

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