Unrelated individuals [5]	1, 158	465	40. 2	557	48. 1	1, 011	389	38. 5	464	45. 9	146	76	52. 1	92	63. 0
Members of family units	18, 511	2, 304	12. 4	3, 530	19. 1	16, 294	1, 512	9. 3	2, 458	15. 1	2, 218	793	35. 8	1, 072	48. 3
Head	3, 050	571	18. 7	812	26. 6	2, 713	410	15. 1	614	22. 6	338	162	47. 9	199	58. 9
Wife	4, 885	601	12. 3	966	19. 8	4, 438	456	10. 3	747	16. 3	447	145	32. 4	219	49. 0
Never-married children [6]	7, 794	857	11. 0	1, 296	16. 6	6, 794	489	7. 2	812	12. 0	1, 000	368	36. 8	483	48. 3
Other	2, 782	275	9. 9	456	16. 4	2, 349	157	6. 7	285	12. 1	433	118	27. 3	171	39. 5
Persons aged 25 to 54	67, 659	7, 700	11. 4	11, 642	17. 2	60, 382	5, 242	8. 7	8, 375	13. 9	7, 276	2, 458	33. 8	3, 271	45. 0

Unrelated individuals	3,770	822	21.8	967	25.6	3,000	553	18.4	655	21.8	770	269	34.9	312	40.5
Members of family units	63,889	6,878	10.8	10,675	16.7	57,382	4,689	8.5	7,717	13.4	6,506	2,189	33.6	2,959	45.5
Head	30,844	3,633	11.8	5,507	17.9	27,615	2,428	8.8	3,917	14.2	3,229	1,205	37.3	1,591	49.3
Wife	28,062	2,568	9.2	4,176	14.9	25,625	1,833	7.2	3,149	12.3	2,437	735	30.2	1,027	42.1
Other	4,983	677	13.6	992	19.9	4,142	428	10.3	651	15.7	840	249	29.6	341	40.6
Persons aged 55 to 64	16,923	2,640	15.6	3,758	22.2	15,392	2,072	13.5	3,011	19.6	1,529	567	37.1	746	48.8
Unrelated individuals	2,524	851	33.7	990	39.2	2,193	688	31.3	801	36.4	326	163	50.0	188	57.7
Members of family units	14,399	1,789	12.4	4,768	19.2	13,194	1,384	10.5	2,210	16.8	1,203	404	33.6	558	46.4
Head	7,490	912	12.2	1,385	18.5	6,820	683	10.0	1,069	15.7	668	228	34.1	315	47.2
Wife	5,647	671	11.9	1,092	19.3	5,231	541	10.3	914	17.5	416	130	31.2	179	43.0
Other	1,262	206	16.3	291	23.1	1,143	160	14.0	227	19.9	119	46	36.7	64	53.8
Persons aged 65 and over	17,645	5,279	29.9	7,397	41.9	16,252	4,494	27.7	6,466	39.8	1,394	786	56.4	933	66.9
Unrelated individuals	4,679	2,692	57.5	3,129	66.9	4,268	2,376	55.7	2,786	65.3	411	317	77.1	345	83.9
Members of family units	12,966	2,587	20.0	4,268	32.9	11,984	2,118	17.7	3,680	30.7	983	469	47.7	588	59.8
Head	6,895	1,521	22.1	2,446	35.5	6,348	1,245	19.6	2,103	33.1	547	276	50.5	344	62.9
Wife	3,514	749	21.3	1,285	36.6	3,338	649	19.4	1,164	34.9	176	100	56.8	120	68.2
Other	2,557	317	12.4	537	21.0	2,298	224	9.7	413	18.0	260	93	35.8	124	47.7

[1] Noninstitutional population.
[2] Families in poverty and families above poverty but below low-income index.
[3] Includes never-married own children of head and all other never-married relatives under age 18. Excludes 346,000 children under age 14 who live with a family of nonrelatives.
[4] Includes heads, wives, and other never-married relatives under age 18
[5] Includes 93,000 unrelated individuals aged 14 to 17 of whom 90,000 had incomes below SSA poverty level in 1965 and 2,000 were above poverty but below low-income level.
[6] Aged 18 to 21.

Source: Derived from special tabulations by Bureau of the Census from the Current Population Survey for March 1966.

TABLE 4.—*Poverty and low-income status in 1965 of all persons in the noninstitutional population,[1] by age and family status*

[In millions]

Age and family status	Total noninstitutional population [2]	Poverty criterion			Low income criterion		
		Poor	Nonpoor (including near poor)		Poor and near poor	Nonpoor (above low-income level)	
			Total	Hidden poor [3]		Total	Hidden poor [3]
Number of persons	191.5	32.7	158.9	2.8	47.3	144.3	2.8
Unrelated individuals	12.1	4.8	7.3	----	5.6	6.5	----
Under age 65	7.5	2.1	5.3	----	2.5	4.9	----
Aged 65 or over	4.7	2.7	2.0	----	3.1	1.6	----
Members of family units	179.4	27.8	151.6	2.8	41.6	137.8	2.8
Children under age 18 [4]	69.6	14.3	55.4	.6	20.4	49.3	.6
Own children of head (or spouse)	66.1	13.0	53.2	----	18.7	47.5	----
Other related children	3.5	1.3	2.2	.6	1.7	1.8	.6
Persons aged 18–64 [5]	96.8	11.0	85.8	.5	17.0	79.8	.5
Head	41.4	5.1	36.3	----	7.7	33.7	----
Wife	38.6	3.8	34.8	----	6.2	32.4	----
Never married children aged 18–21	7.8	.9	6.9	----	1.3	6.5	----
Own children of head (or spouse)	7.2	.7	6.5	----	1.1	6.1	----
Other related children	.6	.1	.5	----	.2	.4	----
Other relative	9.0	1.2	7.9	.5	1.7	7.3	.5
Persons aged 65 or over	13.0	2.6	10.4	1.7	4.3	8.7	1.7
Head	6.9	1.5	5.4	----	2.4	4.4	----
Wife	3.5	.7	2.8	----	1.3	2.2	----
Other relative	2.6	.3	2.2	1.7	.5	2.0	1.7

[1] Income of family unit or unrelated individual below Social Security Administration poverty index fo family size and sex of head, or, alternatively, at the low income level, roughly 30 percent higher in cost.

[2] As of March 1966, there were 2,133,000 persons in institutions, including 279,000 children under age 18; 117,000 persons aged 18–64; and 737,000 persons aged 65 or over. These persons as well as the 346,000 children under age 14 who live with a family of nonrelatives are not represented in these indexes, because income data are not normally collected for inmates of institutions or unrelated individuals under age 14.

[3] Individuals or subfamily members with own income below the poverty or low income level but living in a family above that level. A subfamily represents a married couple with or without children or a parent and 1 or more children residing in a family as relatives of the head.

[4] Represents never married children. Excludes 625,000 children, 279,000 in institutions and 346,000 under under age 14 in households of nonrelatives, all of whom are likely to be poor.

[5] Includes any persons under age 18 living in families as heads, wives, or never married children.

Source: Derived from special tabulations by the Bureau of the Census from the current population survey or March 1966.

TABLE 5.—*Living arrangements of persons aged 65 and over in the noninstitutional population, by sex and poverty status in 1966*

Family status	Number (in millions)			Percentage distribution		
	Total	In poor households [1]	In nonpoor households	Total	In poor households [1]	In nonpoor households
Persons aged 65 or over	17.6	5.3	12.4	100.0	100.0	100.0
Living alone [2]	4.7	2.7	2.0	26.5	51.0	16.1
Living in family units	13.0	2.6	10.4	73.5	49.0	83.9
Head	6.9	1.5	5.4	39.1	28.8	43.5
Wife	3.5	.8	2.8	19.9	14.2	22.4
Other relative	2.6	.3	2.2	14.5	6.0	18.1
Poor by own income [3]	2.0	.3	1.7	11.3	5.9	13.6
Not poor by own income	.6	(4)	.6	3.2	.1	4.5
Men	7.7	1.8	5.9	43.7	34.7	47.6
Living alone [2]	1.3	.6	.7	7.2	11.0	5.7
Living in family units	6.4	1.3	5.2	36.5	23.7	41.9
Head	5.8	1.2	4.6	32.7	22.6	37.0
Other relative of head aged 65 or over	.1	(4)	.1	.8	.4	1.0
Other relative of head under age 65	.5	(4)	.5	2.9	.8	3.9
Women	9.9	3.4	6.5	56.3	65.3	52.4
Living alone [2]	3.4	2.1	1.3	19.3	40.0	10.4
Living in family unit	6.5	1.3	5.2	37.0	25.3	42.0
Head	1.1	.3	.8	6.4	6.2	6.4
Wife, husband aged 65 or over	3.2	.7	2.5	18.3	13.0	20.5
Wife, husband under age 65	.3	.1	.2	1.6	1.2	1.8
Other relative of head aged 65 or over	.4	.1	.3	2.5	1.9	2.8
Other relative of head under age 65	1.5	.2	1.3	8.2	3.0	10.4
In household with head aged 65 or over:						
Male head	10.5	2.5	8.0	59.6	47.3	64.9
Female head	4.9	2.5	2.3	27.6	47.8	18.9
In household with head under age 65:						
Male head	1.8	.2	1.6	10.2	2.9	13.3
Female head	.5	.1	.4	2.6	2.0	2.8

[1] Income in 1965 of persons living alone or of family unit below the Social Security Administration poverty index.
[2] Includes those living with nonrelatives only.
[3] Income of other relative aged 65 or over in 1965 less than $1,500.
[4] Less than 50,000.

Source: Derived from special tabulations by the Bureau of the Census from the Current Population Survey for March 1966.

USDA HOUSEHOLD FOOD CONSUMPTION SURVEYS AND THEIR USES

By Faith Clark, Director, Consumer and Food Economics Research Division, Agricultural Research Service, U.S. Department of Agriculture

The U.S. Department of Agriculture has a continuing long-term program of research on food consumption and dietary levels of households and individuals involving nutritionists, food economists, and statisticians. The Department began making studies of food consumption of population groups at the beginning of the century when Dr. W. O. Atwater became the first head of the Office of Experiment Stations. The first nationwide food consumption survey was made in 1936–37 as a part of the Consumer Purchases Study. Since then four large-scale studies have been made in 1942, 1948, 1955 and in 1965–66. In between the nationwide studies, a number of small-scale special-purpose surveys have been made.

1965–66 Nationwide Survey

The objectives of the 1965–66 survey were to obtain information on current food consumption and also to compare the results with those of the 1955 and earlier surveys. Hence the methodology was kept quite comparable to the 1955 survey and yet we did introduce many improvements. The sample was designed to be representative of housekeeping households of one or more persons in the United States. A household was defined as housekeeping if at least one person had at least 10 meals from home food supplies during the past seven days.

Thus we excluded families who ate few meals at home the week before their interview and the institutional population of the U.S.

The 144 dots on the map represent the 144 areas from which households were selected to provide a nationwide sample (Fig. 1). The sample was designed to provide classification of data for four Census regions—Northeast, North Central, West and South—and for three urbanizations—farm, rural nonfarm, and urban—and by income groups. To assure adequate farm coverage, the farm population was over-sampled.

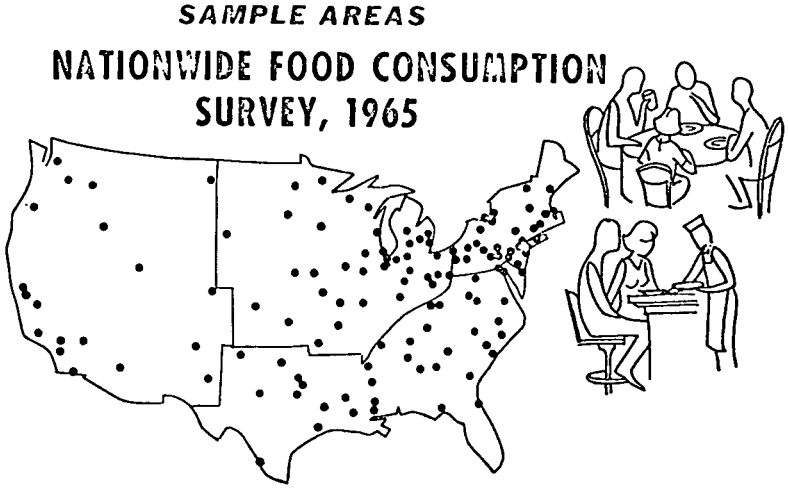


Figure 1

A new feature of the 1965–66 survey is coverage during all four seasons. The survey was designed to include 7,500 households in spring 1965 and 2,500 in the summer and fall of 1965 and the winter of 1966 (Fig. 2). In 1965 also for the first time in a nationwide dietary survey, we collected information on the food intake of individuals. About 13,000 family members of the households sampled in the spring were covered. Persons of all ages from birth on were included.

