Volume Title: The Role of the Computer in Economic and Social Research in Latin America

Volume Author/Editor: Nancy D. Ruggles

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/rugg74-1

Publication Date: 1974

Chapter Title: The NBER Time Series Data Bank

Chapter Author: Charlotte Boschan

Chapter URL: http://www.nber.org/chapters/c6614

Chapter pages in book: (p. 33 - 56)

THE NBER TIME SERIES DATA BANK

CHARLOTTE BOSCHAN*

National Bureau of Economic Research

1. INTRODUCTION

One of the favorite occupations of practicing economists is forecasting aggregate economic activity and its components. Many of these forecasts involve short-term predictions of major indicators of aggregate economic activity, either as final products or as inputs to a "dependent" forecast of other, disaggregated minor variables. The methods used vary widely. Some base their forecasts on the indicator approach, some on econometric models, ranging from single equations to very complicated and elaborate multi-equation models, some on a sectorial approach, some on intuition, and some on a combination of different approaches. Whatever the methods, almost all forecasters use time series of economic variables, which are collected, maintained, and published by a wide variety of government and private agencies. If each individual economist had to keep track of these data—of their revisions and updates, changes in definition, lack of comparability, new information and discontinuances—the resources spent on such housekeeping tasks would be enormous.

For a large number of time series these tasks were facilitated by the data collections and analyses published in the *Survey of Current Business* which, for many years, provided the main source of current economic and business information. Since 1961 the Census Bureau's monthly publication *Business Cycle Developments*—now named *Business Conditions Digest*—has provided numeric and graphic information for the analysis and forecasting of current business conditions. The organization and the approach of this publication is an extension of the National Bureau's work on indicators, diffusion indexes and recession-recovery analysis. BCD contains adjusted numerical information for the three most recent calendar years, graphs from 1950 to date, certain statistical summary measures and some analytic measures. This publication serves a very useful purpose and is indeed very popular among business economists and other forecasters. However, it still has several shortcomings. Publication and distribution of such a data collection necessarily involves delay, and thus some time series may be out-of-date by the time BCD reaches the economist's desk. Also, the number of time series published is severely limited and may not include time series important for given users. Finally, the data are not immediately in machine readable form, ready for input into computer programs. A data bank residing in a time sharing system, in which series are updated and revised as soon as new data are available from the publishing agency, is free of many of these shortcomings. The latest data are always in the bank and available to users. Series important to users can be stored, listed, plotted, and used as input into canned general purpose programs, or into the user's own

special ones. Such a time series data bank, operated and maintained by the National Bureau of Economic Research, is the subject matter of this paper.

The most interesting aspect of maintaining a data bank is the number and variety of problems that arise and must be solved. Hence this paper is largely problem oriented. First, there are problems arising in connection with the selection of series to be included, their time coverage, periodicity, and seasonal adjustment status. Second, the National Bureau has a special problem in having to deal with several time sharing systems simultaneously and to fulfill a unique obligation of being scrupulously nondiscriminatory between the small and the large, the good and the bad, the wise and the foolish. A third set of problems are operational and revolve around checking and updating, having the very latest information at the very earliest moment, and generally maintaining time series on-line. After discussing the documentation for such a system (the minimum requirements as well as the ideal envisaged for the future) we will give an overview of the economics of running the venture. We hope that a description of the problems we faced and are still facing, as well as an indication of how we tried to solve them, may be helpful to others who want to set up, maintain, and use similar banks. We also hope that readers who can suggest better solutions than the ones we tried will give us the benefit of their thoughts.

Before we start with the description of the specific problems, a short explanation of the historical and institutional background of the particular data bank is in order. Early in 1967 a group of economists from about twenty large New York firms, all engaged in short term forecasting, became aware of the extent of duplication in their efforts. Therefore, they started an experimental cooperative venture, called "Project Economics," consisting of a time series data bank on General Electric's Mark I system. Each member was assigned a set of time series to be maintained in the bank, with updates and revisions as published by the respective government agencies. Manipulative computer programs, largely written by General Electric, were pooled and kept in common storage. This scheme worked tolerably well for a while; but problems arose: data handling expertise varied among members, quality control of the data in the bank was never fully developed, large revisions became too cumbersome for members to enter, and the quality of the data deteriorated.[1] This is when the National Bureau of Economic Research entered the picture. The provision of economic data to economists has long been an accepted function at the National Bureau. Data developed and selected by the Bureau play a prominent role in such government publications as *Historical Statistics, Business Conditions Digest*, and *Indicators of Economic Growth*. The Bureau has published several source books, such as the second volume of *Business Cycle Indicators* (1958) and a volume on construction statistics (1966), and has always made its unpublished material freely available to other economists. Also, we felt that, in the long run, nobody else could insure the standard of accuracy that we think appropriate. Thus it was a natural step for the Bureau to develop and maintain a machine readable data bank. In 1969 the Bureau took over the maintenance of the Project Economic's data bank, expanded its coverage and made it available to other time sharing systems and user groups.

[1] For a description of the original Project Economics see Robert E. Lewis, "Operating a Cooperative Economic Data Bank on a Time Sharing System," *Business Economics*, May 1970.

2. CONTENT OF THE DATA BANK

The two main purposes for which the data bank was designed are the analysis of current business conditions and short-term forecasting. These purposes have several implications for the content of the bank: the bank contains practically all indicators, aggregates, and other time series which may be considered important in evaluating current business conditions; most series are either monthly or quarterly, with a few annual ones included for the convenience of model builders; all time series start in 1946 or later, whenever they become available (none are carried before 1946); all series which have a seasonal component are carried on a deseasonalized basis; the series in the bank are updated as soon as new or revised estimates become available from the primary source.

Time Series Included

In the pursuit of the broad goals outlined above a number of problems arose and a number of decisions had to be made.

(a) What size should the data bank be? For instance, should all or most available time series be included? To the uninitiated user there may seem to be definite advantages in having a large number of series in the data bank: a bank which has twice as many series appears to be twice as good. However, a smaller collection of series is not only less costly to store and easier to update but it also takes less time to access and is more economical to use. Thus we decided to include only series which are clearly useful and which would, in fact, be used. It turned out that this was one decision which was easier to make than to implement—which leads directly to the next question:

(b) How should series be selected for inclusion in the bank? During the first stage of its existence, time series to be included in the data bank were chosen by a committee of users. Later, for a short period of time, we included almost anything anybody wanted, as long as it could be called a cyclical indicator or was part of the National Income and Product Accounts. This procedure had several obvious disadvantages. We could not make any use of the economies involved in updating a number of series from the same source. Code names were determined on an ad hoc basis, without reference to names for other series, let alone a systematic nomenclature. Worse than that, we might find ourselves maintaining esoteric regional or industrial subclassifications, which some user may regularly need (or may have needed for a special purpose in the past), without carrying the more important aggregates in the bank. We, therefore, soon became very careful about inclusion of further series. At present we are working with a "dictionary" or "directory" of significant time series which contains, in addition to all the series presently in the bank, many series we consider to be important enough to warrant inclusion, provided enough users express an interest in them. These series are already ordered by source. Also, orderly and meaningful code names have been determined for all of them.

