versational mode, the user sees his answers almost immediately and can decide on the spot what further equations he may want to run on the same or other data.

3. The program language used is APL, which is particularly suited to the conversational mode and which offers a highly condensed notation for mathematical operations. A typical sixty-line page of FORTRAN statements can usually be expressed in two or three lines of APL statements.

Parts of the above innovations have been successfully tested in a prototype APL regression program written for the IBM-1130.

Sidney Jacobs

A Study of the Properties of the Minimum-Sum-of-Absolute-Errors Estimator

In this study, an attempt is being made to determine the sampling distribution of the minimum-sum-of-absolute-errors (MSAE) estimator of the parameters of a linear regression. The estimator is known to out-perform the least-squares estimator in applications where disturbances are characterized by very dense extreme tails, as, for example, where the disturbances follow the symmetric stable distributions with characteristic parameter α very much less than two. Mandelbrot has argued persuasively that such fat-tailed distributions are the appropriate model for many economic processes, which suggests that MSAE may be a good estimator for economists to employ.

Developing a sampling theory for MSAE is made difficult by the fact that there exists no analytic expression for the MSAE estimator. Instead, it is calculated via a linear-programming algorithm. However, by utilizing some results on the distributions of order statistics, it should be possible to characterize the asymptotic distribution of the MSAE estimator. Monte Carlo experiments can then be used to supplement the asymptotic results and enable us to assess how seriously the small-sample distributions seem to depart from the asymptotic distributions.

Reports on some preliminary work on this topic are contained in last year's *Annual Report,* as well as in a forthcoming volume of *Econometrica.* The study is being conducted jointly with Robert Blattberg of the University of Chicago.

Thomas J. Sargent

10. ELECTRONIC COMPUTER SERVICES IN SUPPORT OF ECONOMIC RESEARCH

Introduction

The Bureau's electronic data processing operations encompass a large variety of activities, such as programming, consulting, and other services connected with individual research projects; improvement of data storage and retrieval; and development of programmed approaches to statistical problems. While major services are provided in response to internal demands, we are increasingly attempting to make our resources available to outsiders.

In the supporting operations there are three developments worth pointing out: (1) with the increasing size and complexity of data sets received from government agencies and other sources, data retrieval has become more important and more difficult; (2) the fact that programming is taught in schools and universities has led to an increase in the importance of our consulting functions as compared with our programming activities; (3) easier access to various computer systems, via remote terminals and time-sharing, has increased the necessity for system selection, job channeling, and similar operations. These changes are described below in the report on the activities of the E.D.P. unit.

The availability of ever-increasing quantities of statistical data requires the development of efficient methods for data retrieval and manipulation, as well as of the documentation describing the meaning, the coverage, and other characteristics of the data. New methods of storage, retrieval, presentation, and documentation are described by Richard and Nancy Ruggles in their report on project RIPP. Our effort to facilitate the accessibility and manipulation of the Bureau's time series collection is described in my report on the National Bureau's data bank. A third approach to developing more efficient methods is that pursued in the newly organized Universities-National Bureau Conference on the Application of the Computer to Economic Research, reported on in Part III of this report.

Several new programs and improvements of old programs were developed during the report period, particularly in reference to business cycle analysis. A description of our efforts in this area is found below.

Another effort is aimed at the improvement of nonlinear regression programs. A major shortcoming of existing programs is that the accommodation of nonlinear relationships involves large expenditure of effort and machine time. Our attempts to remedy this condition are described by Sidney Jacobs in Section 9 of this report.

Charlotte Boschan

Operations of the Data Processing Unit

The use of electronic computers in the Bureau's research operations continues to increase. Our small IBM-1130 computer is being used for at least two shifts. In addition, we are using computing facilities of Yale, Columbia, various other universities, and some commercial service bureaus. The most important of these operations, those at Yale, are described below by Sanford Berg. Access to some of these facilities is via remote terminal. In view of the increasing importance of time-sharing arrangements, particularly in connection with data bank operations, we are always experimenting with new systems in order to establish their relative efficiency and cost. These experiments are helping us to select the systems most appropriate for our operations.