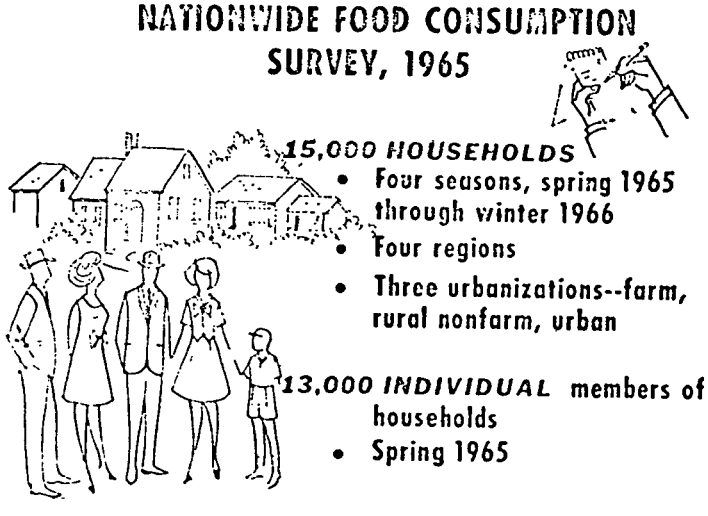


Figure 2

The information on household food consumption obtained in the 1965–66 survey included the types and amounts of foods used in the home during the seven days preceding the interview (Fig. 3). The information included a description of each food sufficient to calculate its nutrient contribution. The source also was obtained, i.e., whether it was purchased, home produced, Federally donated, or received as gift or pay, and information on the price paid for purchased food was requested. Other basic information included the age, sex, height and weight of persons eating from home food supplies with the number of meals each one had, home practices in food production, canning and freezing and the 1964 or 1965 income of the family.

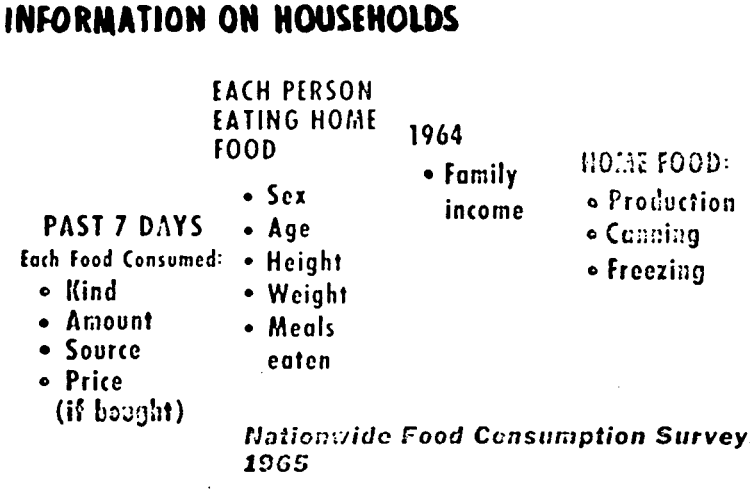


Figure 3

The information about the food intake of the individuals in the families was requested after the completion of the household questionnaire during the spring interviews. Information on individuals concerned all the foods eaten at home and away from home during the day preceding the interview (Fig. 4).

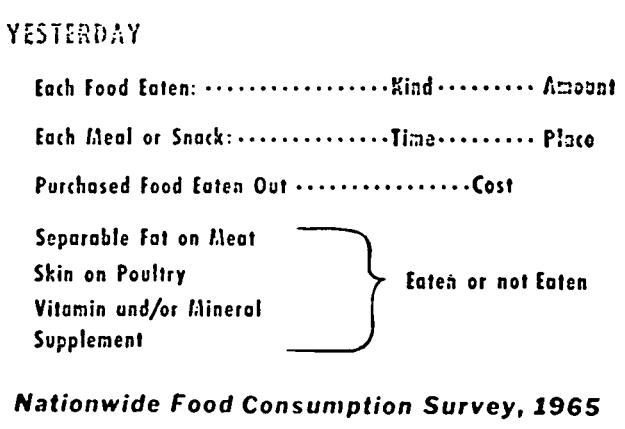


Figure 4

We asked for the time of day at which any food was eaten, whether any vitaman or mineral supplements were taken during the day, the types of places where away-from-home food was eaten and the expense for purchased meals and snacks eaten away from home.

Tabulation of the household data is under way. We have issued a preliminary report on the money value of food used by households in the U.S., spring 1965.[1] This report provides data on the money value of food used at home—bought and home produced—and the expense for meals and snacks eaten away from home.

In our basic statistical reports, we will have three types of information for about 200 foods or groups of foods: the percentage of households using each food during the week, and its average quantity and its average money value. Where pertinent, these data will be shown separately for purchased, home-produced and Federally-donated food. In the publications primarily concerned with dietary levels, there will be data on the average nutritive value of the week's food, distributions showing percentages of households with diets at specified levels of calories and each of 9 or 10 key nutrients, the contribution of selected groups and items of foods to total calories and to the total of each nutrient, and average quantities consumed of about 50–60 foods grouped for their nutritional value.

The major classifications for reporting these data in the first 10 volumes to be published will be region, urbanization, and income. Some charts based upon the preliminary report on total money value of food used by households in the spring of 1965 indicate the types of tabulations that will also be available on the consumption of foods and the nutritive content of diet.

Regional differences in family food supplies were quite large in 1965 as shown by the data in the preliminary report (Fig. 5). The average money value of food used at home in the Northeast was $32 per family per week, almost a fourth higher than the $26 per week reported in the South. Expenditures for food away from home were $7.25 in the Northeast, a third higher than the $5.35 a week in the South.

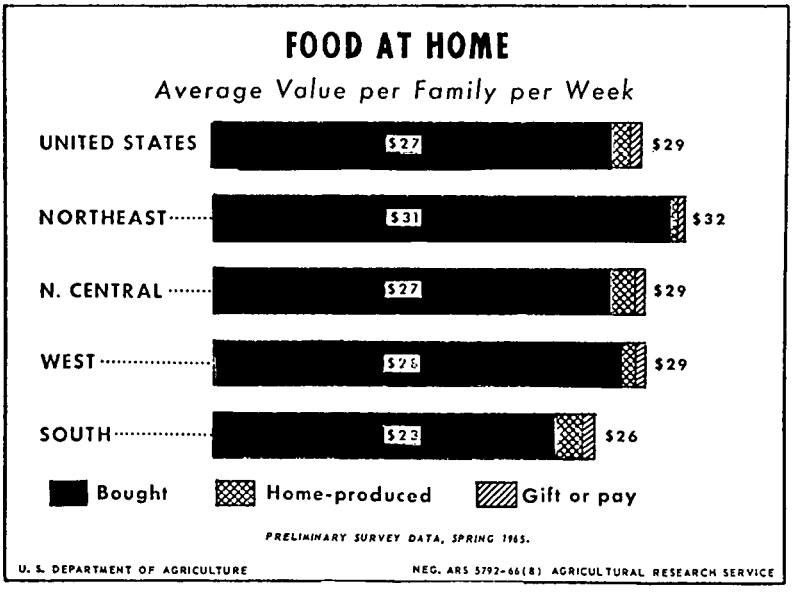


Figure 5

Although farm families now spend more like urban families, there is still considerable difference between them especially in the proportion of the total food money that is spent on food away from home and in the proportion home produced. Almost a fifth of the total money value of food of urban families went

[1] Money Value of Food Used by Households in the United States, Spring 1965. CFE (Adm.)–300. Food Consumption Survey, 1965–66, Preliminary Report. September 1966.

for food away from home compared to about a tenth of that farm families (Fig. 6).

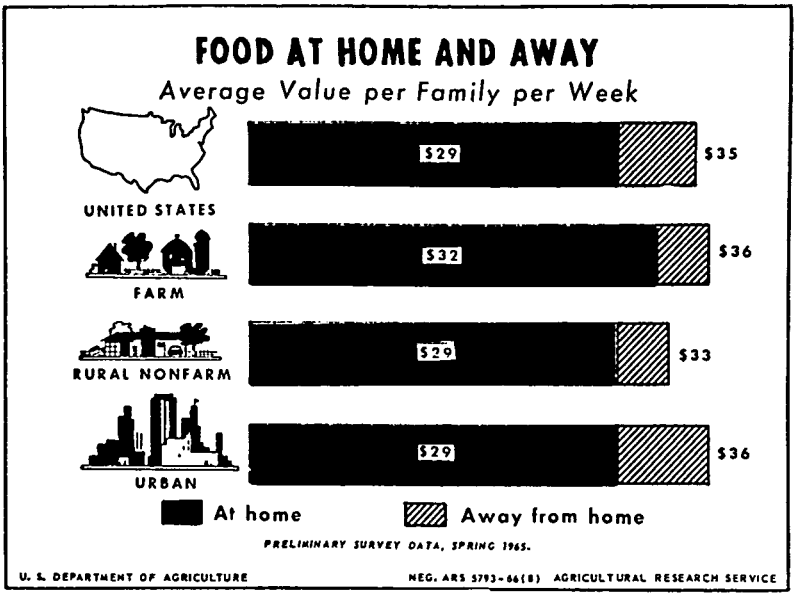


Figure 6

Between 1955 and 1965, the proportion of the food used in farm homes that was home produced declined considerably—from 41 percent in 1955 to 31 percent in 1965 (Fig. 7). The proportion that was purchased increased from 56 to 67 percent. These figures support observations that farm families are becoming more like city families in their way of life.

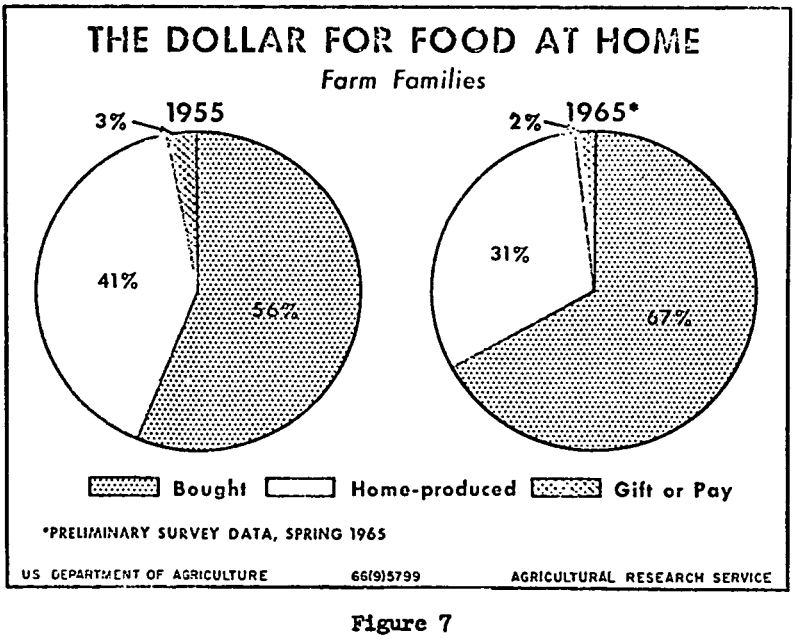


Figure 7

The effect of income as a factor affecting families' food expenditures is shown by the next chart (Fig. 8). In this chart, it is important to note that factors other than income have not been held constant. The average family size for the two lowest income classes shown on the chart is considerably lower than that of the top three classes. In comparing income-expenditure relationships between 1955 and 1965, we have made some adjustments for difference in family size and have plotted the data for food at home and for food away from home on double log scale (Fig. 9). It appears from these data that the

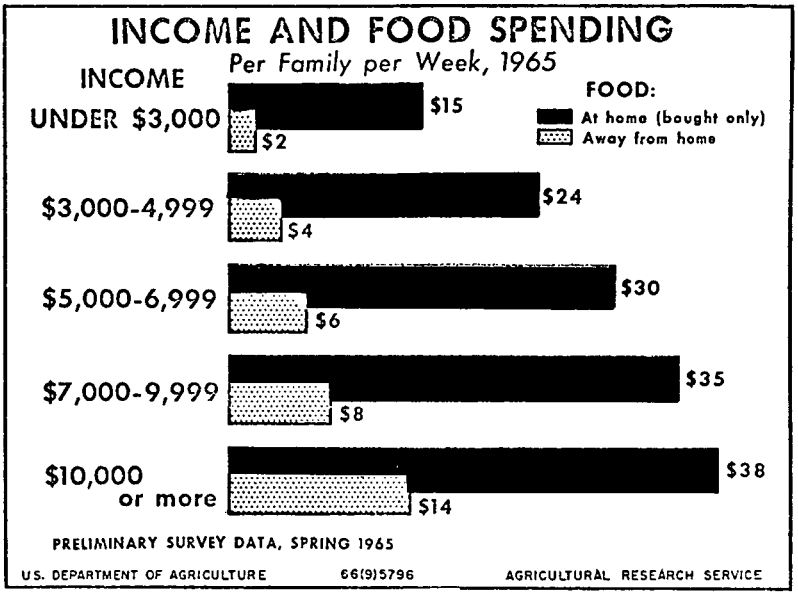


Figure 8

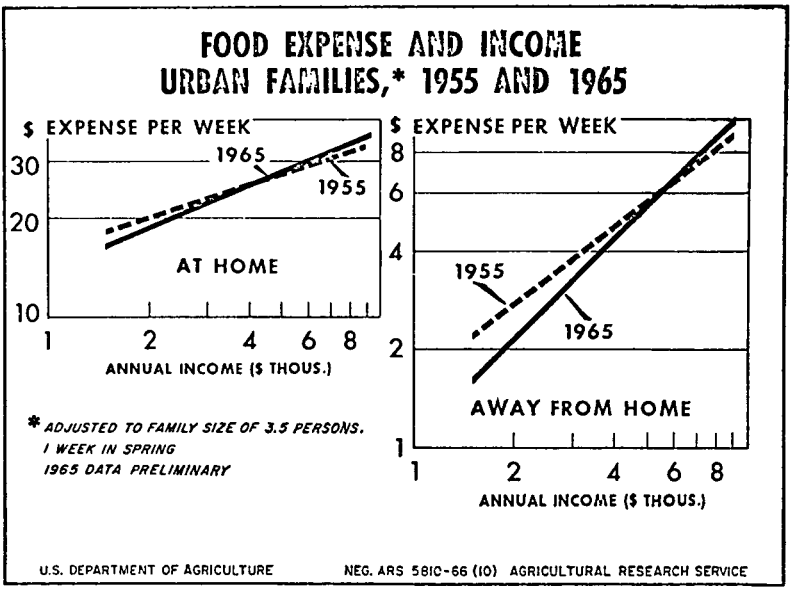


Figure 9

income-expenditure curves for both food used at home and food away from home are steeper in 1965 than in 1955. Because of several differences in the two sets of data, however, we are not ready to conclude that there is a real difference in these income-expenditure relationships. Some of these differences are explained in a paper given at the National Outlook Conference last fall.[2]

Uses of Data From Food Consumption Surveys

Results of food consumption surveys are of value to many public and private agencies and individuals including Congress, USDA, and other Federal agencies, the food industry, and educators. The information provided helps guide farm and food policies, and the appraisal made of the nutrient adequacy of diets furnishes a basis for consumer education and action programs. The data also have wide use in economic and marketing research on the demand for agricultural products. To be more specific, I will speak of five uses.