During the past year we have continuously added new series. Now that we think we have a fairly reasonable selection, we will only add new series if new statistics are compiled and published by issuing agencies (job vacancy statistics of the BLS, for example, and the new production series by the FRB) or if we think

that existing series, or series which can be constructed, have become important (deflated indicators are a good example or, if many users need them, the flow of funds series may be a case in point).

(c) How should series, now in the bank, be selected for deletion? This problem is difficult to handle. Since none of the systems we are using is designed to keep track of how often a particular series has been accessed, it is quite likely that some of the series which we painstakingly bring up-to-date have never been used and will never be used at all. In order to avoid this waste we are planning to experiment with several schemes such as transitory second class storage (where a user must request series before he can use them and where one can monitor the rate of usage), low priority updating, announcement of plans for discontinuance, and perhaps some others. Users would be invited to comment and we hope that their reactions would give us some clues to the importance they attach to individual series.

(d) What should subscribers do if they need series which are not of general interest and, therefore, not included in the bank? This problem is not difficult since in all the systems we are using—and probably in all time sharing systems in which one can use a data base—it is possible to maintain private files of series and use them in conjunction with data in the bank. As a special service, we sometimes provide subscribers with exact source information for data they want to keep on line and update privately; occasionally we even give them the actual figures.

At present the bank contains some 2000 time series, covering all U.S. series shown in the Census Bureau's publication *Business Conditions Digest*, almost all series shown in the Council of Economic Advisers' *Economic Indicators*, and all series included in the Bureau of Economic Analysis' "National Income and Product Accounts" as published monthly in the *Survey of Current Business* as well as most of the national income series published only in July. The series can be roughly classified as follows:

	Number of Series
National Income and Product Accounts	950
Manufacturers' Shipments, Inventories, Orders	140
Retail and Wholesale Trade and Inventories	30
Construction and Housing	30
New Plant and Equipment Expenditures	30
Industrial Production	70
Population, Labor Force, and Employment	120
Productivity and Unit Labor Cost	30
Prices	80
Financial Series	220
Balance of Payments	30
Federal Fiscal Operations	20
Miscellaneous	250
	2,000

Time Coverage and Reporting Span of Data

Since the data bank was originally conceived to be a tool of current conditions analysis and forecasting, none of the time series start before 1946. For an on-line time sharing data bank this is probably sufficient, and longer series might be difficult to handle. However, there is no compelling reason why the Bureau's collection of pre-World War II data should not be available on magnetic tape, to supplement data bank information. Many of these series provide back data for series presently included in the bank, and may constitute the best available time series for the earlier years. They could be used for longer period analysis, for an analysis of the stability of relationships, for comparisons of pre- and post-war cyclical and other behavior, and for a variety of other research purposes.[2]

At the moment the data bank contains 1350 monthly series, 500 quarterly, and 150 annual ones. Most annual series are part of the National Income Accounts, some 25 series are population and labor force data. For many purposes it is desirable to work with data of uniform periodicity. The software provided by time sharing companies usually includes functions which take quarterly averages or totals from monthly series, and annual averages from quarterly or monthly series. We are providing straight line interpolation subroutines and are in the process of incorporating a spline function interpolation. This function has the property that the arcs, on which the interpolation is performed, join smoothly.

Seasonal Adjustment

For the analysis of current business conditions, particularly if the "indicator approach" is used, one obviously needs time series which are free of seasonal fluctuations. We are, therefore, carrying almost all series which have a seasonal component in seasonally adjusted form. However, model builders often prefer to use unadjusted data, with dummy variables to represent the seasons; and for current business conditions analysis it is desirable to know the magnitude of the seasonal fluctuations, their variability, and the reliability of seasonal indexes and seasonal adjustments. But the inclusion of all original data and major analytical measures would, of course, more than double the amount of work and space needed. Hence, we decided, early in the life of the bank, that the cost of carrying all original data on line is too high and included only a few series in their unadjusted form. (Whenever only original data are published by the primary source, we usually provide our own seasonal adjustment and carry both versions in the bank.)

This decision does not mean, however, that we are not searching for some solution of the problem. The question of the reliability of seasonal adjustments occurs too often to be ignored. Perhaps we should include original data for all important series which have a large seasonal component and for all those whose seasonal factors are unreliable, fluctuate much or are in any other way problematic. Another possibility would be to include recent seasonal factors and confidence ranges or other variability measures of the seasonal factors in the description of

---

[2] If enough internal and external demand exists to make key punching and other chores worthwhile, we would certainly consider such a venture. The data would not need to be brought up-to-date and the probability of revisions is extremely small. Thus, the production of a machine readable tape would be a one-time task.

series. Such variability measures are, of course, quite problematic when we deal with moving seasonal patterns.

We cannot ignore the needs of model builders either, since TROLL,[3] one of the systems which uses our data bank, was especially designed to facilitate model estimation and simulation. We will probably provide them with a selection of unadjusted series on a low-priority update basis.

Confidential and Proprietary Data

Our bank includes several "confidential" series, originating in the Bureau of Economic Analysis, which may be used for analytic purposes in combination with other series, but may not be published. Before we release the code names and thus permit retrieval of these series to individual users, they must agree not to publish these data. We also carry several sets of proprietary data of Dun and Bradstreet, the Conference Board and the F. W. Dodge Corporation. We have permission to carry these series but, again, users must not publish them.

Discontinuities, Overlaps, and Missing Observations

Economic time series frequently contain discontinuities of various kinds. The underlying sample may change from one period to the next, concepts are improved or updated, the geographic coverage changes, or the method of collection varies. Statistical agencies sometimes carry these data with an overlap, sometimes with a footnote indicating the break and including perhaps a few observations on both bases; sometimes two separate series are published. As far as carrying these series in the data bank is concerned, we have conflicting goals. On the one hand we want to preserve data in the "purest" possible form, that is without any manipulation. On the other hand, we want to make time series easy to use in statistical analyses of various kinds, and for this purpose they must be continuous over the period of analysis. We could, of course, let each user do his own splicing, estimating, and manipulating of segments, but we prefer to avoid duplication of effort, use of several, slightly different, spliced versions by different subscribers, and waste of time we might have to spend in advising people. We therefore decided to compromise. Whenever possible we carry both segments: the old segment to the latest available date, and the new segment from the earliest available date. Continuous with the new segment we carry the old one, spliced and "adjusted" to the level of the new. The particular splicing method used is, of course, fully documented. A sophisticated user can do his own thing, while the analyst in a hurry has a ready-made adjustment.