A significant change in the operations of the E.D.P. unit has been brought about by the trend towards large data sets. Much of the work done in the Bureau's research projects requires retrieval of data from magnetic tapes and involves various storing and merging operations. These tapes are originally generated by a large number of different systems with different characteristics; the lack of standardization, or even compatibility, makes merging information from two different sources a somewhat complicated undertaking, particularly if neither source is compatible with our own computer system. These conditions emphasize the need for standardization and documentation, planned for the future by Richard and Nancy Ruggles. At present, however, a large part of our time is spent in writing tape utility programs for various machines and in helping researchers to handle heterogeneous data sets.

Because many research projects now have their own programmers, whose skills range from simple coding to sophisticated programming, our central programming staff, instead of writing ad hoc programs for one-time use, spends considerable time in teaching, advising, and assisting. Among our supporting operations is electronic charting on a Calcomp Plotter attached to the IBM-1130. This operation is increasingly replacing hand charting, both for analysis and for publications.

In principle, all members of the data processing unit are taking part in all its operations. To the extent that there is specialization, Lora Weinstein and Susan Crayne do counseling; mathematical and complicated statistical matters are handled by Sidney Jacobs; Antonette Delak takes care of charting operations including some programming; and Irene Abramson helps with operations, programming, and maintenance of the 1130. Assistance and advice

concerning the operation of the IBM-1130 is given by Martha Jones, who also handles the unit-record equipment, and Dora Thompson takes care of card and paper-tape punching.

Charlotte Boschan

NBER Computer Operations at New Haven

The NBER computer operations at New Haven consist of three parts: (1) development of administrative information processing, ADMIN (2) development of research information processing, RIPP, and (3) support of current research activities.

In the ADMIN project, we are exploring ways to computerize administrative tasks and information dissemination processes. At the present stage of development, computer programs exist to create labels for mass mailings to specified classes of NBER subscribers and to facilitate budgeting operations. A description of the RIPP project is contained in the report by Richard and Nancy Ruggles.

The use of Yale computer facilities by NBER research projects has increased. Daily United Parcel delivery service is connecting New York and New Haven operations. While we have not the staff to handle substantial debugging of programs, all output is examined so that minor changes might be made if it appears that this will result in a successful run. A series of seminars has been initiated in order to introduce NBER staff members in New York to Yale's IBM-360/50. Warren Sanderson presented the first seminar on the general principles of this system. Later seminars will examine programming techniques and research strategy within the framework of an interactive computer system.

Sanford Berg

Progress Report on Project RIPP

The Research Information Processing Project (RIPP) has been set up by the National Bureau to develop computer techniques for large scale storing and retrieving of information useful in economic research. The function of the project is to provide for the long-run development of computer capabilities rather than to assist in the day-to-day problems encountered by specific research programs.

A major portion of the project is focused on the development and utilization of machine readable documentation systems. These documentation systems are concerned with handling three kinds of information. First, the time series which the National Bureau has on hand are being documented in such a manner that they can be processed and retrieved by standard programs and used in conjunction with existing statistical packages designed for analyzing time series. Second, methods of generating worksheets and tables in machine readable form are being developed so that the research analyst can call on such information and produce tabulations suitable for further research or publication. Finally, work is under way on the development of machine readable documentation describing the organization and content of magnetic tape and disk files so that they can be accessed and operated upon by standard programs.

The project is considered to be part of the development work underlying the National Bureau's effort to create a machine readable data library for research purposes. Although the problems involved are basic to any computer configuration and to the statistical processing of any body of research data, the work is geared to random access third generation computers, in particular the IBM-360/50 at Yale University.