Profile of national dietary situation. The surveys provide statistical profiles of the dietary situation of population groups in the United States. Comparisons of surveys show trends in food selection and nutritional quality of diets. A brief look at the situation found in earlier surveys shows that in:

1936–37—A third of the households were in great need of better diets.

1942—Considerable dietary improvement had taken place that was credited to higher incomes, concerted nutrition education programs, enrichment of bread and other cereal products and improvements in transporting and distributing foods.

1948.—Continued but less dramatic dietary improvements.

1955—Diets of about a tenth of the households were still short in one or more nutrients on the basis of the same criteria used in 1936–37. Shortages were still found among households in the upper third of the income scale.

1965–66—?

Results of the 1936–37 survey and related findings gave impetus to the enrichment of white bread and flour with three of the B-vitamins and iron, stimulated programs of nutrition education and provided a basis for the school lunch programs that we now have.

The results of the 1955 survey indicated need for dietary improvement in spite of relatively high incomes in the U.S. The data were used to develop new educational materials for families of both low- and middle-income levels. They also provided baseline data for the pilot food stamp program in 1961.

Food budgets. The food budgets at different cost levels developed about 35 years ago by Dr. Hazel K. Stiebeling of the USDA are revised periodically using survey data as a point of reference for family food habits. These food budgets are widely used in counseling with families on making wise use of available food money and also by welfare agencies as a basis for their food allotments for needy families. Costs are estimated quarterly and published in *Family Economics Review.*

More recently costs of the low-cost food budgets have been used by Mollie Orshansky of the Social Security Administration in developing the poverty yardstick now widely used by the Office of Economic Opportunity. The Department's food budgets are also incorporated in the Bureau of Labor Statistics' City Workers' Family Budget and Budget for a Retired Couple. Still another example of use is in the periodic cost-of-living budgets for employed women developed by States for use in the deliberation of minimum wage boards. In several research projects in our own Family Economics Branch, the food budgets are being used as important reference points.

Control and regulatory use of the data. During World War II, data from the 1942 food consumption survey were used in developing the rationing and price control systems. Of course, we hope we shall not have to use the data for that purpose again, but up-to-date data should be available in case the need arises.

In regulatory work the Food and Drug Administration and the Public Health Service both make use of our data on diets and food consumption in making estimates of amounts of residues, food additives and radioactive fallout in total diets. The statistics from these surveys provide weights for the amounts of various foods to use in estimating total amounts of the elements with which the regulatory agencies are concerned. The data on the food intake of individuals will be especially useful in ascertaining differences in amounts of these elements in the diets of different age groups.

---

[2] Changing Patterns of Family Food Spending, talk by Faith Clark at the 44th Annual Agricultural Outlook Conference, November 16, 1966.

When the cereal enrichment program was being developed, the survey data were used to show the effect of different levels of enrichment on different income classes in the population. One of the arguments for enrichment was the evidence from the USDA food consumption survey that low-income families, whose diets were most in need of improvement, would be especially helped by the program.

Market research. Government economists and market analysts from food industries use information from the surveys in developing estimates of the present and future size and location of food markets. Such estimates are used in decision making about food marketing such as relate to new products, new processing plants or firms, or the acquisition of additional processing outlets. The information provided by these studies on the percentage of households that use different kinds of foods or foods in different forms gives a measure of market penetration for existing products and suggests targets for new products.

Many segments of the food industry are eagerly awaiting new information on household and individual food consumption from our 1965–66 survey. It is quite possible that the National Industrial Conference Board will recompute the data to provide "a share of the market type" analysis as they did for the 1960–61 Consumer Expenditure Survey data.

Consumption research. To gain better understanding of the factors that influence consumption, Government, industry and university economists have made many analyses of the data from USDA's food consumption surveys. The data have been used to measure income elasticities of expenditures for food at home and away from home and of the quantities and expenditures for individual foods. Other factors studied have been place of residence, education and employment of the homemaker and size and composition of families. With better understanding of the effect of these factors on the demand for food, better forecasts are possible under different assumptions of income, price, and population. The new data on the food intake of individuals classified by age and sex will provide a better basis for projections as the population distribution by age changes. Such estimates, of course, are required for public policies and programs that affect agricultural production, and the marketing and distribution of food.

In summary—our surveys of food consumption are multipurpose, providing basic data for many action and educational programs and for research needed for policy determinations. The results are important in relation to nutrition and food programs, consumer use of food, health protection programs, and for economic and marketing research that pertains to production, marketing, and distribution of food.

Marketing Uses of Consumer Expenditure Survey Data

By Helen H. Lamale, Chief, Division of Living Conditions Studies, Bureau of Labor Statistics, U.S. Department of Labor

The BLS Survey of Consumer Expenditures in 1950 was described by Vergil Reed as "a gold mine of market data." [1] More recently, the National Industrial Conference Board, in releasing the results of a special tabulation of data from the nonfarm portion of the 1960–61 BLS-Agriculture expenditure survey, predicted that "for many years to come this book will be an indispensable source of information—for those sections of the business community that are involved in any aspect of marketing consumer products and services." [2]

Although large-scale Federal surveys of consumer expenditures, incomes, and savings date back to the 1880's, their widespread use by the business community for marketing studies seems to have begun with the 1935–36 Consumer Purchases Study. This survey was conducted by the Bureau of Labor Statistics and the Bureau of Home Economics (Department of Agriculture), in cooperation with the National Resources Committee, the Works Progress Administration, and the Central Statistical Board. The two bureaus prepared separate reports on the distribution of income and expenditures in individual cities or rural areas which their respective surveys covered. The National Resources Committee utilized the results in the preparation of estimates of national consumption as related to the socioeconomic distribution of the total population. The data were used extensively in developing the household sector of the National Income and Product Accounts

---

[1] Reed, Vergil, "Don't Miss This Gold Mine of Market Data," *Printers' Ink,* December 18, 25, 1953. See also, Printers' Ink Advertisers' Annual-1954 number.

[2] Linden, Fabian, (Editor), "Experience Patterns of the American Family," prepared by the National Industrial Conference Board, 1965, Foreword. See also, "A Graphic Guide to Consumer Markets, 1965," NICB-Life.

of the Department of Commerce, Office of Business Economics. The hearty reception these data got from market analysts probably stems from two factors: (1) it was the first survey covering the total population, and (2) the detailed statistical reports became available at a time when the business community was acutely aware of the need for a better understanding of the characteristics of consumer demand.

The titles of a series of reports by the Curtis Publishing Company in 1938 to 1940 indicate the kinds of analyses which were made, for example, "Market Patterns in Eight Large Cities" and "Who Owns Fire, Health and Accident, and Automobile Insurance." The 1935–36 survey data and the small-scale nationwide survey in 1941 were also extensively used in more scholarly studies which had relevance to market research, for example, the National Bureau of Economic Research, "Studies in Consumer Installment Financing;" the Twentieth Century Fund 1947 study, "America's Needs and Resources;" and "The Economics of Installment Buying," by Reavis Cox.

The BLS Survey of Consumer Expenditures in 1950 was limited to urban areas, but the sample included 91 metropolitan areas and urban places. Through a grant from the Ford Foundation to the Wharton School of the University of Pennsylvania, the expenditures, income, and savings data were tabulated and published in highly disaggregated form—one-variable classifications for large cities and three-variable classifications for 9 classes of cities and for the urban United States.[3] In addition, much of the supplemental information, such as ownership of durables and quantities purchased, was published or made available to analysts through special tabulations. Thus, the 1950 data found their way by many routes into a wide variety of marketing studies which ranged from analysis of the market for a single item, or class of items, in a single city, or class of cities,[4] to the calculation of demand relationships for major classes of commodities and services[5] and the development of consumption models.

SURVEY OF CONSUMER EXPENDITURES, 1960–61

Before discussing some of the market uses which are being made of data from the most recent survey, the Survey of Consumer Expenditures in 1960–61, I will review its essential features. In this survey, detailed information on annual expenditures, incomes, and changes in assets and liabilities was obtained from 13,728 families and single consumers. The Bureau of Labor Statistics collected information for 1960 and 1961 from 9,476 families residing in 66 metropolitan areas and small urban places, and for 1961, in cooperation with the Department of Agriculture, from 2,285 rural nonfarm and 1,967 rural farm families. Information for 1959 for Anchorage, Alaska, was included in the national summaries. Thus, for the first time since 1941, information is available for a representative cross section of all American families—urban, rural nonfarm, and rural farm. Information for 1959, 1960, 1962, or 1963 is also available for 10 other cities which were not a part of the national sample.[6]

In addition to the detailed expenditure, income, and saving data, information on 1-week's purchase of food items was obtained for urban and rural nonfarm families. Information on many family characteristics, their living arrangements, and a selected inventory of household durables was also recorded for each family.

The primary purpose for the urban portion of the survey was to revise the Consumer Price Index, and this use was a controlling factor not only in the design and content of the survey but also in priorities established for processing and disseminating the data. From the beginning of the project, the need for general purpose tabulations and for making the data available to special purpose users was recognized, and a substantial volume of data was provided in the General Purpose Tabulation Program.[7] The data have been tabulated separately for the urban segment of 66 metropolitan areas and urban places and have been combined for the 4 broad census geographic regions and the total United States—

---

[3] "Study of Consumer Expenditures, Incomes, and Savings," University of Pennsylvania, 1956–57, Volumes I–XVIII.

[4] For example, see Smith, Arthur L., "Good and Bad Cosmetic Markets," *Drug and Cosmetic Industry*, August 1954.

[5] For example, see Lippitt, Vernon, "Determinants of Consumer Demand for House Furnishings and Equipment" in Consumption and Savings, Volume I, University of Pennsylvania, 1960.

[6] Cincinnati, Ohio (1959); Houston, Tex., Kansas City, Mo.-Kans., Milwaukee, Wis., Minneapolis-St. Paul, Minn., and San Diego, Calif. (1963); Fairbanks, Alaska (1959); Juneau and Ketchikan, Alaska (1960); and Las Vegas, Nev. (1962).

[7] For description, see "Survey of Consumer Expenditures, 1960–61: List of Statistical Reports," Bureau of Labor Statistics, U.S. Department of Labor, September 1966.

for urban, rural nonfarm, total nonfarm, rural farm, and total U.S. families. Reports for the metropolitan areas, regions, and the United States include summaries of major categories of expenditures, incomes, and savings for families classified by annual money income after taxes; family size; age, occupation, and education of head; housing tenure; race; family type; and number of full-time earners, i.e., one-variable tables, as well as many two-variable classifications, e.g., income and family size. Detailed tabulations of expenditures for subgroups and individual items of goods and services, sources of income, and nature of change in savings and debts have been prepared for families classified by income and family size, separately for urban, rural nonfarm, total nonfarm, and all U.S. families, at the regional and U.S. levels.

The Bureau is currently using the survey data to revise the "modest-but-adequate" standard budgets for a 4-person city worker's family and for a retired couple. Estimates of the autumn 1966 costs of these revised budgets will be published in August and September, separately for 23 metropolitan areas, for medium-sized and small cities in 4 geographic regions, and for urban United States. In addition, budgets for a lower and a higher standard are being developed for both family types, and estimates of their spring 1967 costs for the same metropolitan areas and classes of cities are scheduled for publication by the end of 1967. These budgets provide the basis for preparing comparative indexes of living costs. Both the standard budget cost estimates and the indexes of place-to-place differences in living costs based upon them are used extensively by the business community.

Despite the rather substantial general purpose tabulation and publication program, the Bureau has received numerous requests for special tabulations. In cases where these requests could be met from existing unpublished data, we have provided copies at cost through the BLS Regional Offices. One such tabulation, which was planned especially to meet repeated requests of market analysts, provides for each metropolitan area surveyed average expenditures for individual items, or classes of goods and services, by families classified as "below" or "above" the median income for the area. Such a tabulation of the 1950 data for New York City was the basis for a market profile, entitled "New York's 19¼ Billion Dollars—Who Spends It and How," prepared by the Research Department of the New York Times as a guide to advertisers and advertising agencies.