Our treatment of missing observations is somewhat less exemplary. In the few cases of this kind we provide interpolations. However, because of the inability of our software to deal with flags and footnotes, a user has to consult the source notes before he can tell that a particular value represents our estimate rather than an original observation.

3. Operational Problems

Any data bank run by the National Bureau, a non-profit publicly supported organization, must be accessible to anybody who wants to use it. Since we do not have a large enough computer system of our own—or, at least, did not have one at the time the data base was set up—and since we cannot give an exclusive franchise to any one computer manufacturer or time-sharing service, it follows that the data bank has to be accessible through different time sharing systems; and it must be available on magnetic tape in a "standard" easy-to-read format, for batch processing in any kind of system.

Participating Time-Sharing Systems

At the present time our data base can be accessed through three time sharing systems: Rapidata, General Electric MAP, and TROLL. A fourth one, Interactive Science Corporation's PDP-10 system, is in the process of developing software for using and updating our time series efficiently. Other groups are planning to participate.

The problem of dealing with several not always compatible systems may be unique to our particular set-up, since producers of data banks usually deal with one system only. It is obviously very expensive and time consuming to enter additions and revisions by hand via a system's editor, for each of the varying systems. We have solved this problem by writing "update" programs for each system, which can interpret the same paper tape input. Thus, the paper tape is punched once and read into each system. Once it has been read, the update program in each system takes care of the rest. The provision of an update program able to interpret our standard paper tape is now a condition for the participation of any new time sharing system in our data bank. One other condition for adding a new system is that there be at least ten paying users, since it would otherwise be too costly to update—even with the same paper tape input.

Series Code Names

Historical accidents sometimes determine how systems will operate in the future. In the General Electric Mark I time sharing system, the first system in which our data bank was used, each time series constituted a file and each file was retrievable with a six-letter file name. This is the origin of our system of six-digit code names for each series. Since most computers can handle six characters in a word, there has been no reason to change.

In naming our time series we have attempted to attain some hierarchical order and some consistency, apart from mnemonic considerations. Thus, all series starting with G are part of the GNP—National Income Accounts. All financial series start with F, all industrial production series with I, all labor series with L, and so forth. All annual series have the letter A in the second position and most series which are not deseasonalized, the number 6. Whenever possible, industrial subclasses carry the SIC code number.

Although the six letters obviously do not permit a lot of flexibility or a really useful structure, at least our naming conventions insure a certain amount of

orderliness and facilitate manipulation : When the data bank became so large that the Rapidata system had trouble accommodating all series in one "library," we simply put all National Income Account series into a separate library, and changed the software to look in the second library for all series starting with G. This principle can, of course, be carried further if necessary. Since, in the Rapidata System, individual time series form individual files on a disk which are accessed directly by the system looking through a table to find their address on a disk, access time can also be cut by using several separate "libraries."

The TROLL system permits a much more sophisticated explicit hierarchical structure, which should not only speed up retrieval but also permit the user to locate series whose code names he does not know.[4] Time series in TROLL can be stored, referred to, and retrieved by using an explicit decision tree. One can have as many "nodes" as one wants. Obviously, the system will not be better than we design it, and we plan to work on this as soon as more people use our data base on TROLL.[5]

Internal File Organization

Originally, the last lines of each file contained the name of the time series, the units of measurement, and some other information. Later on these last lines were dropped so that each time series could be read entirely, to the last available figure, by a retrieval program, without explicit specification of the length of the series (which was not always known by the user). Instead, a "directory" or "dictionary" was provided which contained the code name of the series and other pertinent information and which could be accessed on line. We have since had occasion to regret the decision to have files which contain nothing but numerical information. When we wanted to "export" the bank from one system to another, for example, we were unable to just dump the contents of the disk on a magnetic tape, since the name of the file would not be dumped and there would have been no way to identify the numbers. Instead we had to write a program which first read a file name from the index, then searched for that file on the disk, and copied it on magnetic tape—together with the pertinent information from the index. Also, for listing and plotting time series it is more convenient to have the title together with the numerical information.

Fortunately, Rapidata is now in the process of changing its software, so that it can read several lines of non-numeric information (including the number of such lines) in the beginning of a file. This means that identifying information can now be carried in the beginning of each series. We hope that other systems will follow suit.

On the Rapidata system, the numerical information within each file is organized as follows: Monthly series are stored and listed six items per line with 3-digit identifying line numbers. The first two digits of the line number indicate the year and the last digit (1 or 2) indicates whether the data refer to the first or second

[4] For a description of how this can be accomplished on the Rapidata system see *NBER Data Bank for Project Economics*, National Bureau of Economic Research and Rapidata, 1971.

[5] For a description of the theoretical background of the methods and algorithms used in the estimation capability in the TROLL/1 system see Mark Eisner and Robert S. Pindyck, "A generalized approach to estimation for the TROLL/1 system," April 1971, mimeo., NBER.

half of the year. For example, the line number of January–June, 1956 would be 561. July–December of 1956 would appear on line 562. Quarterly data are stored four items per line with a 2-digit line number indicating the year, and annual data are stored one item per line with a 4-digit line number.

Routine Updating and Checking Procedures

Up-to-dateness is a very important aspect of our data bank. Our contract specifies that all new and revised data be incorporated into our bank within two working days of their official release. We are located in New York City and most governmental data are generated and published in Washington, D.C. This creates some problems. If we relied on ordinary mail service we would sometimes receive a release several days later than some of the data bank users. We therefore employ a "press service" which picks up specified releases from government offices and sends them by airmail. During any particular month we may use as many as a hundred different source documents.

Once the release is received, figures have to be entered into the different time sharing systems. A paper tape[6] is produced, which contains the name of the series, the date for which an addition or revision is made, and the new figures to be used. For each of the systems there exists a program which can take the information from that tape and make the necessary changes and additions to the affected time series.

Entries are checked from a print-out of the paper tape and then again from a listing produced by the update program. This listing is cut up and pasted on the time series listing, which is kept on file for each series. Periodically, new listings are obtained and checked against primary and secondary sources. The secondary sources we use are mainly the *Survey of Current Business*, the *Federal Reserve Bulletin*, *Business Conditions Digest*, and some others. Whenever one of these appears, print outs are again checked against these sources. (This is one way in which we find errors in Government publications!) Consistency checks and checks for the adequacy of seasonal adjustments are made frequently and, if necessary, the source agencies are informed of their outcome. In many instances inconsistencies can be resolved (or at least explained) and bad seasonal adjustments can be improved.