Richard Ruggles
Nancy D. Ruggles

The National Bureau Data Bank

The National Bureau makes a part of its collection of time series available in the form of a machine readable data bank. Users may be out-

siders as well as members of the Bureau's staff. At present this data bank contains about 1,000 monthly and quarterly time series. We plan to increase the collection to about 1,500 series by the end of 1970. Series are updated and revised as soon as the data become available. A list of the available series can be provided.

Typically, the outside user belongs to a group centered around a computer system; access to the data base is obtained via remote terminals or other means. This arrangement is advantageous for the outside user as well as for the Bureau. For the user it permits the sharing of transmission cost, storage cost, software development, and so forth; for the Bureau it reduces the effort involved in the mechanical transmission of updating. At present, only one user organization (Project Economics, on the Rapidata System) uses the NBER data bank; the Information Service Department of General Electric has about completed arrangements for a similar relationship on their own computer system, and other user groups have approached the Bureau. Some outsiders use the data bank as individual subscribers, e.g., the Department of Economic Research of IBM.

The National Bureau is compensated for its services through a fee paid by the individual user; each user must agree to a set of specified terms of cooperation. Academic users pay a nominal fee for once-a-month tape service. These fees do not include any costs incurred by users in connection with the services of their computer system.

Day-to-day operations of the data bank are supervised by Peggy Cahn. Hanna Stern is working on source material and data documentation. Others involved in the project are Constance Lim, Young Lee and Wan-Lee Hsu (updating and revisions), Dorothy O'Brien (seasonal adjustments, checking for comparability), and Lora Weinstein (computer operations).

Charlotte Boschan

Programmed Determination of Cyclical Turning Point and Timing Measures

Determination of turning points. The program for the determination of specific cyclical turns in individual time series, developed by Gerhard Bry and myself, has been used in various Bureau projects and stands up rather well.[1] There are, however, a few shortcomings which should be known to prospective users, although they show up infrequently.

The program may designate shallow fluctuations as cyclical, even in series that typically exhibit steep cycles. On occasion, this might be undesirable. However, we hesitate to modify the program, since shallow fluctuations might well have become a characteristic of present-day economic developments.

Another problem arises in connection with the use of the program for the analysis of current business conditions; the program reflects considerable caution in recognizing turns close to the end of a series. In order to avoid recognizing temporary reversals as cyclical, the program requires at least six months of reversed direction. This conflicts with the need for early recognition of cyclical changes in individual series and in the economy as a whole. Rather than change criteria, we shall provide options to the user for varying the duration of reversals required for recognizing turns.

Finally, in excluding turns associated with "short" phases (less than five months), the alternative peak selected by the program may be lower than the eliminated one, or the alternative trough may be higher. This result is justified if the eliminated turn is randomly extreme but not if it reflects a cyclical reversal. Further experience with this problem may lead to program modification.

Timing relationships. Timing comparisons describe the relationship among cyclical turning points in different economic time series.

[1] The program is described in the forthcoming book by Gerhard Bry and Charlotte Boschan on "Cyclical Analysis of Time Series: Selected Procedures and Computer Programs."

These comparisons are carried out by determining leads, lags, and coincidences of cyclical turns in specific time series relative to reference turning points (which may be turns in business conditions at large or turns in a designated reference activity).

It might appear that the determination of timing measures is simple, once the specific and reference turns are established. This is indeed the case, if the series is identified as positively or inversely conforming, if every specific turn can be matched with a reference turn, and vice versa, and if no opposite turns occur between matching peaks and between matching troughs. However, a number of difficulties occur when these conditions are not met; these, in fact, led to the prediction that timing comparisons would prove intractable to a programmed approach. The primary problem is to determine which specific turn should be related to which reference turn. This decision may involve many considerations, such as proximity, typical timing behavior at other turns, and amplitudes. In the present program, which is still in the early stages of development, we restricted the decision rules to proximity and typical timing, but we shall consider other criteria if this proves necessary. We are fairly confident that timing comparisons can be successfully programmed.

Charlotte Boschan