ETHICS AND EFFICACY IN RELEASING OFFICIAL STATISTICS TO THE PUBLIC AND DATA USERS

Nancy M. Gordon
Associate Director for Demographic Programs
U.S. Census Bureau, Room 2061/3, 4700 Silver Hill Road, Stop 8800, Washington, D.C. 20233-8800

The substantial contributions of Wendy Alvey, Joan Bernard, and Marjorie Hanson are greatly appreciated.

This paper has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. It is released to inform interested parties about the Census Bureau's commitment to protecting confidentiality while facilitating access to data, and to encourage discussion of these important issues.

KEY WORDS: U.S. Census Bureau, Confidentiality, Data Collection, Ethics, Privacy

INTRODUCTION

An inherent tension exists between an individual's right to privacy and the country's need for extensive, timely, and accurate data that enables government, business, academia, and myriad organizations to function effectively. The U.S. Census Bureau is at the center of this tension, since its mission is "to be the preeminent collector and provider of data about the people and economy of the United States," and to do this it must maintain the public's trust that it will safeguard the confidentiality of the information it receives. In other words, it must fulfill its ethical responsibilities to individuals and firms who respond to questions, but it must also enable the country to make use of the information embedded in its data.

In 1971, the President's Commission on Federal Statistics defined the right to privacy as individuals' right to decide whether and to what extent they will reveal their views to the government or answer specific questions about themselves and their circumstances. The commission noted that confidentiality requires that restrictions be imposed on how information can be transmitted and used. Its report said that a commitment to privacy requires that publicly released information not enable someone to identify respondents or to harm them, which includes a prohibition on the use in legal proceedings of information collected for statistical purposes about a particular person or entity.

The American Statistical Association (ASA) has been a leader in promoting ethical conduct for data protection. Its 1997 "Report of the Ad Hoc Committee on Privacy and Confidentiality" documented the increasing public concern about issues of personal privacy, confidentiality, and freedom of information arising from the information gathering processes of governmental institutions.

These issues were further articulated in the ASA's "Ethical Guidelines for Statistical Practice," prepared by the Committee on Professional Ethics and approved by the Board of Directors, August 7, 1999. While noting that governmental decisions regarding a wide array of important matters (e.g., public health, criminal justice, education, and the environment) depend in part on sound statistics, the guidelines stress the importance of protecting the privacy and confidentiality of research subjects and data concerning them, and they specifically mention the statutory responsibility of the Census Bureau to maintain the confidentiality of a respondent's data.

As concerns grow about threats to the privacy of information posed by advanced technology and the capabilities provided by the internet, and about the intrusiveness of government inquiries, the Census Bureau must effectively collect greater quantities of high quality information that will provide data users the accurate and timely information they need to do their jobs efficiently and effectively.

By law, the Census Bureau may not publish data about a particular establishment or individual in a way that allows them to be identified. This "zero tolerance" for disclosure is backed up by criminal penalties and reinforced by a strong commitment to ethics and professionalism in the Census Bureau's culture. The Census Bureau works constantly to improve its policies, systems, and procedures, and continually examines the ethics, efficiency, and practicality of various means to enhance the utility of its data files to government, researchers, and the public, while meeting its requirements for nondisclosure.

The Necessity for Privacy

Protecting respondents' privacy is ethical conduct to which the Census Bureau is fully committed. When the Census Bureau asks for information, it promises the respondents that their answers and identity will remain confidential. The Bureau gives its word and keeps it.

All employees and non-employees who have access to confidential data take an oath that they will protect its confidentiality, and this responsibility permeates the Census Bureau's policies and ethos. Even without the threat of legal sanctions, the Census Bureau's own standards of professionalism lead it scrupulously to protect the confidentiality of its data because that is the ethical thing to do.

Under Title 13, U.S. Code, the Census Bureau is legally required to maintain the confidentiality of private information, and personnel who violate this law are subject to fines and even imprisonment. Some other federal statutes also apply, such as the Privacy Act and the Paperwork Reduction Act, which require confidentiality and informed consent in data collection and related activities. By agreement with the National Archives and Records Administration, decennial census and household survey information is protected for 72 years after collection and business data cannot be made available for 30 years. Furthermore, in many cases – particularly in the economic area – Internal Revenue Service restrictions also apply when federal tax information is used to improve and evaluate Census Bureau programs.

