

Cynthia Z. F. Clark and Jerry L. Coffey, Office of Management and Budget

INTRODUCTION

As members of the statistical profession are well aware, data collection and analysis has acquired vast significance in assessing the status of our society. In order to set public policy and meet the data needs of the general public, the United States government has established statistical agencies and units within many executive departments and independent agencies. The data collections of these agencies serve their specific policy and program interests, but many components of the collections overlap.

In the collection, processing, and analysis of these data, protecting the confidentiality of individual responses to statistical inquiries is of paramount concern to these statistical agencies. Protection of statistical data is considered to be essential to ensure a high level of voluntary participation. Accurate survey data obtained at a high response rate are needed to make statistical programs reliable.

To establish confidentiality protection we define: for what purposes the data may be used, by whom, and under what conditions. A dilemma arises in formulating confidentiality protection because such protection requires specifying permissible access to protected data. The statistician has long asserted that data collected for statistical purposes should be used only for statistical purposes. Different agencies have used different means to restrict use of its data and to safeguard its data. To provide tight controls over dissemination of data, agencies have often adopted restrictive policies proscribing release of data outside their agency--policies independent of the purposes for which, or conditions under which, that data ideally would be used. Some agencies feel that such restrictions are necessary to maintain the agency's image as a protector of confidential data. Other agencies have developed procedures that have permitted limited data exchanges for statistical purposes under stringent controls.

Because the Federal statistical system is decentralized, statistical programs are scattered throughout the various departments and independent agencies. The gathering, processing, and statistical analysis of individual agency data sets results in much duplication of effort (for example, in

sample frame development and coding operations). Many statistical programs would benefit if it were possible for statistical agencies to share data files and yet maintain their confidentiality. This exchange of data files would have to be limited to actions that further the statistical missions of the agencies involved, and should be accompanied by strong safeguards.

If strong and uniform confidentiality protection were available across statistical agencies, it would create an environment in which data could be shared with minimal risk to the confidentiality afforded individual data subjects. Such protection would require that each agency establish appropriate administrative, technical, and physical safeguards for the statistical data it collects or maintains. Procedures would need to be specified to effect secure data exchanges among these statistical agencies. This framework would enable statistical programs to produce more reliable data more efficiently.

Statutory confidentiality protection has been sought by agencies collecting statistical data. They see it as the most desirable means of demonstrating that assurances made to respondents will be honored. Some agencies have already been given statutory confidentiality protection. The protection most publicly recognized is the Title 13 confidentiality statute of the Bureau of the Census. The National Center for Health Statistics in the Department of Health and Human Services also has statutory protection, as does the Bureau of Justice Statistics in the Department of Justice and the Bureau of Economic Analysis in the Department of Commerce. Agencies collecting statistical data that do not have such statutory confidentiality protection must rely upon general statutes and agency regulations or policies. A number of agencies in this position have made a case for specific statutory protection in terms of both public perception and administrative costs.

Both the Bureau of Labor Statistics in the Department of Labor and the Statistical Reporting Service in the Department of Agriculture have proposed legislation that would satisfy the specific needs of their agency. Single agency proposals suffer from a common weakness--in addressing only the needs of that agency, their cumulative effect may be to severely restrict or handicap any sharing between the statistical agencies. A preferable legislative

approach seems to be to provide uniform statistical confidentiality protection for qualifying data files with appropriate sanctions and penalties for infractions of the law. Such protection could be drafted so that it would not affect existing agency exchanges of data for statistical purposes.

Some agencies that already have statutory confidentiality protection are concerned that such broad legislation will weaken the protection afforded by existing statutes and, thus, adversely affect their ability to carry out their responsibilities. Extending confidentiality protection similar to that provided by Title 13 to named statistical agencies, among whom statistical record files could then be exchanged, does expose confidential data to additional risks from several sources. Would an increase in the number of people with access to any given data file affect the confidentiality of that data? Would the existence in different agencies of somewhat different procedures and confidentiality safeguards (or administration of those procedures or safeguards) lessen the protection for exchanged data? What would be the public perception of the potential for misuse of data exchanged among these agencies? If it is to be perceived as providing strong protection, a confidentiality statute needs to mitigate these risks.

#### HISTORY OF CONFIDENTIALITY LEGISLATIVE INITIATIVES

The earliest statute to control government uses of private information was the Revenue Act of 1864. This statute, designed to deter corruption of tax agents, was re-enacted on several occasions and remained in effect until 1948. Early in this century, a permanent organization, the forerunner of the Bureau of the Census, was established to conduct censuses. This action was accompanied by enactment of permanent law to protect confidentiality of census data.