Major Revisions

At the present time, government agencies do not release machine readable revisions in advance of or concurrently with printed revisions. It seems that revisions are still computed by long hand and key punched afterwards. The Bureau of Labor Statistics is a possible exception, but their tapes are incompatible with our time sharing systems and very difficult to read. Until such time as procedures are modernized and communications standardized we have to enter all revisions by hand—punching paper tape, punch cards or magnetic tape casette. This is time consuming and error prone, and we hope it will soon be a thing of the past.

[6] We may switch from paper tape to a magnetic tape cartridge, which is easier to edit and to handle.

Software for Statistical Analysis

The analytical software available to the user of the data bank is, of course, different for the different systems. We added to the software offered by the various time-sharing companies only if we needed something for our own updating procedures or if generally useful programs (such as the X-11 seasonal adjustment program) could be incorporated.

We also have several FORTRAN programs which are specifically designed for the analysis of current business conditions, such as a Recession-Recovery analysis program, a program which determines cyclical turning points, and several diffusion index programs.[7] These programs, being written in minimum FORTRAN, are compatible with most time sharing systems and can be made available to users.

In the case of Rapidata, we were involved in the development of their language PLEA (Prototype Language for Economic Analysis) from its very beginning. We cooperated with Rapidata's programmers on specification, documentation and testing. They now have a very useful language and they are continuing to make improvements and refinements. PLEA deals with time series as vectors, which can be taken either from our data bank or a user's own files for a specified time period. It contains (1) a simple but powerful data manipulation section, including arithmetic operations, averaging, concatenation, deviations, extractions, mathematical functions, creations of dummy variables, (2) a facility to shift series in time, (3) the ability to perform a variety of statistical tasks, such as seasonal analysis, correlations, regressions, analysis of variance and factor analysis, and (4) several ways to output results, including plotting and the generation of new data files. These features are described in two manuals: "PLEA for Statistical Analysis," Rapidata, 1971 and "The PLEA*ECO Statistical Library," Rapidata, August 1970.

General Electric's MAP (Management Analysis and Projection) System runs on a GE 605 under the Desk-Side Time-Sharing System. It consists of a collection of programs that allow easy manipulation of data, analysis and forecasting. Input to those programs can be obtained from our data bank or from the user's own data. MAP includes various types of display of data—lists, tabulations, graphs, and a transformation package which can compute arithmetic operations, moving averages, leads and lags, various mathematical functions, growth triangles, curve fits, correlation matrixes, auto-correlations, and so forth. The mainstay of the system is GE's factor analysis forecasting system. All these programs can be used interactively if desired. In addition, several FORTRAN subroutines are available to interact with the data base. The system is described in "MAP, An Information Management Analysis and Projection System, Economic Data Base," General Electric Manual, November 1970.

Tape for Academic Institutions

Colleges and universities are interested in using the time series data bank for various purposes. Courses in econometrics, economic statistics, forecasting, business cycle analysis and model building can all use a data bank of time series,

[7] For a description of some of these programs see Gerhard Bry and Charlotte Boschan, *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*, NBER Technical Paper 20, 1971.

particularly if it is accessible by their computer. Some of the research may also utilize some of the time series included in the bank. For all these purposes it is usually not necessary that the data be very current and updated to the very last month. Therefore, we are providing some universities with magnetic tapes containing our data base. The format of these tapes is a compromise between the new "data transmission standards" set up by Ruggles and Sadowsky[8] and cost considerations in producing the tape from the existing data bank. (For format see Exhibit 1 in the Appendix.)

4. DOCUMENTATION

This is one of the most problematic areas. How much documentation should there be? How much of it should be on-line, how much hard copy? Is it necessary to document each figure in the bank? Can other source descriptions be used without actually being reproduced?

Hard Copy

The primary document describing the content of our data bank is a printed dictionary or directory consisting of four major parts: The main part lists all series included in the bank, with the six letter code name assigned to it, its starting date, periodicity, units, seasonal adjustment status, and a list of the source documents from which it was taken. Whenever possible, documentation is presented in tabular form, with a table showing all series from one source pertaining to a particular subject matter. Exhibit 2 is a sample page containing manufacturing and trade inventory and sales series published by the Department of Commerce. The series code names are shown in capital letters next to the descriptions and the starting years next to the names. Footnotes, including cross references, and source notes are shown at the bottom of the page. If the material is not easy to put in tabular form, it is just listed sequentially, as in Exhibit 3. It will be noted that the source documentation is somewhat unusual. The numbers on the extreme right hand side indicate the successive source documents from which the data were obtained. Thus, on line #09688, for "Number of Job Vacancies in Manufacturing" #45 refers to "U.S. Department of Labor, Bureau of Labor Statistics," *Employment and Earnings*, and #45B to the monthly news release from the same agency; "Job Vacancies, Hires, Quits and Layoffs in Manufacturing." (The key to these source code numbers is, of course, also included in the dictionary.) Thus, the source documentation in the dictionary refers to the printed releases which underlie the data shown in the bank. It does not, unfortunately, contain any economic or statistical description of the series, informing the user whether, for instance, it is based on a full count or a sample, when, where, and how data are collected, how reliable the data are and what their coverage is. Nor does it, at the present time, refer the user to the publication in which these characteristics of the time series are discussed and explained. This is a serious shortcoming of our documentation and we are in the process of correcting it. As a first step we will add code number and

[8] See Richard Ruggles and George Sadowsky, "Standards for Time Series Interchange." Mimeograph, NBER.

date of the document containing the description (or descriptions) to the source part of the index for each series or for entire sections of series. Since this is not a trivial task, it will take some time and effort. We have started to collect this information and to write up our own descriptions when none are published anywhere. Strangely enough very few users have ever asked us for this type of information. The reason is probably that we are dealing primarily with well known macro economic time series and most users know what they represent or at least where they are described.

The second part of the printed dictionary consists of a facsimile reproduction of the tables contained in *Economic Indicators* published by the Council of Economic Advisers. Instead of data, the associated code names are shown. The third and fourth parts consist of similar reproductions of the "Series Finding Guide" and "Titles and Sources of Series" of *Business Conditions Digest* and of the National Income and Product Accounts as published in *The Survey of Current Business.* In each case our code names are shown next to the descriptions of the series. Sample pages are shown in Exhibits 3 to 6. The facsimile reproductions included in our directory are intended to facilitate the use of the data bank for people who are accustomed to working with these publications. A table of contents, by subject matter, and an alphabetical index are also contained in the dictionary.