Beyond professional ethics and legal strictures, however, protection of privacy is efficient. From a purely practical viewpoint, the Census Bureau needs the cooperation of everyone in the country to carry out its mission. If individuals refuse to answer questions or to complete surveys or census forms truthfully, the validity of the data is undermined. People need to trust that the information they are revealing will be held in strictest confidence and will not be used for any purpose other than anonymous statistical information.

The Congress and the Executive Branch are also concerned about the availability and validity of data and about maintaining the privacy of individuals' responses. Their constituents forcefully convey their privacy concerns and, at the same time, elected officials, the administration, and other government agencies depend on accurate data as the basis for decisions on crucial matters ranging from redistricting to social programs and economic policies.

Data from the Census Bureau are the foundation for efficiency in governing, in creating a strong economy and social fabric, and in enabling the private sector to thrive. The private sector requires accurate information in order to allocate resources efficiently and effectively.

An interdependence links the Census Bureau's need for complete and accurate responses, the country's need for Census Bureau data, and the benefits individuals reap from living in a country which has efficient government and a strong economy. This mutually beneficial system will only work if individuals know that their privacy is protected and their information will remain confidential.

THE CENSUS BUREAU'S RESPONSE

The following section describes the Census Bureau's approach to protecting respondents' confidentiality and then illustrates its application in a specific policy area.

Administrative, Technical, and Programmatic Controls

The Census Bureau goes to great lengths to ensure that individuals and businesses cannot be identified through data that are publicly released and that the data are protected from unauthorized use. This responsibility begins as soon as the

data are collected, by whatever means, and continues through the processing, release, and archiving of the data. For example:

- The Census Bureau carefully trains its field staff about the importance of Title 13 confidentiality and the need to protect data.

- The Census Bureau has a broad range of technical controls in place to monitor its technical systems for potential vulnerabilities. Physical and information technology guidelines specify the physical constraints to maintain a secure environment. Tests and audits ensure that adequate security controls are being implemented on existing systems, and new technologies are pre-tested before deployment to demonstrate safety and security.

- The Census Bureau has developed a set of procedures to identify potential unauthorized intrusions into its network systems. The procedures cover the identification, notification, recovery, and recertification for any level of attack reported. The Census Bureau employs multiple Intrusion Detection Devices, both outside and inside its firewall. These systems compare inbound and outbound traffic against a regularly updated table of attack signatures that can identify known attacks.

- The Census Bureau uses comprehensive and consistent disclosure avoidance techniques to reduce the probability of a respondent being identified. Examples include stripping identifiers, recoding, top coding, data swapping, geographic limitations, and suppression.

- The Census Bureau limits the geographic area that is identified for released data. In general, microdata files are not released for areas with a population less than 100,000 or 250,000. Tabular data are subject to additional monitoring by the Disclosure Review Board for "complementary disclosure," instances when confidential information could be identified through subtraction. Filtering techniques are also applied to prevent disclosure in vehicles such as the American FactFinder.

Policy Controls: the Data Stewardship Executive Policy Committee (DSEP)

In addition to administrative and programmatic controls, the Census Bureau has established the Data Stewardship Executive Policy Committee (DSEP), as part of its continuing effort to improve the quality of its policies and procedures. The DSEP's mission is to ensure that the Census Bureau can effectively collect and use data about the country's people and its economy, while fully meeting the legal and ethical obligations to respect respondents' privacy and protect the confidentiality of their information. To do that, the DSEP serves as a focal point to address a broad range of data stewardship issues related to privacy, confidentiality, security, and administrative records.

One of the DSEP's first efforts was to adopt a set of Data Stewardship Principles, which support the Census Bureau's mission. These principles and the accompanying commentary clearly articulate the Census Bureau's longstanding position on privacy and confidentiality, and provide the basis for its data stewardship policies and procedures.

The policy issue described in the next section illustrates how the DSEP balances the need to protect privacy with the Census Bureau's mandate to release important information on a broad scale to the public, in this case by allowing people who are not employees of the Census Bureau to have access to confidential data under carefully regulated circumstances. It is often efficient and desirable for non-employees to be allowed access to confidential data, but that is only permitted under special procedures that have been designed to ensure that the Census Bureau's standards for privacy and nondisclosure are maintained.

Policy on Non-Employee Access to Title 13 Data

The Census Bureau's policy on non-employees' access to confidential data provides a case study of proactive, thorough attention to the details required to maintain the privacy of the individuals and firms who provide confidential information. The policy places a codified responsibility on the Executive Staff to make deliberate decisions based on sound criteria to

protect the confidentiality of the respondents. It was created to coordinate the approval process among all parts of the bureau, to codify the procedures, and to ensure consistency in decisionmaking.