The 1938 confidentiality law for the Bureau of Foreign and Domestic Commerce (later the Bureau of Economic Analysis) had consequences far beyond the Commerce Department. Its premises were recognized in the 1948 Trade Secrets Act, which added "confidential statistical data" to the categories of business information protected under the Criminal Code. This Act combined elements of the Revenue Act, a Tariff Commission statute and the Commerce Department statute to produce a broader law than any of its components. Covering all Federal officers and employees, it prohibited any disclosures

that were not authorized by law. This was the high water mark for confidential treatment of statistical data gathered from business establishments.

But other policies were already beginning to impinge on confidentiality. The Federal Reports Act of 1942 authorized interagency disclosures. Though the bounds of this authority were never clear, it was used by OMB on rare occasions to overrule the disclosure policies of Federal agencies. In the early 1950's, the authority of cabinet officers was strengthened throughout government, sometimes at the expense of agency-oriented confidentiality law. In the 1954 revision of Title 13 responsibility for censuses was assigned directly to the Secretary of Commerce, reflecting the Secretary's authority under a 1950 reorganization to use any officer or employee of the department to perform these functions. Conforming changes made the confidentiality restrictions of Title 13 binding on the entire Commerce Department and extended the statutory boundary for access to the entire department where it remains today. The Census Bureau and other statistical agencies dealt with this restructuring of authority by advancing exclusive delegations of authority and other restrictive agency and departmental policies to prevent improper use of statistical data in the new legal and organizational environment.

In the early 1960's, Congress became disenchanted with the failure of the Administrative Procedures Act to increase the flow of information to the public and thus mandated a broad range of disclosures in the Freedom of Information (FOI) Act of 1966. This reversal of previous disclosure policies put the burden of proof on agencies desiring to protect information from public scrutiny and inaugurated an era of uncertainty which affected most statistical agencies. Many agencies have now tested their confidentiality policies in Congress or the courts. The Bureau of Economic Analysis' exemption was affirmed in 1976, Title 13 was amended in 1980 to block FOIA access to Shippers' Export Declarations, and the Bureau of Labor Statistics' policies, based on the Trade Secrets Act, were found to support an FOIA exemption in a 1981 case. In spite of these positive actions, public confidence in statistical confidentiality still waivers hanging on the outcome of each new FOI court case.

Concern about the treatment of personal as well as financial information produced the Privacy Act of 1974. The law is much less binding than the Freedom of Information Act in the sense that it allows a great deal of

agency discretion. It also relies on individuals to enforce their own rights rather than imposing rigid standards on Federal agencies.

In 1976 policies concerning the disclosure of tax return information were codified in the Tax Reform Act. While this Act laid out specific prescriptions for handling tax information, Congress strained credulity by declaring tax returns confidential. This action probably did more to undermine the meaning of confidentiality than it accomplished in protecting return information, as the patchwork of exceptions to the general rule included many disclosures which could adversely affect the individual interests of taxpayers.

As information law became more complex, several commissions and a Presidential task force attempted to define appropriate policies for statistical information. In 1971 the President's Commission on Federal Statistics recommended that confidentiality be exclusively linked to specific legal authority and that all agencies collecting data for statistical purposes be granted such authority. It also advocated a strict definition of confidentiality (linking it to non-threatening uses of information) that would have avoided the Tax Reform Act's assault on public credulity. Finally, it recommended that the Statistical Policy Division of OMB be assigned broad powers to authorize data transfers for statistical purposes.

The Privacy Protection Study Commission and the Commission on Federal Paperwork both released reports on privacy and confidentiality in July of 1977. The Privacy Commission elaborated the doctrine of "functional separation" (the principle that information collected or disclosed for statistical purposes should not be used for any other purpose) and recommended detailed prescriptions for the protection of statistical data (and administrative data disclosed for statistical purposes). The Paperwork Commission generally endorsed the findings of the Privacy Commission but recommended simpler legislative solutions.

President Carter's Task Force for Reorganization of the Federal Statistical System (the Bonnen Task Force) developed the Privacy and Paperwork recommendations into a comprehensive legislative proposal in 1979. However, further pursuit of this initiative was deferred in favor of the Paperwork Reduction Act of 1980. This Act thus became the only legislation to reflect to any extent the work of the three Commissions. It clarified and thus gave substantial new force to OMB's Federal Report's Act authority to direct

interagency disclosures, and required that "legislative proposals to remove inconsistencies in laws and practices involving privacy, confidentiality, and disclosure of information" be submitted to the President and Congress.

