

also State and local government groups, businesses, and other Federal agencies, often place a severe and unwanted burden upon data-processing facilities and the time and energy of specialized personnel. Even when such work is done on a reimbursable basis, limitations of staff due to overall personnel and budgetary considerations and the ability to hire people making the filling of special requests a burden. Outside requests for data are often uninformed, unreasonable, and in view of the Federal agency, not worth while. Few outsiders can know enough about the data, their nature and characteristics to make sensible requests, or to have a realistic appreciation of the analytic limitations which the data impose. As already indicated, the disclosure problem is formidable and causes considerable uneasiness on the part of the responsible people in the data-producing agencies, but it also may be very useful as a shield to protect them from the nuisance of dealing with individual requests. The problem, as seen by the Federal Government as a whole when contemplating a request for data, could be reduced if research workers asking for data could get together and coordinate their requests.

#### ECONOMIES OF SCALE AND THE NEED FOR COORDINATION

Many of the requests for basic information on a specific subject by different research scholars are duplicative. However, since each research project will be designed in somewhat different terms and has different objectives, it is inevitable that the independent requests for information will not be identical. On the other hand, it is also quite possible that, if careful consideration were given to the matter, general master tapes might be designed in specific areas which would meet the needs of a large number of research projects. One of the difficulties with published information is that different research workers want different types of aggregations and classifications. Since it is now possible to provide data on a disaggregated basis, these differences are no longer relevant, and it becomes necessary only to specify the basic items of information to be included in the body of data.

This basic similarity in the demand for information on a given topic implies that considerable economies of scale could be achieved by coordination. If a single master tape would fully satisfy the demands of each user, designing special tapes for each user would be unnecessary. For this reason, the committee has undertaken a preliminary survey of 10 major areas of economic data to see whether or not it would be possible to construct sets of such basic tapes in these areas. On the basis of this examination it is the committee's considered conclusion that this construction not only is feasible from the point of view of economic research needs and objectives, but also would go a long way toward improving access to major bodies of data for scholars, and toward reducing the costs and alleviating the burden placed on the Federal statistical agencies.

#### DEVELOPMENT OF TAPES FOR SPECIFIC RESEARCH AREAS

The committee circulated to a group of research scholars working in various areas copies of the preliminary inventory of machine readable data recently collected by the Bureau of the Budget and the National Archives. In a large number of cases, these scholars prepared suggestions as to bodies of data currently in existence, which should be developed and made available to universities and research institutions on a low-cost basis. Included among the suggested data files are some which are currently available to research scholars and which certainly should remain available. For example, the Bureau of the Census has developed a program of making available for purchase large bodies of unpublished data in the form of computer tapes. The Internal Revenue Service and the Bureau of Labor Statistics have also developed specific tapes for sale. However, even where unpublished material is available on computer tape, it is often not in a form which is directly useful to the research worker. Data reduction to prevent disclosure or to select a manageable sample of data may be necessary. It would also be very useful if the research community could be better informed about what tapes exist in the various Federal agencies, and the cost of obtaining them.

A number of different data characteristics which are important for research purposes have been mentioned by research workers. (1) They point out that sets of data that are continuous over time are particularly valuable. This is especially true where information relating to a specific reporting unit is obtained at regular intervals so that changes taking place at the micro level could be ob-