

DISCUSSION

Joseph W. Duncan, The Dun & Bradstreet Corporation

The papers presented in this session represent a thoughtfully organized presentation of critical issues and concerns about the confidentiality of Federal statistics. The organizer, Wendy Alvey, deserves a special compliment since the papers from this session should provide an excellent reference concerning the historical perspective on the issues, examples of pressures on the statistical system and a clear articulation of critical issues.

Concern about the confidentiality of Federal statistics begins with the fact that few agencies have secure statutory protection to back up their promise to respondents that information provided will be held confidential. As stated in the Clark/Coffey paper, the legislation is thus both "a shield and a tool;" a shield for respondents and a tool for sharing among agencies.

Before discussing the individual papers, let me set forth my statement of the purpose of proposed confidentiality legislation. This legislation should be designed to achieve three goals:

1. To assure the cooperation of respondents by protecting information from unauthorized access. Statistical activities should not affect the rights, benefits and privileges of individual respondents.
2. To improve statistical surveys by exchanging information concerning both sample selection frames and key data parameters.
3. To assure consistency and integration in statistical concepts, definitions and classifications, e.g., by assuring that industry is uniformly defined in various surveys.

The proposed legislation has not been released for public comment, therefore the degree to which these objectives are achieved is still uncertain. Nevertheless, in the current political context, it is evident that the decision process for achieving coordination of data exchange among the "protected" or enclave agencies must be strong, with both independence and clout, for the coordination function to ensure a.) the integrity of statistical activities and b.) agency coordination and collaboration.

Before reaching some general conclusions about the overall issue, I will discuss briefly several points covered in the individual papers.

SPECIFIC COMMENTS

Clark/Coffey - "How Many People Can Keep A Secret? Data Interchange Within A Decentralized System"

This paper provides a useful overview of the soon to be released Office of Management and Budget (OMB) draft bill. Clark & Coffey agree with my introductory point that a key issue is the "role for strong, centralized decision-making authority," especially to resolve conflicts among the agencies with respect to whether or not data is to be released. Unfortunately, the paper creates some confusion concerning the manner in which the strong, centralized decision-making authority is actually implemented in the proposed legislation.

As described in the paper, the agency objections have been met by providing for decentralized authority to release files; authority which is essentially delegated to the Cabinet Secretaries. It is unclear, given this decentralized authority, how the centralized role of OMB will be exercised and, consequently, how proper coordination will be achieved.

This dilemma is illustrated in the paper's discussion about the current difficulty of obtaining agreement among agencies concerning the identity of firms necessary for the "directory information," which has been developed by the Census Bureau through the Standard Statistical Establishment List (SSEL). Apparently conflicts remain among the Internal Revenue Service, Social Security, Census and other agencies, in determining whether or not the information on firm identity can be released. With decentralized authority resting with Cabinet Secretaries, it is not clear that the proposed legislation will meet the second and third objectives which I mentioned earlier.

Alexander - "There Ought To Be A Law..."

It is noted that the proposed legislation would implement the Paperwork Reduction Act provision for assuring the authority of the Director of OMB "to compel disclosure of data to Protected Statistical Centers from essentially any record source, if need were demonstrated." This feature of the proposed legislation is especially important. This quote implies that the proposed legislation establishes a strong centralized control; not the decentralized authority discussed by Clark & Coffey.

The Alexander paper does a fine job of illustrating some of the problems which may

be associated with requests for disclosure. The case of the linkage of death certificate information from state vital records to identify fraud in the Social Security Administration clearly illuminates the potential conflict.

After a useful discussion of issues and a presentation of several very interesting examples, the paper does not present direct conclusions on the pros and cons of the proposed legislation. Specifically, the role of the Director of OMB and the coordinating function merits considerably more attention than is presented in the paper.

This paper underscores the importance of reviewing the principles underlying any proposed confidentiality legislation. The specific examples of problems provide practical demonstration of the issues which must be resolved.

Mugge - "Issues In Protecting Confidentiality In National Health Statistics"

Mugge's policy objectives for confidentiality policy at the agency level merit repeating. He states that agencies must set policies to:

- (1) provide informed consent to study subjects
- (2) provide physical protection of records
- (3) assure training and supervising of staff on confidentiality matters
- (4) avoid statistical disclosure through published tables or data tapes
- (5) establish appropriate limits to sharing information with the researchers
- (6) determine what else is necessary to meet the letter and spirit of existing laws and regulations, and
- (7) determine what changes should be requested in laws and regulations.

Mugge characterizes the position of NCHS on the legislative proposal as follows: "It is my view that even if NCHS got nothing directly out of the new law for itself, it should still support it, because it is in the best interest of the total Federal statistical system." This reflects the Center's view that it presently has sufficient legal protection for its confidentiality status.

Mugge notes that the new bill should provide a benefit to NCHS by providing a more direct access to a sampling frame for household related surveys. This potential benefit underscores the fact that, while the Center has reasonably strong legislation at the moment, it could also directly benefit from the proposed legislation.