In March of 1982, OMB circulated for agency comment a draft bill quite similar to the one developed earlier by the Bonnen Task Force, a bill supported by the statistical agencies in the context of the other recommendations of the Task Force. This version envisioned a key role for strong, centralized decision-making authority with respect to permitted data exchanges. Since the statistical policy function had been transferred (in the Paperwork Reduction Act) to the Office of Information and Regulatory Affairs (OIRA) in OMB rather than the independent central statistical office recommended by the Bonnen Task Force, the statistical agencies did not find such centralized authority to their liking. Other comments received from the agencies suggested that other parts of the draft legislation needed to be rethought. Later in 1982, the confidentiality and statistical data exchange needs of the major statistical agencies were brought to the attention of the reconstituted Working Group on Economic Statistics of the Cabinet Council on Economic Affairs, chaired by William Poole of the Council of Economic Advisers. The interest of this group, coupled with the directive to OMB found in the Paperwork Reduction Act, led OMB to launch a major effort to rewrite the Bonnen confidentiality proposal early in 1983.

#### DESCRIPTION OF OMNIBUS CONFIDENTIALITY INITIATIVES

All versions of this legislative initiative involved the creation of two new legal concepts: "protected statistical file" and "protected statistical centers." The procedures for designating such "protected statistical files," and the statutory prohibitions against release of information from these files, are designed to restrict their use to statistical purposes only. A limited number of statistical agencies would be designated as "protected statistical centers," and would be authorized both to create and to receive protected statistical files. Other agencies would be authorized to protect their own statistical files, but would not have any privilege of access to protected statistical files other than their own.

Protected statistical files could be created by (1) designation prior to the collection of the information which constitutes the protected statistical file; (2) incorporation of information

from a protected statistical file (with or without information from any other source) into a different file; (3) acceptance, by a protected statistical center, of information to be used exclusively for statistical purposes, when such information is derived from a file other than a protected statistical file; (4) designation of existing files containing information collected under a pledge of confidentiality if that pledge is not inconsistent with the provisions of this legislation.

Once a file had been designated as a protected statistical file, it could be released only to employees of the agency that created it, to a protected statistical center, or, for certain limited nonstatistical uses specified in the legislation, to other parties. Such permitted nonstatistical uses include referrals to the National Archives, internal audits of statistical programs, investigation of violations of the confidentiality statute, situations involving disclosure with consent, and prosecutions under statutes mandating truthful response. No exceptions for general law enforcement use, for national security investigations, or for Freedom of Information requests have been granted in any of the versions of this legislative initiative.

Since the protected statistical centers would potentially have access to most Federal files, their designation cannot be taken lightly. The most credible means are (1) initial designation in the legislative proposal, or (2) later action taken by the President. Candidate centers must reflect a high degree of functional separation of statistical activities from those that are nonstatistical. Comments on the legislative proposal circulated in March 1982 reflected a strong interest in making the individual protected statistical centers more accountable, as well as an interest in reducing the pressure to designate new centers. Interest was also shown in developing a mechanism for a quick suspension of the access privileges of a protected statistical center pending a review to determine whether protected statistical center status should be revoked.

Several operational mechanisms have been considered for achieving the objectives of confidentiality and data exchange. The approach chosen maintains the framework of the original proposal of the Statistical Reorganization Project, but reduces the reliance on OMB decisions by delegating decision-making authority for data releases to the Secretaries of the Departments in which the protected statistical centers are located. This shifts the balance between confidentiality and data exchange so as

to place greater emphasis on reducing confidentiality risks (consequently reducing data exchange).

Following that scenario, each Department or Independent Agency head would be given the authority to designate protected statistical files, and to voluntarily disclose information (for specific statistical purposes) to a protected statistical center. New powers are to be structured and understood as an extension of the responsibilities imposed by current law derived as they are from authority to protect the confidentiality of data used for statistical purposes. Numerous tests have to be met and determinations made before a protected statistical file could be released by the agency head. These would include limitations on disclosures from such comprehensive statistical or administrative files as those compiled from censuses or tax returns. In a case of refusal to disclose, a protected statistical center could appeal to the Director of OMB to exercise the authority of the Paperwork Reduction Act to order disclosures not inconsistent with applicable law. Such a mandated release could be accompanied by restrictions extending beyond those required by the proposed confidentiality statute.