The make-up of the alphabetical index, incidentally, is somewhat problematic. An alphabetic listing of the code names is obviously useful only for people who know the code names, and if they know the code names, they don't need an index. For finding code names what needs to be alphabetized is the series title. But a series title usually contains many "key words" and if you list each series under each of the key words you get a very voluminous index. To take a relatively harmless example, the series called "Manufacturers' unfilled orders, durable goods industries" would have to be listed under manufactures, under orders, under unfilled orders, and under durable goods. Still worse would be the listing of a series compounded from heterogeneous components, such as "Manufacturers' machinery and equipment sales and business construction expenditures (industrial and commercial construction put in place)." In addition, the selection of key words is hard to program if we want to avoid innumerable listings of series under "goods," under "industries," under "sales" and under "expenditures." One could, of course, let the computer do the alphabetizing and then take the nonsense items out, but in any case the resulting index would be quite voluminous. (N.B.: One well known data bank produced such a massive computerized compilation and its salesmen point with pride to the comparative richness of its collection and the comprehensiveness of its cross-classification system by reference to the weight of the printed index.) We are presently omitting an alphabetical index with multiple listings for each individual series. Our index is alphabetized by subject matter (orders, sales, construction, etc.) and refers the reader to a page in the printed dictionary.

On-line Documentation

Although hard-copy documentation is important, a certain amount of information must be available on line. One reason is that users, accustomed to the

convenience of man-computer interaction afforded by time sharing, often do not bring their manual to the terminal when they are using the data bank. The second, perhaps more important, reason is that in a data bank such as ours, changes in documentation are needed quite frequently and printed copy would have to be kept up by some looseleaf arrangement or similar scheme. In cross-sectional or other single-time-period data banks, where the information usually does not change once it is in the bank, the documentation obviously need not change either. Our data bank, however, is a growing, constantly changing collection. Not only is the information contained in individual time series frequently revised and updated, but new series are added, old ones discontinued, samples, base periods, sources, and source descriptions are changed and so forth. The user must be informed of these changes as soon as they occur.

    We are taking care of these needs in two ways. First, we keep an entire listing of all included time series by subject matter on-line in the data bank. Instead of looking up series titles in a printed book the user can let the computer list parts in which he is interested. Since most time sharing systems have editor programs which permit the listing of all lines containing specified combination of letters, it is usually quite easy for the user to find the series in which he is interested. (See pages 6–8 of the directory for an example of how this is done.) These on line files, called GINDEX for GNP series and INDEX for all others, are kept up-to-date: new series are added, old ones deleted, source notes revised and brought up-to-date, and so forth.

    In addition to this kind of inventory of all series in the bank, we used to keep a record of "transactions." A file called "change," written by the update program whenever a series was revised or updated, contained changes made, by date and series code name. Ideally, a user could let the edit program list all the changes made to a particular series by just listing each line on which the code name of the series appeared. We thought that this would be a very useful thing to have. However, it turned out that nobody ever seemed to use it. Specifically, when a bug in the program resulted in the deletion of most of the file, no one ever brought this to our attention. Since there were also some technical difficulties caused by the file's becoming too large too fast and thus quite expensive, we decided to discontinue writing it. No one has ever missed it, or if they did, they spared our feelings. All the same, we still think that this type of information—or at least the date of the last change in a series—ought to be available to the user. If we find an inexpensive way to provide it we will certainly do so.

    In the meantime, we have several other, perhaps minor, pieces of useful information on line. A file called "INFORM" contains such information as dropped series, revisions of whole batches of series, prospective revisions, etc.

Special Services

    Users of our data bank vary widely with regard to experience, sophistication and expertise in economics, statistics and computer use. They need varying amounts of assistance in finding series, using series, writing and using programs, interpreting results and many other things. Some of the more technical computer and programming assistance is usually provided by the time sharing company,

but many questions arise which we have to answer. Typically, somebody misreads a code name (we now try to avoid l's and 1's, as well as O's and zeros), confuses monthly and quarterly series or uses wrong starting dates. Occasionally users want to know when certain series will be updated and are very much upset if the dates cannot be predicted with accuracy since they vary somewhat from month to month. Other users enquire about the purpose and justification of particular revisions and changes in seasonal factors. Most of these questions are relatively easy to handle. But every once in a while somebody asks a really tricky question. Why do the turning points listed in the BCD differ from the ones given by us? What exactly is the difference between the insured unemployment rate and the total unemployment rate? How can he use the Durbin–Watson Statistics given in the Rapidata regression program? What variables should he use to forecast the sales of his company? Why do we not carry them in the data bank? We try to answer most reasonable questions to the best of our ability, particularly if they refer to data in the bank. Often questions are asked only because our documentation is not good enough to obviate them. This, we hope, will change in the near future. However, some questions will always need to be asked and, in principle, we welcome this fact, since it brings us in contact with the users of our data bank and provides a certain amount of feedback on how the data and the data bank are used.

5. THE ECONOMICS OF THE NBER DATA BANK

General Considerations

The National Bureau is a nonprofit publicly supported research organization. We do not intend to make any profit on the data bank operations, nor do we want to incur a loss. The entire venture is planned as a cooperative enterprise, with each member paying an equal share of the cost. However, breaking exactly even is an unusual managerial goal which is very difficult to attain. Originally, our price of $750.00 per user per year was set with the intention of breaking even, assuming a constant number of users. In principle we intended to add more series to the bank as more users joined the project and thus continue to break even. In practice, series were added much faster than users, i.e., the costs increased faster than the revenues. Furthermore, the original costs were somewhat underestimated, and thus we are losing money. The loss would be somewhat smaller if the imputed benefits derived by internal use were explicitly considered.

What can be done about the loss? There is the possibility of raising the unit price (before the price-wage freeze). But we do not know by how much since we are trying to reduce machine cost by shifting to the night shift, and we do not know yet how much we can shift. Furthermore, development expenditures on new time series and the directory will be reduced. On the other hand, we will have to spend more time on source documentation. Thus, it takes a lot of juggling and close watching to stay in the proximity of the break-even point.

User Characteristics

A large number of our users are members of the original "Project Economics." One reason for this may be that we have not advertised the availability of the bank, and many potential users do not know that it exists. There are now about 50

users who participate in the cost of maintaining the data bank. They consist of the economics departments of 15 banks, 13 large industrial corporations, 2 insurance companies, 3 public utilities, 3 economic consultants, 3 stock brokers and investment advisors, and several others.

We originally intended to make data bank tapes available to academic users at incremental costs, but to charge the full price (average estimated cost) if an academic user wanted fast updating and on-line services. However, if an academic user participates in a time-sharing system which contains our bank and happens to get fast updating and on-line services, our incremental costs are of course zero. We are therefore not charging anything to universities who want to use our data through one of our affiliated time sharing systems. Their only costs are computer costs, and whatever other costs they incur in their time-sharing system. However, some universities are interested in incorporating our series in their own system—time sharing or otherwise. For these users, we periodically dump the data base on tapes and they share in the cost for this service. At present 7 universities have a version of our bank.

Another, not unimportant, group of users is the National Bureau research staff. People whose work is in the field of business cycle analysis and current business conditions analysis are obviously heavy users, but model builders and others interested in time series make frequent use of our machine readable data.