It became evident very early in the processing of the 1960–61 survey data that the Bureau did not have either the staff or computer facilities to fill requests for special tabulations, even on a reimbursable basis. The BLS made the special tabulations for the NICB (mentioned above) with the understanding that the Board would provide the machine programs and make the data available to others.

Since many of the requests were from other Federal agencies having data processing equipment, the Bureau rented the basic data tapes to them under a contract which complies with our nondisclosure regulations. At present, six Federal agencies (in addition to BLS and USDA) have these master tapes. For users outside the Federal Government, the Bureau prepared a General Purpose Tape, consisting of three standard-length (2,400) feet reels of magnetic tape, which is sold on a restrictive contract basis. This tape contains for each of the 13,728 urban and rural consumer units in the sample most of the family characteristics codes; expenditures for major groups, subgroups, and classes of goods and services; and details on sources of income and changes in assets and liabilities. Twenty-one universities, trade associations, and business organizations have purchased this tape.

The requests for special tabulations and for the basic data tapes have revealed the wide range of uses which are being made of the data, many of which are directly or indirectly marketing uses. Probably, from the statistics producer's viewpoint, some of these are misuses. Before making the 1960–61 basic data available, the Bureau explored with its advisory committees and with such other groups as the Social Science Research Council and the Federal Statistics Users Conference, the problems of misuse which could result from such dissemination of disaggregated consumer expenditure and income information from a sample survey. The consensus of these discussions was that, in the present computer age, the need for such data and advantages of their general availability far outweigh the danger and the disadvantages; that the statistics producer's responsibility is to provide as complete descriptions, interpretations, and evaluations as possible; but that uses should not, and misuses cannot, be controlled.

Since the 1960–61 data have been so widely distributed and many, perhaps most, of the business contacts are made through the Bureau's regional offices, the

following discussion is not intended as a comprehensive survey but only to point out various types of marketing uses with a few illustrations. I shall try to summarize the advantages and disadvantages of the Bureau's expenditure survey data for these purposes from the producer's point of view. As a market research user of the data, the discussant may well see both the uses and my appraisal of the data in a different light.

TRADITIONAL MARKET USES

Until recent years, market use of consumer expenditure survey data was rather generally thought of as the use of the data by business organizations to define the market for a specific and, usually, a narrowly defined product, e.g., washing machines, toilet soap, or men's shoes. Such analyses are used to estimate the share a manufacturer has in each product market; to establish sales quotas or reorganize sales territories; to plan advertising, select or appraise advertising media and sales promotion and to select cities, or consumer groups, in which to test new products. Such uses still constitute a major part of business demand for the Bureau's survey data. The special NICB tabulations of the 1960–61 survey data (referred to above) were designed to serve these purposes, and the Board is planning additional publications on the "share of the market" theme.

However, such uses of the data are not limited to trade association and business analysts. Government agencies, especially analysts in the Departments of Commerce and Agriculture, and university researchers are citing or using the data for the same kinds of studies. For example, individual metropolitan area data from the 1960–61 survey are listed in the U.S. Business and Defense Services Administration "Facts for Marketers," and the General Purpose Statistical Reports are cited in this agency's "Guide to Negro Marketing Information." Andrew F. Brimmer, when Deputy Assistant Secretary of Commerce, used the survey data in a speech before the National Association of Market Developers on "Economic Trends in the Negro Market." As a part of its planning service for industrial and financial clients, the Stanford Research Institute used the survey to prepare a report on spending of "Upper Income Families," and the Super Market Institute, Inc., of Chicago cites the survey data in "A Guide to Source Material for Store Location Research."

Another traditional market use of consumer expenditure survey data is in the development of new products or services. The Federal Government and general public have an interest in some of these studies. For example, the 1960–61 survey data were used as a part of a study of the demand for passenger transportation in the Washington-Boston corridor in the investigation of the feasibility of high-speed railway facilities and demand for supersonic transport. The potentials of the survey data for estimating the costs of operating combustion-engine autos under various assumptions, as compared with anticipated costs for operating electric autos are being studied.

Some of the market studies involve analyses of spending of different types of families for classes or items of goods and services in relation to their spending for related or competing items. The Division of Marketing Research of the Life Insurance Agency Management Association is considering a study of variations in life insurance spending for families having various levels of housing expenditures.

The National Association of Real Estate Boards, of which the discussant is Research Director, is making special tabulations of the survey data to compare homeowner expenditures for various housing-related items with similar expenditures of renting families.

The data are being used in several studies on the impact of taxes of various kinds, with particular emphasis on the differential effects of State and local taxes of various types, e.g., sales and real property taxes, on various groups in the population and among regions.

FORECASTING MARKET TRENDS

A major use of the CES data by market analysts and business in general is as background on the overall picture of consumer spending in relation to income, saving, and other socioeconomic characteristics of families, and within the broader social and economic scene. Increasingly, business analysts recognize that they need such general background from which to plan their own detailed market research studies and general business programs. Such general business economic analyses do not usually require as current or detailed data as some of the more specific market studies mentioned above. Frequently, the historical cross-section

CES data can be used effectively with time series aggregate expenditure data, Census business and demographic statistics, and a wide range of social statistics for such studies and for developing forecasting models.

In commenting on a paper on the changing pattern of consumer expenditures between 1950 and 1960, as revealed by the BLS surveys,[8] a market analyst said it was "as useful as a map is to a navigator." He observed that in nine instances, in addition to wars and taxes, the changes in spending were linked to influences which result "from deliberate decisions of either government or business management," (i.e. homeownership; retirement; Social Security and private insurance; school- and employer-provided lunches; public and private transportation; urban renewal; education; household equipment furnished by builders; and medical care). He stressed the importance of such analyses because they point up the "larger forces" which those "who study the consumer in a sales context are prone to overlook" but which "generally take precedence over influences of the marketplace."

A study by F. G. Adams and D. S. Brady of the University of Pennsylvania utilized the published BLS expenditure survey reports back to 1918 to trace the diffusion of new durable goods among various groups in the population and their impact on consumer spending.[9]

The 1950 and earlier BLS survey data were used, in conjunction with time series and similar data, for several popular studies which described in depth the past, present, and probable future consumer markets within a broad social and economic framework. "The Changing American Market" and "Markets of the Sixties" by the Editors of Fortune magazine are examples of such studies.

The 1950 BLS survey data were published in greater detail than any previous study, and the Ford Foundation grant to the Wharton School provided for a series of monographic studies, many of which were directed toward identifying demand relationships for major components of total consumption expenditures.[10] This work and related analyses in private research agencies, other universities, and Federal agencies have resulted in socioeconomic models designed to forecast aggregate consumer spending and its composition in the 1970's.

One such model is the Program Analysis for Resource Management (PARM) model of the National Planning Association which made extensive use of the tabulation of the 1950 detailed expenditure data by income and family size.[11] Similar 1960–61 data, as published in Supplements 3A to the BLS Report Series 237 at the national and regional levels, are being used to check the forecasting equations based on the 1950 data.

In the September 1966 issue of the Battelle Technical Review, Joseph W. Duncan describes the Socio-Economic Model of Battelle Memorial Institute.[12] In this model, projections of consumer spending patterns were developed by projecting the structure of consumer spending as found in the 1950 and 1960–61 BLS consumer expenditure studies. Mr. Duncan announced that "a recently completed study for long-range planning in several major U.S. corporations included a detailed projection of 252 socioeconomic classes," i.e. defined in terms of the income of the household, and the occupation and educational attainment of the household head.

Similar projects, designed to wed the cross-section and time-series data and to develop forecasting models, are going on in the business schools of many of our universities. A recent publication of the Institute of Business and Economic Research of the Graduate School of Business Administration of University of California at Berkeley reports on a series of "Studies in the Demand for Consumer Household Equipment" [13] which relies, primarily, on data from the Michigan University Surveys of Consumer Finances but stresses the value of combinations of subject matter, type of analysis, and sources of data, described as the "utility of

---

[8] Chase, Arnold E., "Changing Patterns of Consumer Expenditures, 1950–1960"; and discussion by Sidney Hollander, Jr., in *1963 Proceedings of the Business and Economic Statistics Section*, American Statistical Association, pp. 65–75, 90.

[9] Adams, F. G., and Brady, D. S., "The Diffusion of New Durable Goods and Their Impact on Consumer Expenditures," American Statistical Association, op. cit., pp. 76–88.

[10] "Consumption and Savings" Volumes I and II, University of Pennsylvania, 1960.

[11] Snyder, Eleanor M. and Edmonston, J. Harvey, "Personal Consumption Model," NREC Technical Report No. 15, National Planning Association, Washington, D.C., October 1963. See also, "Informational Requirements for Planning and Projections—A Syllabus of Background Materials," National/Regional Economic Projections Series, Report No. 66–J–3, National Planning Association, Washington, D.C., 1966.

[12] Duncan, Joseph W., "A Framework for Forecasting Socio-Economic Change," in *Battelle Technical Review*, Volume 15, September 1966, pp. 9–13.

[13] Carman, James A., "Studies in the Demand for Consumer Household Equipment," IBER Special Publications, *Research Program in Marketing*, Graduate School of Business Administration, University of California, Berkeley, 1965.

'panoramic' research." This report also expresses, as do many other business users of consumer expenditure data, the need for continuing expenditure survey data, i.e. panel surveys.[14]

CONSUMER SERVICE USES

The use of consumer expenditure survey data by family counsellors and by families themselves, as guides for appraising family spending and to improve family financial practices, is probably as old as the surveys themselves. (One of the early BLS studies gathered information on the cleanliness and neatness of the home.) The data have been used in unnumber books and articles on the general and specific aspects of careful spending. The most recent such publication by a Federal Government agency is "Helping Families Manage Their Finances" by the Agricultural Research Service.[15]

However, until recently, this use could hardly have been considered a "marketing" use, although for many years research departments of life insurance companies, household lending organizations, and banks have used the data in their operations and publications. During the past year, the computer and publication of the 1960–61 survey data in considerable detail for various types of families have transformed these spending guides into a marketed service. Such a service called "Family Money Profile," was initiated in 1966 by the Pittsburgh National Bank. Families are invited to fill out a confidential application which give monthly income, age, family size, and occupation.

Based on the BLS consumer expenditure survey averages and spending patterns for similar families in the Pittsburgh area, a computer analysis prepares a guide to spending and saving for the individual family.[16]

In the April 1967 issue of "Changing Times," The Kiplinger Magazine initiated a similar service on a nationwide basis.[17] Upon receipt of the completed questionnaire, the 1960–61 published survey data, sorted to match the applicants characteristics, are converted into a "Spending Yardstick" by which the family can compare its spending pattern with the average for similar families. The charge for this service is $1.75. After 6 days, returns were being received at the rate of 1000 per day. One request was from the sales manager of a small concern, asking if he could have 150 copies of the questionnaire for distribution to employees as part of the company's consumer education program. He indicated that the company would bear the cost of the service as a "fringe benefit."

ADVANTAGES AND DISADVANTAGES OF CES DATA FOR MARKET USES

Those of us who have struggled with the problems of collecting, processing, and interpreting consumer expenditure survey data realize the strains that uses, such as those which I have just described, place on data. We are often concerned that these sample survey data, designed primarily for the development of consumer price indexes, standard budgets, and analyses of levels and standards of living, may be inappropriate for many of the uses to which they are put. On the other hand, we are constantly reminded of the inadequacy of data sources on consumer spending otherwise available to the business community and agree that maximum use should be made of data collected at public expense.

The advantages, as I see them, of the Federal consumer expenditure survey data for marketing and general business use are as follow:

(1) The consumer expenditure surveys are unique in that they provide complete and detailed information on expenditures, incomes and changes in assets and liabilities in relation to a great variety of socioeconomic characteristics of families and for many localities, including classifications based on spending for related or competing goods and services. As levels and standards of living and spending habits change, the expenditure data, even though not on a continuing basis, are a unique source of data for

[14] For a discussion of this need, see "Improved Statistics for Economic Growth—A Compendium of Views and Suggestions From Individuals, Organizations, and Statistics Users" and "Comments by Government Agencies on Views" (p. 46), Subcommittee on Economic Statistics, Joint Economic Committee, 89th Congress of the United States, U.S. Government Printing Office, July 1965 and March 1966.

[15] "Helping Families Manage Their Finances," Home Economics Research Report 21, Agricultural Research Service, U.S. Department of Agriculture, Washington, D.C., June 1963. See also, "Socioeconomic Research Abstract Series of the Home Economics Education Service," Miscellaneous 2230–3, U.S. Office of Education, September 1940.

[16] "Family Money Profile," a news release; "How to Manage Your Money to Get More Out of Life"; and "Record Book for Thousandaires"; Public Relations Department, Pittsburgh National Bank, Pittsburgh, Pennsylvania, 1966.

[17] "Check Up on Your Family Spending," in *Changing Times*, April 1967, pp. 48–49.

understanding the changing composition of consumer demand and the socio-economic factors affecting aggregate demand.

(2) They represent the only substantial volume of historical data for such micro analyses that are in major respects comparable—at least as far back as the mid-1930's.

(3) My cursory review of market research literature has revealed a great concern with the quality of other data sources available to the business community for such purposes. An article in the fall 1966 issue of "Business Horizons" published by the Graduate School of Business, University of Indiana puts it this way:

"Most research buyers agree with the adage that, no matter how sophisticated the research design, how elegant the analysis, or how incisive the final report, the results of marketing research are no better than the interviewing that generates the data." [18]

The consumer expenditure surveys of the Federal Government have a long history of careful attention, not only to the quality of the interview but also to the basic design, sample and content, which getting such a study through advisory committees and approved by the Office of Statistical Standards requires.