The March 1983 proposal established a committee to advise the Director of OMB on matters concerning the confidentiality of statistical records and the disclosure of records for statistical purposes. Agency comments recommended that this committee include the heads of protected statistical centers and representatives of other interests. The committee would review any proposal to designate a new protected statistical center as well as proposed confidentiality and disclosure rules, policies, or procedures of protected statistical centers, and other agencies.

An important objective of the omnibus confidentiality approach is to ensure that existing responsible cooperative data sharing arrangements are not impeded. To meet this need, OMB has refined the concept of a "cooperative statistical agreement." Such an agreement extends confidentiality protection and limited access, coupled with legislative sanctions and penalties, to any parties cooperating with a protected statistical center as if such parties were actual employees of the center responsible for the protection of the data.

The potential for the exchange of protected statistical files among protected statistical centers accommodates the interagency use of certain kinds of "directory information". The interagency statistical use of a common business

directory is an example of one type of "directory information" that has been much discussed among the major statistical agencies. Early in 1974 funds were appropriated to the Census Bureau to develop such a list and the Commerce Department was given the responsibility to develop legislation to amend Title 13 so that the list could be shared for statistical purposes among federal and state agencies. (Legislation to accomplish this objective has been stalled during this period, largely because of lack of agreement among the agencies contributing data respecting identity of the potential users and the conditions to be placed on the use of the list information.) Similar exchanges of names and addresses of individuals would make it possible to coordinate complex sample designs among protected statistical centers and would open up the possibility of contracting for high quality sample design and selection services among these centers.

#### CONCLUSION

Although the recent version of the omnibus confidentiality legislation has received support from many of the statistical agencies it does not yet have the concurrence of either the Departments of Commerce or Treasury. Some of the issues which they have raised can, we believe, be accommodated with minor changes in the 1983 draft. Under the assumption that an omnibus approach is preferable for addressing the confidentiality and statistical data exchange issues, OMB has continued to fashion compromises and attempts to negotiate with the affected Departments. The confidentiality proposal would establish in statute, with some generality, the best of the proven policies and procedures developed by the statistical agencies. The posture of OMB has been to create stringent safeguards and procedures for all uses and exchanges of protected statistical files, and enforceable sanctions and penalties for any infractions of those procedures.

Will this approach provide adequate protection to mitigate the risks to the statistical system occasioned by any perceived misuse of confidential information? In the hands of professional statisticians, at least, and for studies and analyses serving statistical purposes, the omnibus confidentiality legislation would be both tool and shield, aiding their labors and protecting them (and the statistical processes they employ) against untoward pressures to make all information available for policing rather than prediction.

Can the public be reassured? We believe it can--if we demonstrate that our government is serious about using the best available data, impartially, in a statistically sound way. That quest would be aided by legislation protecting confidential information being used for statistical purposes from disclosure for other nonrelated uses. This approach recognizes that much more data is needed to provide reliable statistics for our society. It provides more material for the processes involved in compiling these statistics. It protects the information used to perform statistical functions. An advance for both statistics and public confidence can be achieved.

#### ACKNOWLEDGEMENTS

The authors would like to thank Lois Alexander, Catherine Miller, Robert Mugge, and Peter Yates for their contributions over a long period of time in developing the issues discussed in this paper. They also appreciate the full support of Christopher DeMuth, Thomas Hopkins, and Dorothy Tella during the time when the paper was being prepared. Comments from Louis Kincannon and Theodore Clemence helped to clarify key points in the narrative.

#### REFERENCES

Business Directories: Findings and Recommendations of the ASA; ASA Committee on Privacy and Confidentiality, The American Statistician, February 1980, Volume 34, Number 1.

Chrysler Corporation V. Brown, Secretary of Defense, et al. Supreme Court Decision, April 18, 1979.

Confidentiality and Privacy; Report of the Commission on Federal Paperwork, U.S. Government Printing Office, Washington, D.C., July 1977.

Federal Statistics; Report of the President's Commission on Federal Statistics (two volumes), U.S. Government Printing Office, Washington, D.C., 1971.

Paperwork Reduction Act of 1980; House of Representatives Report Number 96-835, and Senate Report Number 96-930.

Personal Privacy in an Information Society; Report of the Privacy Protection Study Commission, U.S. Government Printing Office, Washington, D.C. July 1977.

Privacy and Confidentiality as  
Factors in Survey Response; Committee on  
National Statistics, National Academy of  
Sciences, Washington, D.C. 1979.

Statistics; Report of the Commission

on Federal Paperwork, U.S. Government  
Printing Office, Washington, D.C., 1977

United States Code Annotated; West  
Publishing Company, St. Paul, Minn.