Personnel

With the present configuration of about 2000 series in two time sharing systems, the permanent personnel consists of one senior analyst and three research assistants. However, the availability of a wide variety of specialized talent within the walls of the National Bureau is of great help in the running of the data bank. We can get the help and advice of specialists on many different subject matters with regard to the selection and definition of series to be included, particularly if they are derived and not directly taken from a published source; we can draw on the experience of business cycle analysts with regard to seasonal adjustments and other technical matters; we can get a programmer to do special purpose modifications of the general update program or to write a tape for use on another system; we can have the help of an experienced data librarian to trace discrepancies and write source descriptions; and, in emergencies, we can get the help of a typist to copy data from releases to paper tape. We could not, of course, support any of these talents on a full time basis.

Expenditures

Apart from the advance releases which we get by airmail from Washington, the data bank librarians are, of course, making use of the National Bureau's extensive library of current and historical publications. The availability of the library is important for checking historical data, for determining specific concepts and definitions of series, and for other tasks.

The time sharing systems in which our data bank resides are, like most time sharing systems, accessible through a variety of terminals. Since the Bureau research staff also uses a variety of terminals for access to different computers, it is possible for the data bank operation to use different terminals for different

purposes. (Teletype for paper tape, 2741 for wide carriage, Novar for magnetic tape cassette and high speed transmission). At the present time we use about the equivalent of one-and-a-half terminals full time.

Telephone cost, particularly for long distance communication, is a potentially very expensive item. Fortunately Rapidata's computer is in New York, GE is accessible by a local telephone call from anywhere in the country, and ISC, the new system to be added, has a New York FX line. In addition, the Bureau has a Watts line to Massachusetts which we can use for data bank operations.

6. CONCLUDING REMARKS

I have tried to show you what kinds of problems we faced in setting up and running our data bank, and how we have attempted to solve them. We hope that some of our solutions can be regarded as sufficiently adequate to help those who face similar problems.

Our efforts are, of course, all conducted within the somewhat narrow framework of single private time-series-and-forecasting oriented retrieval systems. It has been suggested that an all-embracing central system, which includes all published macro-economic and related series and makes them widely available, would be greatly superior to the decentralized and often duplicative efforts of individual data banks. Such a central data bank would in all likelihood be government operated. Some countries do indeed have such government-run data banks,[9] or they plan to install them. We are confident that even such central data banks may profit from the experiences of past pioneering efforts—be they governmental or private.[10]

In the United States we are very far from the implementation of a central information system—farther than some smaller countries in which computers are almost exclusively in governmental hands. In the U.S. it is likely that private data banks such as ours will continue to exist and to fill an important function in the provision of data and programs, as well as in the improvement of information technology. On the other hand, several government departments—such as the Bureau of Labor Statistics and the Department of Commerce—installed data banks for their own data and analyses. These efforts vary widely in quality, depending on hardware and know-how available at the time of their installation. In time, these banks might well be consolidated into one system. However, even when a centralized government-run time-shared information system does become available, private banks may continue to play an important role as monitors, innovators, consultants, and providers of special services. I think it is unlikely, though, that the Bureau's bank will be among them. Typically, the Bureau has tended to turn over operational responsibilities for its systems to the government and to restrict its role to methodological and interpretative functions. This was true for the National Income Accounts, cyclical indicators, consumer credit statistics, and many other statistical systems. It is also likely to be true for data bank operations.

[9] The most advanced bank of this kind is probably the Canadian one, run by Statistics Canada, which, incidentally, is not available for time-sharing except through a commercial system.

[10] One of the earliest governmental efforts was made in 1962 by Rudolph C. Mendelsohn of the U.S. Department of Labor, Bureau of Labor Statistics.

APPENDIX

EXHIBIT 1

MAGNETIC TAPE FOR ACADEMIC USERS

Description of Data Base Distribution Tape

I. Physical Attributes (can be changed on request)
   1. 9-track
   2. EBCDIC
   3. No label
   4. 800 B.P.I.
   5. 132 characters per record
   6. 50 records per block (blocksize = 6600)

II. Format of data series

   Each series is preceded by two records with descriptive information.

   The first record contains:
   1. Line number from an index file (ignore)
   2. Series code (up to 6 alpha-numeric characters)
   3. Series name

   The second record contains:
   1. Columns 1–4   Number of observations
   2. Columns 6–9   First year
   3. Column 11   1—quarterly
                    2—monthly
                    3—annual
   4. Columns 13–18 Series name left adjusted.

   Following these two records come series data in one of the following formats:

   1. Quarterly (I2, 1X, 4F11.3), where columns 1–2 contain last two digits of year (e.g. 47 for 1947)
   2. Monthly (I3, 6F11.3), where columns 1–2 contain last two digits of year and column 3 is 1 for first six months and 2 for last six months of the year (e.g. 672 for July–December of 1967)
   3. Annual (I4, F11.3), where columns 1–4 contain last two digits of year plus two zeros (e.g. 6900 for 1969)

   III. End of File

      The end of file is last record written on tape

EXHIBIT 2
SAMPLE PAGE OF DIRECTORY, TABLES

MANUFACTURING AND TRADE INVENTORIES AND SALES

(millions of $, seasonally adjusted)

	Sales		Inventories	
Manufacturing and Trade	MRWT	(48)*	IVMT	(48)*
Manufacturing	MFGS	(47)	IVM	(47)
Durable	MDS	(47)	IVMD	(47)
Nondurable	MNS	(47)	IVMN	(47)
Retail Stores	RT	(46)	IVR	(47)
Durable	RD	(46)	IVRD	(47)
Nondurable	RN	(46)	IVRN	(47)
Merchant Wholesalers	WT	(48)	IVW	(48)
Durable	WTD	(48)	IVWD	(48)
Nondurable	WTN	(48)	IVWN	(48)

See also:
  Change in Book Value, Manufacturing and Trade Inventories IVCMT (48).
  Change in Unfilled Orders, Mfg. Durable Goods Ind. IVCUD (47).
  Ratio, Inventories:sales, Total Manufacturing and Trade IVSRMT (48).
  Ratio, Unfilled Orders:Shipments, Mfg. Durable Goods Ind. MDUXS (53).

    * Billions of $.

    Source: U.S. Department of Commerce, Bureau of Economic Analysis, "Manufacturing and Trade Inventories and Sales." (M).

MANUFACTURERS' INVENTORIES AND SALES: ACTUAL AND EXPECTED

(Billions of dollars, seasonally adjusted, back data, starting in 1947, is actual)

	Sales Total for Quarter	Inventories* End of Quarter
All manufacturers	MFGSAN	IVMANT
Durables	MDSANT	IVMDAN
Nondurables	MNSANT	IVMNAN

    * Inventory expectations corrected for systematic biases.

    Source: U.S. Department of Commerce. Bureau of Economic Analysis. Manufacturers' Inventory and Sales Expectations (0).