On the other hand, many of the *disadvantages* of the data are not as obvious, particularly, when the users get the basic data in magnetic tapes or in a polished summary volume. Many of the deficiencies of the data, particularly with respect to their uses for aggregate estimating, stem from the fact that the basic design, definitions, and classifications are those appropriate for their primary purposes, i.e., the development of consumer price indexes, standard budgets, and analyses of levels and standards of living. For example, expenditures of families for gifts of goods and services to persons outside the family are neither obtained in detail nor tabulated as a part of the expenditure for such items for family uses. Thus, the data may be "gold for market research," but at best they are crude ore which requires considerable refinement and interpretation when used in a specific market analysis.

Development of tabulation designs and procedures appropriate for market studies are difficult and time-consuming and must be delayed until the sponsoring agency has had an opportunity to review and evaluate the basic data problems. The preparation of general purpose tapes of the kind available for the 1960–61 survey for use by other government and nongovernment agencies creates a whole set of new problems and new areas for which different kinds interpretation and appraisal must be developed. Efficient and valid uses of such basic data require a whole lot more than availability of compatible electronic data processing equipment.

In commenting on the demand analyses reported at a conference which followed the publication of the 1950 BLS data, Margaret Reid summarized the problem very succinctly when she said:

"Probably never before has a single conference added so much to a stock of regression coefficents. . . . The next large scale survey will undoubtedly be followed by a similar crops as Univacs with magical speed add, subtract, multiply, and divide, and trace relationships among a multitude of variables. The data flow in, the coefficients roll out. A stock of coefficients may be the beginning, but it is not the end of wisdom. It must provide the evidence that makes a coherent whole. The achievement of coherence is however outside the scope of Univacs." [19]

I would add that it requires the wisdom and skills of discerning analysts.

Data From Tax Returns and Their Users

By Vito Natrella, Internal Revenue Service

The Internal Revenue Service is the repository for the most comprehensive collection of economic data, based on administrative documents, anywhere in the world. This year over 70 million individual tax returns will be filed covering 110 million taxpayers and 80 million dependents. On these returns will be economic data referring to about 95 percent of the U.S. population.

---

[18] Mayer, Charles S., "The Overlooked Ingredient in Survey Research," *Business Horizons*, Volume 9. No. 3. Graduate School of Business, Indiana University, Fall 1966, p. 75.

[19] "Consumption and Savings," Volume I, University of Pennsylvania, 1960; p. 143.

In addition, tax returns will furnish a complete and presumably accurate record of business activity. In 1967 the Service will receive more than 1½ million corporation income tax returns, almost one million returns for partnerships, and about 10 million individual returns reflecting the operations of sole proprietors, both farm and nonfarm.

Although these constitute the main areas, the Service receives many other returns containing useful information. They include fiduciary, estate, and gift tax returns, and returns of tax exempt organizations, as well as the many millions of employment returns, estimated tax returns and information returns.

Coping with all these documents and maintaining standards of fair and equitable enforcement is, of course, a monumental task. To this end there has been, over the past five years, a change to completely integrated automatic data processing of individual and business tax returns and the establishment of master files. The last regions were converted at the beginning of this year so that the system is now operating on a nationwide basis. In addition to the effect on revenue processing, this has some very significant implications for our statistical programs.

Statistics of Income Reports

For the past 50 years the Internal Revenue Service has been publishing a series of statistical compilations under the title *Statistics of Income.* This is in accordance with a requirement of the Revenue Act of 1916, incorporated in Section 6108 of the Revenue Code, providing that:

"The Secretary or his delegate shall prepare and publish annually statistics reasonably available with respect to the operation of the income tax laws, including classifications of taxpayers and of income, the amounts allowed as deductions, exemptions, and credits, and any other facts deemed pertinent and valuable."

While the publication for 1916 consisted of only one volume, Statistics of Income, in response to this mandate, has expanded to four basic series covering individuals, corporations, business, and estate, gift, and fiduciary tax returns. The last 50 years has, of course, seen important changes. Statistics of Income has developed from a tax collection oriented publication to a document of great economic significance. This came about through the extension of the income tax to practically the whole population. and the introduction of new and complex provisions making the tax return a rich source of economic data.

The volume of *Individual Income Tax Returns* presents, in considerable detail, data showing adjusted gross income, taxable income, income tax liability, sources of income, and itemized deductions, classified by size of adjusted gross income. Taxable income and tax are also shown by applicable tax rates. Selected items of income and tax are shown by States, annually, and by metropolitan areas biennially. Some of the special analyses which have been covered in recent reports include self-employed pension deduction, income of taxpayers over 65, contributions by type of recipient, and divident recipients by number of payer corporations.

Corporation Income Tax Returns presents the many items from the income and balance sheet statements, as well as income tax liability and distributions to stockholders. The statistics are classified by industry, and by size of total assets, business receipts, and net income. Most of these data are also shown separately for returns with net income, consolidated returns, and returns of Small Business Corporations electing to be taxed through shareholders. In addition, selected items are available for each Internal Revenue district or region. Other statistics are designed primarily for use in revenue estimating and tax analyses. Special tabulations have included classifications by ratio sizes based on the relationships between net income and business receipts, and business receipts and total assets. Special subjects for which statistics are available include inventory valuation and depreciation methods, paterns of current and prior year net income, unincorporated businesses electing to be taxed as corporations, and investment credit.

The most recent addition to our publications in the regular Statistics of Income series is *U. S. Business Tax Returns.* It was first introduced for tax year 1957 in response to widespread demand for an annual volume of statistics covering all three of the principal forms of business organization, i.e., sole proprietorships, partnerships, and corporations. The primary emphasis of the report is to provide data with industry detail about unincorporated businesses. Statistics of Income is the only reliable source of this segment. Each year, new data are introduced into the report such as financial ratios, depreciation methods, inventory methods, use of the investment credit, statistics on a state-by-state basis,

and sources of farm income. In addition to industry, these data are classified by size of enterprise as measured by profits, losses, receipts, and assets.

The *Fiduciary, Gift, and Estate Tax Returns* book, which is on a three year basis, brings together information from three important sources relating to the wealth of individuals in the United States. The section dealing with fiduciary income tax returns shows sources of income, deductions, and exemptions, as well as tax. Classifications of the data are made by size of total estate or trust income, tax status, and States. The section dealing with gift tax returns shows the reported gifts by types of property, exclusions, specific exemptions and deductions, with various size classifications. The section dealing with estate tax returns shows gross estate by types of property, deductions, exemption, taxable estate, credits, and tax, by size of gross estate.

Supplemental Reports

With the growing need for more precise and sophisticated information about the operation of specific features of the tax laws, it has become increasingly necessary to search out and interpret data from supporting schedules of the tax forms. In order not to delay the regular Statistics of Income reports it has been preferable to supplement them with the data obtained from special studies. For tax year 1959 a supplemental report was prepared showing sales of capital assets reported on individual tax returns. This report was only a pilot study, and was followed by a study for 1962, recently released, presenting considerable detail as to types of assets sold and length of period held.

The supplemental report on *State and Metropolitan Area Data for Individual Income Tax Returns*, 1959–61 presents area data in one place in addition to information not previously published. State and metropolitan area statistics are, of course, included in the regular SOI series.

Depletion Allowance for Mineral Production reported on U. S. Tax Returns, 1960 contains information on the computation of depletion allowances, income and expenditures related to mineral production as reported on schedules filed by individuals, partnerships, and corporations.

Three reports on *Foreign Tax Credits Claimed on Corporation Income Tax Returns* covering 1961, 1962, and 1964 are now planned. The volume for 1961 is completed and is awaiting publication. It presents information by country on income, including dividends, received from foreign sources and the foreign taxes paid on this income. Future reports will also provide information about U. S. corporations owning stock in foreign corporations and about controlled foreign corporations 50 percent or more owned by domestic corporations.

Another supplemental report, recently issued, covers *Farmers' Cooperatives* for 1963. The report presents, for exempt and nonexempt farmers' cooperatives, information on assets, liabilities, receipts, deductions (including patronage dividends), and income tax. Size classifications include total assets, business receipts, and net income. For exempt cooperatives detailed income statements and balance sheets are presented by type of product marketed and by State.

Of considerable interest is a report, almost ready for release, on Personal Wealth estimated from Estate Tax Returns filed during 1963. The regular Statistics of Income volume presents information on the assets shown on estate tax returns. The supplement presents estimates of the number and wealth of all individuals living in 1962 whose assets would have been subject to the Federal Estate tax if they had died in that year. Under the assumption that death is a random sampler, estimates are obtained by multiplying asset data from each estate tax return by the inverse of the appropriate mortality rate for each age-sex stratum. The report presents wealth distributions for the top wealth holders by type of asset and by age, sex, marital status and State of residence of the owners.

Related Activities

In addition to the Statistics of Income volume on corporate tax returns there is available a Source Book which presents considerably more detailed information. This Source Book may be purchased in microfilm form or in printout form by the page. It contains the complete data for income accounts and balance sheets as shown in the published volume subdivided by asset size classes and 4-digit industry group.

Another set of data which has been prepared from the corporation COI covers information by major industry, size group, and State of filing. As regards economic impact, it is recognized that State of filing is not too significant for large, widely spread, corporations. It is necessary to keep this qualification in mind when using the data. We feel, however, that the data are very useful for analysis

of small business where State of filing generally represents location of business, and also are useful in terms of measuring the workload in tax administration.

Use of Tax Return Data for Tax and Economic Policy

One of the major uses of tax return data is for tax research and the development of economic policy. This involves a number of agencies such as the Office of Tax Analysis of the Treasury, the Joint Committee on Internal Revenue Taxation, the Council of Economic Advisers, and the Bureau of the Budget. Special studies and tabulations are prepared for the use of these agencies and are often included in the Statistics of Income program. In order to make it responsive to their needs, the representatives of these agencies are asked each year to review the program.

In addition to the use of these publications and tabulations for this purpose the advent of the electronic computer has made it possible to develop a powerful analytical tool for tax analysis, a tax simulation model. The 1964 Individuals Tax Model consists of three reels of tape containing the data for a one-fourth subsample of the Statistics of Income sample—approximately 95,000 returns from the population of 66 million. This file can be further subsampled through the use of a Select Code which permits random partitioning of each sample class into equal subgroups. The data record contains all items from the tax return required for tax computation together with codes indicating sample class, district, form of deduction, marital status, and use of certain special provisions.

The file can be manipulated through a generalized program using variable parameters so that the tax on each return is computed under different conditions. These might represent a proposed tax plan or a projection to a future level and distribution of income. The changes are introduced by the use of control cards so that a new tape file is produced with the desired changes.

In addition to the manipulation program there is a flexible table generator program which can be used to weight and tabulate the file produced by the manipulation program. It can also be used to prepare tables from the existing 1954 file. For some purposes it is desirable to operate on returns selected for particular characteristics, since some tax proposals affect only one segment of taxpayers. This can be done by both the manipulation program and the table generator program.

The Tax Model has made it possible to determine the revenue effect and taxpayer impact of legislative proposals on very short notice. Also, this can be done, not just for one change at a time, but for combinations of several changes simultaneously, an almost impossible task prior to the Model. These changes might involve new tax rates, exemption allowances, floor and ceiling limitations on certain deductions, partial income exclusion (such as sick pay or dividends), capital gains treatment, and substitution of tax credits for deductions from income.

Earlier IRS Tax Models—those for 1960 and 1962 which were nowhere near as flexible as the 1964 model—have proved invaluable to the Treasury and to the Congress in development of the Revenue Act of 1964 and the Tax Adjustment Act of 1966. In the latter instance innumerable runs were made on various plans for graduated withholding to determine the varying impact of underwithholding and overwithholding. Suggestions to reduce overwithholding made at Ways and Means Committee sessions could be run through the Tax Model and the results presented within a day or two. What this means is that it is no longer necessary to use rough estimates or guesses when faced with the problem of measuring the impact and effect on revenue of new tax proposals or of determining what tax revenues will be in the future under various assumptions as to the tax structure and level of income.

The Individual Tax Model data tape file can be purchased by anyone doing research for about $350. The data record contains no item identifying the taxpayer so that IRS regulations against disclosure are not violated.

Uses by Other Agencies

Data from tax returns are the basis for considerable further analysis and estimation by other Government agencies. Significant parts of the system of national income accounts prepared by the Office of Business Economics, Department of Commerce, find their origin in SOI. The corporate profit segments of the national accounts and the unincorporated business income are based on SOI. Other parts of the accounts use data on interest, depreciation, personal income, the size distribution of income and State distributions of income. In addition, wholesale and retail mark-up rates are determined from tax returns in order to estimate the value of consumer purchases of commodities.

A substantial use is made of tax return data by the Census Bureau. This is in continuation of the policy of relieving business of the expense and work involved in answering questionnaires and surveys. In addition, using administrative documents is more economical for the Government. The Economic Censuses make use of data for small business establishments by substituting tax return information for separately collecting Census reports. In addition Census is experimenting with the possibility of estimating population in intercensal years by using annual data on exemptions available from tax returns. The Internal Revenue Service has been cooperating with the Bureau of the Census in modifying returns and tabulations so as to fit in with data needed for sole proprietors and partnerships.