Non-employees need access to confidential data for a variety of reasons. Examples include contractors who have special skills, firms that can provide special equipment or facilities which would not be cost-effective for the Census Bureau to purchase or develop itself, sponsors who want to participate in the development and assessment of their surveys, and academic researchers who can provide important improvements to the Census Bureau's programs by using the data.

Congress foresaw this need and included safeguards in Title 13 Section 23(c), which requires that non-employees may have access to protected information only:

- If they are conducting work that is authorized by Title 13; and

- If they are sworn to the same non-disclosure oath that employees take to protect the data and are subject to the same penalties: a fine of up to $250,000 or a jail term of up to 5 years, or both, for unauthorized disclosure or use of protected information.

Upon taking that oath, these non-employees are given "Special Sworn Status" (SSS). They are also bound to abide by the Privacy Act of 1974 and OMB's Circular A-130, which require confidentiality and informed consent in data collection and related activities.

This policy establishes criteria and procedures for determining when it is appropriate to confer SSS on non-employees and for deciding when, if necessary, confidential data may be taken to a non-Census Bureau site or facility to carry out the work. For individuals to qualify for SSS, the project must:

- Require access to Census Bureau confidential data – in other words, publicly released data are not adequate;
- Benefit the Census Bureau's Title 13 programs as required by law; and
- Be a viable project within the time constraints and disclosure requirements of the Census Bureau and be consistent with the Census Bureau's Data Stewardship Principles.

In addition, individuals seeking SSS and their organizations must:

- Have a good track record for handling sensitive or confidential data;
- Have no identified conflict of interest in dealing with the Census Bureau; and
- Pass the background investigation for SSS candidates.

If a project and its associated people meet all of these criteria, SSS may be conferred on them.

In most cases, SSS individuals will access the Title 13 data at a Census Bureau facility, but, in a limited number of cases, the work is best done at a non-Census Bureau site. Then, additional safeguards are required. To begin with, only certain types of projects qualify for off-site consideration by the DSEP – that is, they must support the Census Bureau's core work or joint work with another statistical agency that requires off-site access. In addition, there must be:

- A technical or logistical advantage to doing the work off-site (including, but not always exclusively, cost);
- The ability to follow the Required Security Model for Off-site Access that is prescribed in the policy; and
- In the case of a governmental agency or organizational unit, legal or regulatory functional separation of the data collected for statistical purposes.

Functional separation means that information about an individual or entity that is collected for a research or statistical purpose may not be used in arriving at an administrative, enforcement, or other decision about that individual or entity.

The one exception to this rule is to permit off-site access for selected joint projects carried out at the Social Security Administration (SSA). Although the SSA's functional separation is in practice, not in law or regulation, this exception is made in view of the SSA's more than 30-year history of protecting Title 13 data and the integral role that SSA data play in ongoing Title 13 programs.

Finally, the DSEP must approve each off-site request.

This policy ensures that the Census Bureau's decisonmaking about non-employee access to Title 13 data is carried out in a consistent manner, with built-in attention to rigorous and ethical precepts, on issues such as the track record and potential conflicts of interest, as well as all legal requirements. In doing so, it facilitates the goals of allowing maximum access to the information embedded in Census Bureau data, while scrupulously protecting the privacy of individuals and organizations, and the confidentiality of their responses.

CONCLUSION

The growing complexity of our country and its economy creates increasing demands for data. At the same time, technological advances make it possible to collect and analyze data in more efficient ways, but they also require greater vigilance to ensure that confidential information remains protected.

The Census Bureau will continue to maintain its state-of-the art physical and communications security measures and disclosure avoidance techniques.

On the policy and procedures level, the Census Bureau has created an ongoing process to anticipate emerging threats to confidentiality and to ensure consistent responses within and among its program areas. The DSEP focuses the attention of Census Bureau executives on policy issues that need to be resolved, codified, and applied consistently throughout the Census Bureau. The Policy Office has been enlisted to monitor and examine the implementation of these decisions.

The demand for accurate and timely information will continue to grow because efficacy and efficiency in governing, business, academia, and throughout society require it. Through its continuous emphasis on protecting privacy and developing enhanced disclosure avoidance mechanisms, the Census Bureau will maintain its ethical commitment to protect the privacy of respondents, while collecting and disseminating the highest quality data about the people and economy of the United States.

# # #