EXHIBIT 3

SAMPLE PAGE OF DIRECTORY, SEQUENTIAL LISTING

09528-
09529-HOURS & EARNINGS:NBER & OTHER SPECIAL SERIES
09530-

09531	LHCH	AVG. WEEKLY HOURS WORKED, CLF:HOLIDAY ADJ.	HOURSA	47	#45(A-22, A-20):70 NBER 8, 354
09532	LHOURS	MANHOURS PER WEEK, EMPLOYED CLF	MWHOURSA	47	#45:70 NBER 8, 355
09533	LHLOST	POTENTIAL MANHOURS LOST PER WEEK, CLF	MWHOURSA	48	#45:70 NBER 8, 356
09534	LHPOT	POTENTIAL MANHOURS PER WEEK, CLF	MWHOURSA	48	#45:70 NBER 8, 357
09535	LPCNAG	HOURLY COMPENSATION, NONAG EMPLOYMENT	$SA	47	#70 NBER 8, 350
09536	LPCMFG	HOURLY COMPENSATION, MFG. EMPLOYMENT	$SA	47	#70 NBER 8, 332
09537	LEHGF6	FED. EXEC BRANCH, ALL EMPL. AHE	R7	65	#45 (C-3)
09538	LEMXO	MFG. AHE "STRAIGHT-TIME" PROD WORKERS	$	46	#45 (C-4)
09540	LEMXS6	MFG. AHE EXCL. OVERTIME & INTERINDUSTRY SHIFTS	R7	47	#65, BLS
09546-	LHSTOP	MAN-DAYS IDLE DUE TO LABOR-MANAGEMENT DISPUTES	THDAY	47	#93 (NBER 8, 257)

09550-REAL AND SPENDABLE EARNINGS (PRODUCTION WORKERS)

09552	LEW7	REAL AWE PRIVATE NONFARM, PROD WORKERS	$67SA	60	#45C or 50
09553	LPC3	REAL SPENDABLE AWE, PROD WORKER, 3 DEPENDENTS	$67SA	64	#45C or 50
09554	LPC2	REAL AHE, MFG. PROD WORKERS	$67SA	47	#45A and 50

09648-
09649-UNEMPLOYMENT INDICATORS:MANPOWER ADM, INSURANCE & OTHERS
09650-

09651	LUINC	AVG. WEEKLY INITIAL CLAIMS, UNEMPL. INSUR. (STATE PROGRAMS)	THSA	48	#52:70
09652	LNAGP	NONAGRICULTURAL PLACEMENTS, ALL INDUSTRIES	THSA	45-70	#37, 51:65
09653	LJOB	NONAG JOB OPENINGS, UNFILLED	THSA	47-70	#51:65
09654	LJOBXU	RATIO, UNFILLED NONAG JOB OPENINGS:NUMBER UNEMPLOYED	RSA	47-70	#51, 45:65:70
09655	LHEL	NICB INDEX OF HELP-WANTED ADVERTISING IN NEWSPAPERS	R7SA	48	#21:65
09657-	LAUREM	COVERED EMPLOYMENT, UNEMPL. INSURANCE ALL PROGRAMS	TH	46A	#51A(P. 93) OR 72
09659-	LURP	INSURED UNEMPLOYMENT, ALL PROGRAMS (WKLY. AVG.)	WAV	55	#37(P. 85) OR 72
09661-	LURPB	BENEFITS PAID, UNEMPL. INSURANCE, ALL PROGRAMS	M	58	#65, 51 OR 72
09663-	LURPS	INSURED UNEMPLOYMENT, STATE PROGRAMS	WAV:TH	56	#65,45 OR 72
09665-	LUNIC6	INITIAL CLAIMS, UNEMPLOYMENT INSURANCE, STATE PROGRAMS	WAV:TH	56	#65, 45 OR 72
09667-	LUNEX	EXHAUSTIONS, UNEMPLOYMENT INSURANCE, STATE PROGRAMS	WAV:TH	56	#65, 45 OR 72
09669-	LURSP6	INSURED UNEMPLOYMENT AS % OF COVERED EMPLOYMENT	%	49	#65, 45 OR -72
09671-	LURSPB	BENEFITS PAID, UNEMPL. INSURANCE, STATE PROGRAMS, EXT. PROV.	$	46	#51A(P. 135), 35 OR 72
09672-	LURSPW	AVG. WEEKLY BENEFIT CHECK FOR TOTAL UNEMPL. STATE PROGRAMS	$	46	#51A(P. 136),35 OR 72

09673-
09674-LABOR TURNOVER RATES

09675-					
09676 LPACCM	ACCESSION RATE, MANUFACTURING	%SA	47	#37,45B
09080 LPOFFM	LAYOFF RATE, MANUFACTURING	%SA	47	#34,45B
09685-					
09686-JOB VACANCIES IN MANUFACTURING					
09687-					
09688 LVM	NUMBER OF JOB VACANCIES IN MANUFACTURING	TH	4/69	#45,45B
09690-LVMR	JOB VACANCY RATE, MANUFACTURING	%	4/69	#45,45B
09692-LVLM	NUMBER OF LONG-TERM JOB VACANCIES, MANUFACTURING	TH	4/69	#45,45B
09694-LVLMR	LONG-TERM JOB VACANCY RATE, MANUFACTURING	%	4/69	#45,45B

<div align="center">

EXHIBIT 4

SAMPLE PAGE OF DIRECTORY, ECONOMIC INDICATORS

ECONOMIC INDICATORS

</div>

New Housing Starts and Applications for Financing

	Housing starts (Thousands of units)		
Total private and public (including farm)	< H6SF	59	5406
Total private (including farm)	< H6SP	59	5405
	Housing starts (Thousands of units, s.a.a.r.)		
Private			
Total private (including farm)	HSF	59	5403
One unit	HSPl	64	5404
Two or more units	HSF–HSPl	64	Derive
Government home programs (nonfarm)			
FHA	<HSFHA	46	5431
VA	<HSVA	47	5432
New private housing units authorized[1]	HSBP × 0.0879569	46	Derive
Proposed home construction			
Applications for FHA commitments[2]	<HFFHA	45	5433
Requests for VA appraisals[2]	<HFVA	51	5434

[1] HSBP is a spliced index: the authorizations in 10,000 permit-issuing places prior to 1963 and in 12,000 permit-issuing places from 1963–66 have been raised to the level of 13,000 permit-issuing places. The factor converts the index to units.

[2] Units represented by mortgage applications for new home construction.

Business Sales and Inventories
(millions of dollars, seasonally adjusted)

Total business (includes manufacturing)			
Sales[1]	MRWT	48	3251
Inventories[1]	IVMT	48	3261
Wholesale			
Sales	WT	48	3258
Inventories	IVW	48	3270
Retail Sales	RT	46	3255
Durable goods stores	RD	46	3256
Nondurable goods stores	RN	46	3257
Retail Inventories	IVR	47	3267
Durable goods stores	IVRD	47	3268
Nondurable goods stores	IVRN	47	3269

[1] Billions of dollars.
Series marked < are not on line but available if enough users request them.