Other agencies estimate current data from small samples using tax return data which is more complete as the benchmark. For instance, the Securities and Exchange Commission forecasts plant and equipment expenditures and estimates the net-working capital of corporations using Statistics of Income data as the basis for extrapolating.

Nongovernment Users

In addition there are, of course, many studies that have been based on data appearing in Statistics of Income or the Source Book. These range from studies of tax policy, to studies of consumer characteristics and studies of economies of scale in advertising. Data from the Source Book have been republished in the form of performance ratios so that a business can compare its results with those of other corporations in the same size and industry category. There also have been studies of the incidence of capital gain taxation, of interest taxation and studies of the tax burden on stockholders, based on SOI data.

While in the past users have had to depend on data published by IRS, a recent development makes broader use possible. Public Law 87–870 authorizes the Service to accept reimbursement for the cost of doing special studies for private organizations and persons. Major users have been universities, State Governments, private individuals, companies and research organizations.

New Developments and Implications for Users

I have discussed the tax model which the Service has developed based on individual income tax returns. There are two other tax models which have been developed although they are not in as finished state as the model for individuals. These are the tax model for sole proprietorships just completed, and the tax model for corporations soon to be completed. Both these models will make possible research on the impact of alternative tax policies on business.

The use of tax models is not limited to tax purposes, however. They provide the basis of quickly producing tables varying from the ones published in Statistics of Income. Special relationships can be brought out of characteristics not previously considered, without the long lead time necessary for the regular tabulations.

As I mentioned previously, the change to automatic data processing has resulted in the establishment of master files for the major types of returns. The master file makes possible a number of significant improvements. For one, the SOI program can be made much more efficient and responsive to research needs. Significant improvements in sampling procedures are now possible. Whereas up to this time sampling has been done manually on the basis of adjusted gross income classes predetermined for audit purposes, a new plan has been introduced which uses ending digits of Social Security numbers and is independent of audit classes. In effect, the sampling plan can be more closely tailored to the needs of our program. The sampling characteristics used can be specially chosen, as for instance, business receipts or industry for sole proprietorships, in addition to adjusted gross income. AGI can be used in finer detail than the audit classes permit. In this way it will be possible to use a smaller sample with the same reliability as heretofore.

Another advantage of sampling from the computer is the possibility of making special studies when a particular characteristic is needed. This is important for items which do not occur frequently in the population as for instance, moving expenses, sick pay, or child care allowances. Since these items are identified in the master file there would be no problem in selecting an appropriate sample.

The master file can be used to develop data for finer geographic breakdowns than States and metropolitan areas at reasonable cost. For instance, we are currently experimenting with developing estimates for a few items by 5-digit zip codes or the 3-digit sectional centers through use of generalized programs. While it will not be possible to prepare extensive data long these lines, even the few

aggregate items of information will contribute substantially to analysis for marketing purposes, welfare programs and income studies.

CONCLUSION

As the filing date for 1966 tax returns approaches, I would like to point out that millions of Americans have been preparing a document full of extremely valuable economic data. Information of this type can be obtained only by means of expensive surveys or interviews. Even then the cooperation on the part of respondents would be somewhat less than perfect and the quality of the data open to question.

Since 1916 the Service has not only been economically collecting taxes but has also been economically producing statistics based on tax return data. As the scope of the income tax has increased, so has the value of the statistical output. As the methods and organization of the revenue collecting process have improved, the statistical program has changed accordingly.

The adoption of automatic data processing and the master file system have opened the way to great improvements in the statistical operation. Studies now going on point to faster, more economical production as well as more detailed data both as to geographical areas and subject matter. With the newly developed methods we can be even more responsive to the neds of users, both government and private.

APPENDIX II

EXECUTIVE OFFICE OF THE PRESIDENT,
BUREAU OF THE BUDGET,
Washington, D.C., October 1966.

REPORT OF THE TASK FORCE ON THE STORAGE OF AND ACCESS TO GOVERNMENT STATISTICS

Carl Kaysen, Chairman
Institute for Advanced Study
Charles C. Holt
University of Wisconsin
Richard Holton
University of California, Berkeley
George Kozmetsky
University of Texas
H. Russell Morrison
Standard Statistics Co.
Richard Ruggles
Yale University

The Committee was originally charged with the task of considering "measures which should be taken to improve the storage of and access to U.S. Government Statistics." It is the best judgment of the Committee that it can answer this question only in a much broader context, namely, by looking at the question of how the Federal Statistical System can be organized and operated so as:

1. To be capable of development to meet the accelerating needs for *statistical* information, needs that are increasing in quantity, in variety, and in degree of detail with the developing character of American society, and the changing responsibilities in it of the Federal Government;
2. To develop safeguards which will preserve the right of the individual to privacy in relation to information he discloses to the government either voluntarily or under legal compulsion;
3. To make the best use of existing information and information generating methods and institutions at its disposal; and
4. To meet these needs for statistical information with a minimum burden of reporting on individuals, businesses, and other reporting units.

The focus of the committee's concern is the Federal *statistical* system. Although different government agencies may require information about specific individuals or businesses as part of their legal operating responsibilities, the committee was unanimous in its belief that Federal agencies or other users should not be able to draw on data which is available within the Federal statistical system in any way that would violate the right of the individual to privacy. Organizational and legal safeguards should be developed to prevent the use of data which is brought together for statistical purposes as a source of information concerning individual reporting units.

A body of data can provide useful *statistical* information only to the extent that it is live, in the sense of corresponding to a clearly defined and currently comprehensible system of identifying the sources of information, definitions of

quantities being measured, classifications on which groupings of units are based, and the relations of all these categories to those for other information collected on similar units, or the same units at different times. Thus, no discussion of storage of and access to data can be usefully conducted without some consideration of the larger information system—from basic data collection to analysis—of which storage and access are a part.

1. BACKGROUND—THE PRESENT SYSTEM

At present, the Federal Statistical System is decentralized in respect to all its basic functions: collection, storage, analysis, tabulation, and publication. Twenty-one bureaus are shown in the Budget Bureau list of the "principal statistical programs" for FY 1967. Their total estimated budget, including the annual average over recent years of expenditures on periodic programs (mostly Census programs), was about $122 million, of which $96 million was for current programs, and the balance for periodic programs. The four largest agencies, with their shares of the total budget, were: Census, 24 per cent; Bureau of Labor Statistics, 16 per cent; Statistical Reporting Service, Department of Agriculture, 10 per cent; and Economic Research Service, Department of Agriculture, 10 per cent. Their total share was thus some 60 per cent, and the next four agencies—National Center for Health Statistics, Social Security Administration, Internal Revenue Service, and National Science Foundation, accounted for an additional 18 per cent, making a total share for the largest eight of 78 per cent. Decentralization has been increasing. A decade ago, the four largest statistical programs—those of Census, Agriculture (with the Statistical Reporting Service and the Economic Research Service operating as a single unified agency), Bureau of Labor Statistics, and Social Security Administration—accounted for 71 per cent of the total expenditures of the 11 Bureaus which had significant programs.

The increase in dispersion has occurred in a period of increasingly rapid growth in the total size of the System's activities. The total budget for 1956 for the 11 major agencies was $47 million, of which some $37 million was for current as opposed to periodic programs. In the period 1950–56, the (arithmetic) average annual rate of growth of expenditures for current programs was about 2.5 per cent; in the period 1957–60, nearly 7 per cent; for 1961–66, it has passed 15 per cent. Periodic programs are also increasing in scope and cost, and a projection of the order of $200 million for the 1970 level of expenditures for principal programs appears reasonable. Since many of the most rapidly growing programs have been those of new agencies, or agencies mounting major statistical programs for the first time, the process of further decentralization promises to continue, unless action is taken to change the trend. We do not mean to suggest that the opposite extreme of complete centralization of all data-gathering and analysis is desirable. As we explain below, even ignoring the difficulties of scrapping an existing structure and starting entirely afresh, a substantial amount of decentralization is inevitable and desirable, particularly in connection with the administrative, program planning, and program analysis functions of the operating agencies.

2. SHORTCOMINGS OF THE PRESENT STATISTICAL SYSTEM

The high degree of decentralization in all functions of the present statistical system has for some time been recognized as a major obstacle in the way of its effective functioning.

Nearly two decades ago, F. C. Mills and C. D. Long of the National Bureau of Economic Research made a study of *The Statistical Agencies of The Federal Government* (National Bureau of Economic Research, New York, 1949) for the Hoover Commission. They pointed to many problems arising from excessive decentralization and inadequate coordination. The major remedies they proposed included greater centralization—in the Census Bureau, and the creation of an Office of Statistical Standards with great powers to coordinate and unify that which was not centralized. These recommendations were followed to some extent, but the growth of the problem has out-stripped the strength of the remedies applied.

In March 1965, a committee of the Social Science Research Council, in a *Report on the Preservation and Use of Economic Data,* recommended the creation of a National Data Center, in order to remedy some of the most pressing problems arising out of the present statistical system. In a review of that report made for the Office of Statistical Standards (Bureau of the Budget) and completed in November 1965, Dr. Edgar S. Dunn, Jr., of Resources for the Future, endorsed the

substance of these recommendations, and in some respects went beyond them. Dr. Dunn was assisted in this report by a group of experienced professionals drawn from various parts of the Federal Statistical System, as well as by experts in automatic data processing of the National Bureau of Standards.

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.

The inadequacy of the present statistical system has three major aspects. The first is the lag between the receipt of information and its availability in usable form. This is most striking in the case of the Statistics of Income for Corporation Income Tax Returns. There is a one-and-a-half year lag between filing of returns and preliminary summary publication, and a two-and-a-half year lag before final detailed publication. A large part of the problem arises from the variation in filing dates of corporations filing on a fiscal year basis: some may file as much as 10 months after the end of the calendar year under which their returns are compiled. But part of the problem does reflect questions of priority and availability of facilities, and though these reports provide a basic source of economic data of great importance, their reporting function cannot be given first place in the administration of the Internal Revenue Service.

A second and deeper source of inadequacy in the present system is its widespread suppression of micro-information, and its orientation toward publication of necessarily aggregated and tabulated information as its major goal. These are of course intimately related: restrictions on disclosure to the general public or unauthorized persons within the government of information on individual reporting units is a necessary and desirable legal constraint on any official agency collecting information under the sanction of law. So long as publication is thought of as the basic process that makes information available for use, aggregation and the suppression and ultimate permanent loss of micro-information cannot be avoided. The consequence, however, is the necessity of substituting worse for better information, and cruder for more refined analyses, by those who use the data for research and policy purposes. In particular, much ingenuity and effort is spent in the construction of rough estimates of magnitudes and relations that could be measured with much greater accuracy, if the micro-information that present statistical records originally contained was preserved in usable and accessible form. Present technology makes it possible to do this economically and consistently with desirable limits on disclosure.

The growing decentralization of statistical programs has led to another major inadequacy. At the present time different agencies view the problem of the right to privacy very differently. In some agencies the policy of protecting the privacy of the information reported by individuals and businesses is formally stated and protected by law; in such instances the enforcement of such policies has also been found to be very good. In other instances, formal policies regarding disclosure have not been set up, and in many of these cases the protection depends on the judgment of those who are in charge of the different programs involved. Understandably, the growing decentralization of statistical programs has thus led to considerable unevenness in the nature and enforcement of disclosure rules. It is quite possible that without some overall policy which can be responsibly supervised major violations of individual privacy may take place. It should be the function of some group within the Federal Statistical System to ensure that data gathered for statistical purposes or obtained as a by-product of the administrative process is not to be used against an individual or enterprise. Thus at the present time information about individual persons or businesses collected by the Census Bureau cannot be used by the Internal Revenue Service or the Department of Justice against individuals or enterprises in the investigation or prosecution of such things as tax evasion or antitrust violations. This type of protection must be preserved in order both to protect the rights of individuals involved and to avoid falsification of information which might develop if individuals were not given assurance against disclosure.

The major elements of inefficiency to which decentralization has led are of three kinds. The first is duplication in the collection of information. Although the Office of Statistical Standards controls duplication, it is not always successful in eliminating it entirely. Avoiding duplication is especially important in that it needlessly spends not only money but the even scarcer resource of cooperation by the public; households, business firms, and other respondents, in answering enquiries. While duplication within single agencies is not serious, the great degree of decentralization leads to overlaps between programs of different agencies. The problem is less the collection of exactly the same information by two agen-

cies, and more the collection in two surveys or reports of data that could be collected in one. Failure to make the maximum use of each occasion for collecting information may well lead to a burden on respondents which becomes intolerable with growing needs for data. An example of the problem is provided by current practice in connection with sample data on retailing. The Bureau of the Census collects data on retail sales from one sample of retail stores and the Bureau of Labor Statistics collects data on employment, wages, and hours from another. As a result, there are doubts about the comparability of these input and output data at various levels of publication detail. These doubts arise not so much from the differences in the two samples as from differences in the two Bureaus' methods of assigning industry codes and definitions of reporting units. If both input and output data were collected on the same report form and processed by the same agency, these differences in comparability would be eliminated. This situation applies not only to retail sales but also to manufacturing data, where the Bureau of the Census collects monthly figures on sales, orders, and inventories, while the Bureau of Labor Statistics surveys manufacturing employment, man-hours, and wages each month. There is little doubt that a single consolidated reporting system, using one sample, would be both less burdensome, and less costly, and yield better information.