Wilson/Smith - "Access To Tax Records For Statistical Purposes"

The papers provide a useful reference of

much of the debate around privacy, freedom of information and the characteristics of specific legislation. It is important to carefully review recent legislative history and the concerns of Congress to understand the mood that exists around the confidentiality issue. I think that the "mood" is the Achilles heel of the process right now. In this context the Wilson-Smith paper provides an excellent legislative history of the concerns about access to tax records. The comments on the Privacy Act and Tax Reform Act are especially perceptive.

The authors correctly point out that, even if controls are provided concerning redisclosure of information, there is a danger of problems in the public's perceptions. They state, "if the perception is created that tax return information is widely available to others, that perception alone would be very destructive of public confidence in the security and privacy of information provided under the tax administration system, and ultimately destructive of voluntary compliance." This same concern could be related to statistical programs, including the Census and other voluntary information provision for statistical purposes. The authors conclude that the IRS is uncomfortable with confidentiality legislation because it would greatly expand the dissemination of tax return information. Public perception is very important.

The authors also raised the possibility of the Bureau of Census becoming an autonomous but integral part of the Department of the Treasury. Their question is "Will this move result in better reporting once realization dawned that the requirement to file tax returns and census surveys is administered by the same department?" It seems clear on an a priori basis that combining tax collection and census data collection in the same department will intensify public concerns about "Big Brother." Respondents already fear that statistical information is being used for administrative purposes; the linking of pervasive tax collection with statistical collection would certainly intensify that concern. To restate, "Public perception is important."

Spruill - "Testing Confidentiality of Masked Business Microdata"

This paper presents an illustration of the use of statistical techniques for "camouflaging" data, so that the characteristics of individual entities are not revealed. While the author indicates that the results are inconclusive, they do generally demonstrate the fact that statistical adjustments to data permit the development and testing of reasonable hypotheses even though the individual observations have been altered.

The use of statistical techniques, such

as random rounding and data swapping, will frequently make it possible to release individual records without revealing individual characteristics. While such techniques are not an essential element of the proposed confidentiality legislation, they do illustrate the opportunity which statistical agencies have for meeting the demands of external researchers without compromising their confidentiality standards.

Thus, I agree with Louis Kincannon's discussion point that the work represented by this paper is especially crucial to the long term evolution of confidentiality policy. While more research is clearly needed, it is significant that tools such as those explored in this paper will help resolve many of the problems of releasing micro-data sets for deeper analysis.

Spruill makes the important point that individual data releases must be examined and that universal techniques will probably not be suitable. Nevertheless, tailored approaches can be of great value for releasing micro-information to those who seek to benefit from analysis of micro-data.

CONCLUDING REMARKS

Discussions concerning the confidentiality of Federal statistics have been underway for more than a decade. I first presented the concept of statistical enclaves at the ASA meeting in 1975. [1] It remains an important topic to keep in front of us today.

Despite this long and sustained history of interest in the subject, it should be obvious that the year "1984," with both its Orwellian context and political (Presidential election year) fervor, is the wrong time to propose such legislation. It will inevitably be misinterpreted, as the data bank proposals of the mid-1960s were misinterpreted. The fear of the computer, "Big Brother's" intrusion, record matching to reduce fraud, invasion of privacy, and all of the other rhetoric will inevitably submarine what might otherwise be a sound and highly desirable strategy to improve Federal statistics.

In political terms, the initiative represented by the confidentiality legislation should be fostered as an initiative of the "newly elected" administration in 1985 (regardless of whether it is a Republican or Democratic Administration). It should also be

accompanied by a more comprehensive revitalization of the entire Federal statistical system, including the central coordination unit.

Finally, the Wilson paper noted that there is a current proposal to accompany the Trade Reorganization Act with a relocation of the Bureau of Census to the Treasury Department. In light of the concerns which have been expressed in these papers about building public confidence in the integrity and security of Federal data, it seems unwise to move the Census Bureau into "apparent partnership" with the IRS. Wilson's paper underscores the concern of IRS that voluntary cooperation in tax filing would be harmed if individuals felt that data submitted were freely available. Likewise, Census respondents would prove to be reluctant to participate on a voluntary basis in surveys if they anticipated that information provided would be checked with the IRS to determine if their tax returns were accurate. Thus, the discussion of principles of confidentiality legislation is important, in the short run, in addressing the proposed transfer of the Census Bureau. It seems to me that the worst location for the Census Bureau is the Treasury Department.

On balance, the short term issue of Census reorganization needs to be resolved quickly and then proper groundwork must be put in place for considering the proposed confidentiality after the 1984 elections.

Statistical policy must consider not only the philosophical, technical issues that we discussed today, but the practical, political environment that's involved.

FOOTNOTES

- [1] Duncan, Joseph W. "Confidentiality and the Future of the U.S. Statistical System," American Statistical Association Proceedings, Social Statistics Section, 1975, pp. 59-64.

EDITOR'S NOTE

For space reasons the Wilson-Smith paper no longer contains any reference to the proposed relocation of the Census Bureau to the Treasury Department. Readers wishing more information on this topic should consult the newsletter of the Council of Professional Associations on Federal Statistics, News from COPAFS, No. 17, August 1983.