EXHIBIT 5

SAMPLE PAGE OF DIRECTORY, BCD

SECTION III

BCD BUSINESS CONDITIONS DIGEST

NBER TIME SERIES DATA BANK

In the following pages, the series published in the Census Bureau's publication Business Conditions Digest (BCD), NBER's code name, and the starting year are shown in the order of their appearance in the BCD Index. Source references such as "1,1" or "1,14" indicate a table number of the National Income and Product Accounts as published in the Survey of Current Business, and indexed on-line on "GINDEX". Four digit references such as "1521" or "9907" refer to a line number in the on-line NBER INDEX and SOURCE.

The NBER data bank carries time series in units as published by the primary source. It is always advisable to consult the SOURCE listing for specific characteristics of a series.

All series except those marked * are updated in the data bank before they become available in the BCD.

BCD Series

BCD BUSINESS CONDITIONS DIGEST

NBER TIME SERIES DATA BANK

	1948	5309
	1953	99011
	1948*	99012
	1947	99013
	1948	1064
	1947	99015
	1946	1,10
	1947	99017
	1947	760
	1947-1958	4364
	1953	4365
	1956	9369
	1946	1,14
	1948	5200
	1953	4347
	1947	3278
	1950	99026
	1947	5401
	1947	5430
	1948	3262
	1948	99032
	1955	1060
	1947	99037
	1948	1067

EXHIBIT 6
SAMPLE PAGE OF DIRECTORY, NATIONAL INCOME ACCOUNTS

SURVEY OF CURRENT BUSINESS

NATIONAL INCOME AND PRODUCT TABLES

Table 9.—Gross Corporate Product [1] (1.14)

Seasonally adjusted at annual rates	Billions of dollars (1946)
Gross corporate product.........	GX
Capital consumption allowances.......	GCCCA
Indirect business taxes plus transfer payments less subsidies.............	GCTTS
Income originating in corporate business...................................	GXY
Compensation of employees.........	GCCORP
Wages and salaries...............	GCW
Supplements....................	GCSUPP
Net interest....................	GXINT
Corporate profits and inventory valuation adjustment...........	GXIVA
Profits before tax................	GXPRT
Profits tax liability..............	GPTAX
Profits after tax................	GXPAT
Dividends....................	GXDIV
Undistributed profits..........	GXUP
Inventory valuation adjustment..	GIVA
Cash flow, gross of dividends.........	GXFLOW
Cash flow, net of dividends...........	GXFLUP
Gross product originating in financial institutions.........	GXF
Gross product originating in nonfinancial corporations.........	GXJ
Capital consumption allowances......	GJCCA
Indirect business taxes plus transfer payments less subsidies.............	GJTTS
Income originating in nonfinancial corporations.....................	GJY
Compensation of employees.........	GJCORP
Wages and salaries...............	GJW
Supplements....................	GJSUPP
Net interest.....................	GJINT
Corporate profits and inventory valuation adjustment...........	GJIVA
Profits before tax................	GJPRT
Profits tax liability..............	GJTAX
Profits after tax................	GJPAT
Dividends....................	GJDIV
Undistributed profits..........	GJUP
Inventory valuation adjustment....	GIVA
Cash flow, gross of dividends..........	GJFLOW
Cash flow, net of dividends...........	GJFLUP

	Billions of 1958 dollars (1948)
Gross product originating in nonfinancial corporations......	GXJ58

	Dollars (1948)
Current dollar cost per unit of 1958 dollar gross product originating in nonfinancial corporations [2]...................	SDXJ
Capital consumption allowances......	SDJCCA
Indirect business taxes plus transfer payments less subsidies.............	SDJTTS
Compensation of employees............	SDJCOM
Net interest.....................	SDJINT
Corporate profits and inventory valuation adjustment...........	SDJIVA
Profits tax liability..............	SDJTAX
Profits after tax plus inventory valuation adjustment..	SDJPIV

1. Excludes gross product originating in the rest of the world.
2. This is equal to the deflator for gross product of nonfinancial corporations, with the decimal point shifted two places to the left.
3. Personal saving as a percentage of disposable personal income.

Table 10.—Personal Income and Its Disposition (2.1)

Seasonally adjusted at annual rates	Billions of dollars (1946)
Personal income.........................	GPY
Wage and salary disbursements......	GW
Commodity-producing industries..	GWCP
Manufacturing..................	GWCPM
Distributive industries............	GWD
Service industries................	GWS
Government.....................	GWG
Other labor income..................	GPOL
Proprietors' income...................	GPHJ
Business and professional...........	GPHOB
Farm..............................	GPHOF
Rental income of persons.............	GPRENT
Dividends............................	GDIV
Personal interest income.............	GPINT
Transfer payments...................	GPT
Old-age, survivors, disability, and health insurance benefits.........	GPTOAS
State unemployment insurance benefits..........................	GPTINS
Veterans benefits..................	GPTVET
Other..........................	GPTXO
Less: Personal contributions for social insurance...................	GPSI4
Less: Personal tax and nontax payments...............................	GPTX
Equals: Disposable personal income....	GYD
Less: Personal outlays.................	GOUT
Personal consumption expenditures..	GC
Interest paid by consumers.........	GCINT
Personal transfer payments to foreigners..........................	GPTF
Equals: Personal saving................	GPSAV
Addenda: Disposable personal income:	
Total, billions of 1958 dollars........	GYD58
Per capita, current dollars............	GYOPC
Per capita, 1958 dollars..............	GYDPC58
Personal saving rate, [3] percent.... ..	GXSAV

Table 11.—Personal Consumption Expenditures by Major Type (2.3)

Personal consumption expenditures............................	GC
Durable goods............................	GCD
Automobiles and parts..............	GCDA
Furniture and household equipment..	GCDF
Other............................	GCDO
Nondurable goods......................	GCN
Food and beverages..................	GCNFD
Clothing and shoes..................	GCNC
Gasoline and oil.....................	GCNG
Other..............................	GCNO
Services...............................	GCS
Housing..........................	GCSH
Household operation................	GCSHO
Transportation....................	GCST
Other..............................	GCSO

Table 12.—Foreign Transactions in the National Income and Product Accounts (4.1)

Receipts from foreigners...	GEXF
Exports of goods and servi es	GEX
Capital grants received by the United States...................................	GEXCG
Payments to foreigners...........	GEXF
Imports of goods and services.........	GIM
Transfers to foreigners................	CTF
Personal...........................	GPTF
Government.....................	GGPTF
Net foreign investment..............	CFINET