The second source of inefficiency is failure to use as a statistical resource all the information potentially available in the data collected. This, in turn, has a number of sources. (1) Collection of the data on the same reporting units by different collecting agencies operating with different classification systems, unit definitions, and the like, results in inability to match all the relevant available information on a responding unit for analytical purposes. Information on groups of respondents of different, and to some extent imperfectly known, composition cannot properly be compared and correlated. Census, IRS, SEC, and FTC data on business enterprises exemplify this problem. These incompatibilities in definition often reflect the different purposes of the several agencies that collect the data; yet effort directed to resolving these problems can be fruitful and is worthwhile. (2) After separate collecting and processing, agencies assemble data in summary form; the original individual reports are all but unavailable for further use, or available only at prohibitive costs. This effectively prevents different summaries and analyses of the data for other purposes by the same agency or by different agencies. In particular, the efficient use of data for intertemporal comparisons over any but a short time period becomes difficult, as the classifications change over time, and thus much information is irretrievably lost. (3) Confidentiality restrictions as interpreted by different agencies often act as a barrier to the full use of data for statistical purposes inside the government and within the legal boundaries of use.

The third source of inefficiency is that many of the smaller agencies operate on too small a scale to make fully efficient use of modern techniques, professional specialists, and economical large-scale machines. Only further centralization, rather than better coordination, can cure this situation.

The degree of decentralization in the system, and its predominant orientation toward publication as a means of making information available, correspond to a now-obsolete technology of handling and storing information, as well as to a much lower level of demand for detailed quantitative demographic, economic, and social information by policy-making agencies of all levels of government. Our present organization and mode of operation does not take advantage of modern information processing technology, and is not capable of meeting the variety and scale of present day information needs. The deficiencies of the system, and the gap between what it can provide and what would be technically possible under appropriate organizational arrangements will grow rapidly in the near future. As we have already pointed out, the demand for detailed quantitative information will continue to increase at a high rate. Further, the nature of the demand is changing in qualitative terms in ways that are only just becoming clear. The degree of disaggregation now demanded in the data relevant to economic policy has changed greatly in the last decade, even though the policy continues to focus on objectives stated in terms of such aggregate magnitudes as employment, unemployment, output, and the general wholesale and consumer price indices. The demand for comprehensive micro-data will grow explosively as policy becomes increasingly concerned with the micro-effects of the economic system, in terms of particular localities, income, and occupational, age and ethnic groups; as policy instruments become increasingly capable of sensitive and selective application to particular needs, and include a broader range of government actions in such areas as education, research, health, housing, transportation, and resource development. Further, the need for coordination of data collected

at state and local levels, in answering questions concerning specific small geographic regions, is also growing rapidly.

On the side of information processing technology, the last decade has seen great developments in machine processing, storage and transmission of information in machine-readable form. This progress is continuing both in the computing equipment itself (hardware) and in the programmed instructions for directing the machines (software). These developments place the problems of large-scale storage, integrated data files, rapid access, and confidentiality in an entirely new light. In particular, it is now possible, with sufficient effort, to create the capability for combining centralized processing and storage of large bodies of data with decentralized analytical use, subject to the restraints of a uniform system of limits on the disclosure of data on individual reporting units.

In pointing to the shortcomings of the Federal Statistical System as presently organized, the Committee does not wish to suggest either that those charged with its operation are unaware of these problems, or that they are making no attempt to find remedies for them. Quite the contrary. The Committee has the highest regard for the professional competence and dedication of the senior personnel of the major statistical agencies. We can say the same of the Office of Statistical Standards in the Bureau of the Budget, which is now charged with the coordinating responsibilities for the Federal Statistical System, as adviser to the Director of the Budget. These two groups are now making serious efforts to deal with the kinds of problems we described. Recent increases in the level of Census work performed for other agencies on a reimbursable basis is an example of one method of meeting these problems. The cooperative efforts of the IRS and the Census Bureau to use data from the income tax returns for the Economic Censuses is another. Further, the heads of the statistical agencies and the officials of the OSS have been of the greatest assistance to the Committee in making this report. However, the Committee believes both that insufficient resources are being devoted to dealing with the problem, and that the present organizational framework cannot generate improvements in the existing situation fast enough to cope with the growth of the problem. We conclude, therefore, that significant organizational change as well as increased effort are necessary conditions for a successful attack on it.

The building of an integrated body of data combining presently available sources of the appropriate kinds of data, which preserves in usable form the maximum detail of information, stored on tape or other machine-readable form, coded, organized, and indexed so as to be readily accessible, is the minimum step which must be taken to cope with problems sketched above. The existing agencies are now approaching this task slowly, with a scale of effort too small to ensure that it can be completed, and under a variety of inhibitions and constraints. The Census has taken a commendable lead, and already has done a number of useful tasks. However, this job is viewed both by Census and the other agencies as a second-priority activity, which cannot compete for personnel, machine time, or funds with ongoing current programs. This is natural, and indeed inevitable. Simple inter-agency jealousies and rivalries have also created inhibitions on prompt and full cooperation.

The Office of Statistical Standards is too remote from operating responsibility to move forward at the pace at which such integrating activities need to be carried on. It negative powers are, at least theoretically, great, but its ability to promote new programs is limited to what it can accomplish by persuasion in the face of institutional pressures which go in the opposite direction.

Finally, none of the constituent agencies of the System has given the problems described above or the opportunities to deal with them provided by advancing technology, the importance that this committee—following in the footsteps of several predecessors—assigns them.

3. WHAT IS TO BE DONE?

Were the Committee to be designing a Federal Statistical System *de novo*, it would clearly recommend the creation of a single Central Statistical Agency with the following responsibilities:

a. Collecting all systematic, general-purpose, large-scale quantitative information of a demographic, economic, or social nature, insofar as it is not produced as a by-product of the administrative operations of the Government.

The qualifying adjectives are meant to preserve the freedom of operating and policy agencies to make, when necessary, occasional surveys or special-purpose studies for their own purposes when the Central Agency was unable to provide the requisite information, although the Central Agency would have the facilities

and capacity to carry these out as well, on request. Further, the limitation of the Central Agency to dealing with general purpose information is made explicit in order to recognize the variety of operating needs for gathering and processing information of the several branches of Government, which they must be in a position to meet directly in order to carry out these tasks.

b. Receiving and integrating into its general information stock, data which are the by-product of administrative operations of the other Federal Agencies. In this connection, the Central Agency would, through liaison with the other agencies, help to design tax, regulatory, social insurance, and other report forms in such a way as to produce the maximum information consistent with reasonable burdens on respondents.

c. Developing and maintaining appropriate standards of confidentiality in the release of any data, as determined by law, using the basic data within the bounds of these standards so as to minimize the loss of analytically useful information, while at the same time ensuring protection to the privacy and identity of individual reporters.

d. Coordinating its activities to the greatest possible extent with those of the information collecting of states, cities, and other governmental units so as to arrange as far as possible a rational division of labor, a maximum integration of information, and a free flow of useful information in both directions.

e. Organizing and storing information in such a way as to provide maximum legitimate accessibility, for both governmental agencies and other users.

f. Providing computing, tabulating, and analytical capacity for all government users. Insofar as it proves economical, these facilities could be drawn on by other agencies for any special computations, analytical studies, etc. In other words, the Agency should maintain the central general-purpose large-scale data-processing center for general Government needs. Non-confidential data from the Center in standard documentary or machine-readable form should also be available, on a suitable compensatory basis, for the research uses of academic and other private groups, when such use serves a public purpose.

Neither this function, nor that of *a.* above is intended to preclude other agencies from maintaining independent computing facilities and independent analytical capabilities of an appropriate sort. In particular, it is obvious that every policy agency, as well as many administrative units, will require analytical capabilities for program planning and program evaluation. But current technologies, including distant consoles connected with a central computer facility by telephone links, make compatible decentralized use of data for analytical purposes, with highly centralized data storage and processing.

g. Studying methods of improving the protection of individual privacy and the confidentiality of data while at the same time providing use of it for legitimate analyses. Both the screening of analyses before release, and the camouflage of the basic data itself offer promising paths for exploration.

h. Improving methods of data collection, techniques of sampling, and opportunities for maximum use of by-product information, both for economy's sake, and to minimize the burdens on respondents of increasing demands for information.

i. In cooperation with the analysts who use the data, both within and without the government, defining and refining the standards and bases on which information is collected, and determining the probable development of information needs.

j. Improving the techniques of data handling, storage, and computation, in cooperation with appropriate technologically competent public and private agencies.

k. Securing the research and development contributions of university, business and other groups to the effectiveness of all these functions. This can be done both by making grants and contracts and by providing facilities and capacity which such research personnel could use on a variety of financial bases. Outside research and development assistance might usefully cover the whole range from long-term basic research to assistance in the solution of an immediate problem.

The Committee is not starting with a clean slate. Realistically, the question before us is how to proceed from the present situation of too much decentralization and insufficient coordination. We have not attemped to judge either the wisdom or the feasibility of attempting to create, at one blow, the kind of Central Statistical Agency we have described above. Rather we have sought to make a step, of sufficient magnitude to inject a genuinely needed new element into the system to help it adapt more rapidly to the growing problems it faces. Following this step we envisage further adaptive evolution in the direction of a stronger and more centralized system by an experimental process.

What first step is sufficiently large so as to promise a good prospect of further development? The Federal Statistical System has three basic functions; namely, (1) collection, (2) integration and storage in accessible form, and (3) analysis, tabulation, and publication. It is reorganizing the second that offers the most promise. This function is now the least well-performed of the three, and it is the one which is most easily separated out from the present organizational structure. However, it must be done on a substantial scale, and in such a way as to recognize the interaction of this function with the other two. Further, the new organization must not be confined to a merely archival function. If it is defined along the lines suggested below, it offers the best promise, in the judgment of the Committee, for starting the development of the Federal Statistical System toward a more integrated and efficient form.

Accordingly, the Committee proposes the creation of a National Data Center. This Center would be given the responsibility for: (1) assembling in a single facility all large-scale systematic bodies of demographic, economic, and social data generated by the present data-collection or administrative processes of the Federal Government, (2) integrating the data to the maximum feasible extent, and in such a way as to preserve as much as possible of the original information content of the whole body of records, and (3) providing ready access to the information, within the laws governing disclosure, to all users in the Government and, where appropriate, to qualified users outside the Government on suitably compensatory terms. The Center would be further charged with cooperation with state and local government agencies to assist in providing uniformity in their data bases, and to receive from them, integrate into the federally generated data stock, store, and make accessible, the further information these agencies generate. The funding and staffing of the Center should recognize both these functions.

In more detail, the functions of the Center would be:

(1) To establish and maintain an inventory of all available data in the relevant categories in the Federal System.

(2) To set and enforce uniform disclosure standards so that the legal requirement of confidentiality can be met with no unnecessary sacrifice of analytically useful information.

(3) Similarly, in cooperation with the state and local government units, to perform similar tasks for information generated at those levels of government.

(4) To assemble centrally the data from all these sources, integrate it to the maximum feasible extent, and preserve it in usable and accessible form. This will involve:

The maximum ability to exhibit the interrelations of various kinds of data.

The preservation of detail in basic records, and the avoidance of the loss of information in the storage, manipulation, and retrieval of information.

The ability to produce the full measure of inherent information which is computable from the basic records.

(5) In cooperation with users in and out of government and collection agencies, to set the standards for further collection efforts, so as to make maximum use of administrative information and provide maximum cross-linking of different bodies of data.

(6) To provide facilities—from working space to linked input-output consoles—for major users within government to facilitate their access to the data and improve their ability to work with it.

(7) To develop software and hardware, especially input and output devices.

(8) To define the regulations and compensation arrangements under which non-government users would have access to data in the Center. In general, subject to disclosure restrictions, standard tabulations and tapes could be made available at cost to private users for research and analytical purposes. However, the Center should not become a service bureau or data-processing agency selling special order analyses to private users in competition with firms and individuals in the information processing industry.

In full operation, the National Data Center would provide the following benefits:

(1) Reduce the collection effort and particularly the burden on respondents required to secure a given amount of information.

(2) Improve the protection of individual privacy by developing standards of disclosure, techniques of preserving confidentiality and supervision of enforcement of disclosure rules.

(3) Preserve for continued use all or nearly all the relevant detailed information contained in the original data, as compared with the present situation in which much of the detailed information is irretrievably lost, or become retrievable only at prohibitive cost.

(4) Reduce the processing costs associated with the use of a given amount of information.

(5) Store information in more accessible forms at lower per unit costs and with a comprehensive index or bibliography.

(6) Make much information accessible to non-Government users which now is too expensive or too cumbersome for them to use, even though it is legally available and its use would benefit the general public. This is especially relevant to users in state and local governments, academic and other non-profit research users, and business users. In this connection, the Center should develop extensive working relations with academic users, of the sort which the Census has done to a much greater extent than other agencies. Even these are limited and currently are handicapped by lack of physical facilities, programming capacity, and organizational capacity for dealing with them. As the working relations of the physical science establishments of the Federal Government with the academic and industrial scientific communities demonstrate, such cooperation is of great benefit to the Government in performing its tasks effectively.

(7) Provide improved analyses of existing data for all users. The facilities for cooperative efforts are highly relevant to this point as well.

(8) Facilitate greatly improved coordination of statistical data between the Federal Government and the states and localities, and internationally as well.

(9) Create a repository of technical competence in statistical services, and computer software and hardware, that would be available to the whole Government establishment.

4. PROPOSED ORGANIZATION

In order for the National Data Center to function properly, it must be given a proper position in the Federal Statistical System, and sufficient authority, leadership, trained personnel, and funds to perform its mission. The Committee has given special attention to the problem of finding the organizational arrangement most conducive to the successful functioning of the Center, and attaches great weight to its organizational recommendations. We recommend the creation of a new position, Director of the Federal Statistical System, in the Executive Office, and the placement of the Census Bureau and the National Data Center as coordinate units under his direction.

The Bureau of the Census is the largest, most widely experienced, most professionally competent, and broadest in scope of all the present statistical agencies. To the extent that any agency in the System attempts to perform the functions described above, it is the Census. The data Center will require close cooperation and support from the Census in order to function effectively. For all these reasons, it appears desirable to put the Center in close organizational and physical proximity to the Census. On the other hand, the Center's tasks are not the present tasks of the Census; the Center will be a new organization with the difficult problem of establishing itself as a going concern and making its way in the complex of agencies producing and using large bodies of quantitative information—its suppliers and customers, so to speak. Thus it does not appear appropriate to subordinate the new agency to the existing Census organization. Further, the establishment of smooth working relations between the Center and the other elements of the Federal Statistical System might well be easier if the Center is a new, coordinate agency rather than a part of the Census Bureau.

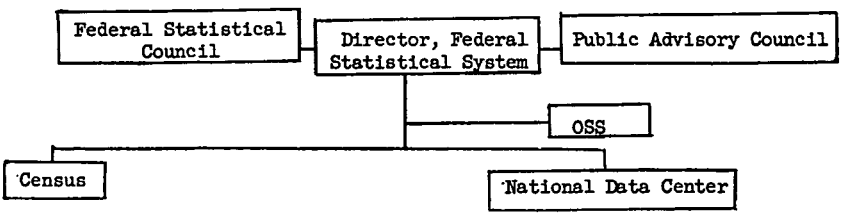
If the new data Center is to have specially close but coordinate relations with the Census and similar, if organizationally less intimate, relations with other data collecting and using agencies, some method must exist to regulate and oversee these relations. This coordination function is now assigned to the Office of Statistical Standards, but in the Committee's judgment, that Office is not placed so as to be able to carry it out effectively. We propose that a new position be created with this function, entitled Director of the Federal Statistical System, to be filled by presidential appointment. The new Director would exercise, by delegation or new legislation, as seemed appropriate, the coordinating powers over Federal statistical programs provided for in Sec. 103 of the Budget and Accounting Procedures Act of 1950 (P.L. 784, 81st Congress). The Office of Statistical Standards would accordingly be transferred from the Bureau of the Budget to become a staff office of the new Director, to assist him in carrying

out these responsibilities. The Census Bureau and the National Data Center, each under its own director, would report to the Director of the Federal Statistical System.

In addition, we propose two councils advisory to the director. The first is a Federal Statistical Council, representing the major data producing and using agencies in the System and reporting directly to the Director. Thus the Commerce Department might be represented by the Assistant Secretary for Economic Affairs, the Council of Economic Advisers by one of its members, etc. The second would be a public advisory council, with members from outside the Federal government representing both the public in general and particular users of information such as business, labor, state and local governments, and the academic community. This council would advise the Director particularly on such matters as the burden on respondents, the protection of confidentiality, and the satisfaction of user needs. The uses and possible abuses of information collected by the government are so important in our society that continuous public scrutiny of these problems at a high level in the Federal System is desirable.

The proposed new office, with its two operating elements and its government-wide coordinating functions could not readily fit into any of the existing Cabinet departments. Rather, its natural home would be the Executive Office of the President, and the Committee recommends that it be placed there.

The organizational relations of these elements are shown in the chart below.



The National Data Center itself might be organized in two main branches: an operations division, and a research division which would consider sampling methods, analytical and computing techniques, methods of protecting confidentiality, and liaison with extra-governmental research enterprises and with users. The research division could also cooperate with the relocated Office of Statistical Standards for data collection and for integration of existing bodies of data. The operations division might be so organized as to provide for the inclusion of sections on assignment from both the major data producing agencies, such as the Census itself, the Bureau of Labor Statistics, the Internal Revenue Service, etc., as well as from the major users within the government, such as the Office of Business Economics, the Council of Economic Advisers, the Tax Research Division of the Treasury, etc. These transplanted sections would greatly facilitate the free flow of information to and from the Center, on which its entire function would depend. They could also take responsibility, as a transition measure, for seeing that the confidentiality requirements of their own agencies were properly applied to the data collected by them.

The provision of space, funds, technical personnel, and machine capacity in the Center must from the first be such as to allow a good deal of flexibility in meeting demands on it by both governmental and non-governmental users. Too close a calculation of capacity, especially in machine time and programmers could prove fatal to the Center's ability to establish itself. So could too early an expectation of visible results.

It would be the major responsibility of the Director of the Federal Statistical System to see that the proper division of labor, coordination of information, and utilization of the Data Center were made by the constituent agencies in the System. In the furtherance of this responsibility, he might be asked by the Budget Director to review for him the statistical budget of each agency, in much the way that the Director of the Office of Science and Technology assists the Budget Director in reviewing the science budgets of each agency. Yet the basic functions in the System would be determined more by what might be called "market forces" than by the fiat of the Director. To the extent that the statistical agencies under his direction provided quicker, cheaper, and better sources, to meet rapidly expanding demands, the rest of the government would increasingly rely on them, rather than on the expansion of their own in-house

capabilities. The Committee is strongly convinced that the proposed new organization would encourage such a path of development. We do not believe that a mandatory reshuffling of the existing tasks of the other statistical agencies is either necessary or desirable, in the face of what we are confident will be great increases in demands on all statistical agencies, provided only that the new organization is started on a sufficiently firm basis to permit it to survive the inevitable birth trauma.

5. INITIAL STEPS

A prerequisite to the creation of the National Data Center is a re-examination and consolidation of the laws and rules governing intragovernmental disclosure of information on individual respondents, and their recodification in terms which will make the operation of the Center feasible. The Committee is neither sufficiently competent nor informed to judge whether this will require new legislation, or whether it can be done by executive authority within the framework of existing legislation.

As soon as the Center is created, it should be put to work with existing personnel, and machine capacity borrowed chiefly from Census, to deal with those tasks of inventorying data and creating integrated files which are already most advanced. A sufficient beginning has been made—as detailed in the Dunn report and the appendices thereto—to permit the Center a running start.

Initial funding for the Center must be large enough to ensure its viability, and to attract to it good people from within and without the government. As a crude guess, the Committee suggests an initial budget of the order of at least $2 million per year, with the prospect of rising to $5 million within three years, exclusive of buildings and computing equipment. A generous allotment of supergrades is as important to the new agency as money, since a large proportion of its tasks, especially in the initial stages, will require high technical and professional skills.

ANNEX: THE RIGHT TO PRIVACY, CONFIDENTIALITY AND THE NATIONAL DATA CENTER

After this Committee was convened and well into its work, Congressman Cornelius E. Gallagher, as Chairman of a Special Subcommittee on Invasion of Privacy of the Committee on Government Operations of the House, has raised questions about the possible threats to privacy and freedom that a National Data Center might present. These are serious questions, that deserve to be met squarely.

In general, our Committee believes that the problem of the threat to privacy can be met best by Congressional action, which defines a general statutory standard governing the disclosure of information that is collected on individuals either as a by-product of administrative, regulatory and taxing processes, or through Census or sampling procedures. The Director of the Federal Statistical System should then be given the responsibility for monitoring compliance with this standard, not only by the Data Center, but by all the parts of the System.

The problem of disclosure of confidential information about individuals and businesses is not new. It has long been recognized that the information which individuals and businesses provide under law to the Bureau of the Census, for example, is confidential. This means that no other Federal agency is permitted to see or use the individual records, and even Congress itself cannot obtain census information on any individual or company. In fact, this confidentiality has been guarded so zealously that Congress and the other agencies of the Federal government have been enjoined from obtaining from companies duplicate copies of those records which were submitted to the Census Bureau. The disclosure rules are meant to safeguard individuals so that they can feel sure that information which they give to the Census Bureau will never be used against them for such purposes as tax enforcement, antitrust, or Congressional investigations. The disclosure rule has not been interpreted, of course, as preventing the use of Census information for analyzing policy or providing information about specific groups, regions of the country, performance of industries, etc. In making tabulations of data, however, the Census Bureau carefully omits those classifications which might enable anybody to figure out information about individual firms or persons.

There are, of course, other Federal agencies which must by their very nature use information about individuals and firms for their operations. Thus for example the Internal Revenue Service not only must collect information about people's income and the taxes they pay, but this information can and should be used to

prosecute tax fraud or tax evasion. Similarly, the Social Security Administration must process information about each individual over a period of years, recording his job status, family status, etc. This information is necessary for the determination of social security payments. Such use of individual information is of course justified, necessary, and legal. On the other hand, it is a real question whether tax returns or social security records should be turned over to other groups who may wish to use them for other purposes if the persons or firms to whom the records refer may individually be affected thereby. The question of the proper or improper use of information by different agencies is indeed a ticklish one, and procedures should be developed by both the executive branch and the legislative branch which will protect confidentiality and insure the privacy of the individual. In a great many instances, agencies may wish to obtain information not for operating purposes, but in order to make policy decisions and to guide future operations. Thus the Office of Education has a real interest in knowing how college enrollments may be expected to develop in the future. Those concerned with questions of poverty wish to know the dimensions and structure of this problem. In a great many of these instances, the agencies in question have contracted with the Census Bureau to provide them with such general information based upon sample surveys. In these instances, a disclosure and confidentiality rule must be developed which will protect the individual and yet yield the general information which is required.

The enforcement of a statutory obligation as the primary method of dealing with the problems of safeguarding privacy can work excellently, as the experience of the Census Bureau shows. Indeed, the present situation, in which there exist a variety of different disclosure standards, some statutory and some executive, is much less conducive to protection of individuals' privacy than would be a situation in which, as our report suggests, the Director and the Data Center would have the obligation of enforcing a uniform standard over the whole system.

The Subcommittee has also raised the question of the creation of a vast file of individual "dossiers" incorporating police and FBI information, Armed Service and government personnel records, and the like. This is not the purpose of the proposed Center at all, and it is clearly within the power of Congress to distinguish between the collection and organization of general economic, social, and demographic information of the sort that Federal statistical agencies have traditionally collected—much of it on a sample basis—to which our proposed National Data Center is directed, and assembly of the sort of personal history information on named individuals that is contained in a personnel file or police file.

Finally, the Subcommittee has raised certain questions as to the technical security of data stored in machine readable form, and accessible through machine operations. Here again, this is not a new problem, and both organizational and technical means are available to control and limit the risks. Though bank robbers have not been totally eliminated, we have not on that account abandoned banks and banking, and the analogy seems to be perfectly appropriate. We think that the maintenance of privacy against both unwitting and illegal disclosure of information made available to the Government are real problems, to which our proposed new Center must direct attention and effort. However they are neither insoluble problems, nor ones of such magnitude as to make the organization and effective functioning of a National Data Center possible only at the expense of significant inroads on liberty and privacy.

APPENDIX III

COLGATE-PALMOLIVE COMPANY,
New York, N.Y., May 23, 1967.

HON. HERMAN E. TALMADGE,
Subcommittee of Economic Statistics,
Joint Economic Committee,
Congress of the United States,
Washington, D.C.

DEAR SENATOR TALMADGE: The testimony of Mr. John H. Aiken of the Federal Statistics Users Conference and Professor Frederick F. Stephen of Princeton University on May 18, 1967 before your Subcommittee made it apparent that a clear understanding and definition of available economical statistical resources, their uses, economic statistical needs and potential uses are needed.

Accordingly, I suggest that your Subcommittee consider proposing to Congress a comprehensive survey of:

1. Economic Statistical Resources and Services.
2. Description of actual uses and actual application
3. Description of users
4. Statistical needs
5. Description of potential users and potential applications among non-users

A survey of this kind would in part follow through the requests received by various agencies to determine the use to which the statistics were applied and possibly their contribution to a specific action. In order to sample properly the various services, institutions, population groups and resources I recommend that Census Bureau statisticians be consulted in view of their proficiency in sampling.

National Data Center

The need for a central file for statistical data is obvious to me for it would make available in relatively convenient form the statistical resources of the nation and minimize, if not eliminate, any duplication that may exist. A National Data Center would perhaps be best as a coordinating body rather than a controlling one. The individual agencies would continue to collect, analyze and control their data while filing with the National Data Center.

The Library of Congress would seem the most appropriate governmental body for maintaining the Central Computer file for data.

Individual Confidence

Much of the data collected for government and other institutions are dependent upon the voluntary cooperation of individuals and the organizations they represent. Legal safeguards should be provided to insure cooperation for surveys to be used for statistical purposes. The Bureau of Deceptive Practices of The Federal Trade Commission can provide you with further background information on the potential dangers when these confidences are violated.

With the increased electronic data processing and growing sophistication in statistical techniques applied to government, military, business, education and welfare and agriculture our needs are becoming such that the costs of conventional collection and analysis of data may far exceed our ability to meet them. For this reason and the pressing problems in every sector of the nation, dependent upon impartial data, in need of solution I cannot stress enough the importance of your committee and how appreciative many of us are of your efforts.

Very truly yours,

ARTHUR KOPONEN.

○