

Official Statistics and Statistical Ethics: Selected Issues

William Seltzer

Fordham University, Department of Sociology and Anthropology

441 East Fordham Road

Bronx, NY 10438, USA

seltzer@fordham.edu

This paper addresses three questions related to statistical ethics and official statistics: (1) Do ethics have a role in official statistics? (2) What are the main ethical challenges that arise in official statistics? and (3) How may one deal with the ethical problems that arise in official statistics?

I base my consideration of these questions on my experience in national and international statistical work, the results of research into the misuse of population data systems and other threats to the integrity of official statistics that I and others have carried out, and insights gained through my nearly six years as chair of the American Statistical Association's Committee on Professional Ethics. However, the views expressed here are my own. Because of limitations of space, this paper primarily cites specific examples from the experience of official statistics in the United States. Nevertheless, the ethical challenges discussed have general applicability.

1. Do ethics have a role in official statistics?

It is widely recognized that government (that is, official) statistical agencies need to carry out their functions by responding to the needs of data users and more general government priorities, guided by law and science, and with due regard to the safety and good will of their respondents (that is, the data providers), all within allowed budgetary and administrative constraints. In these pressing circumstances, the explicit consideration of ethical issues is often reflexively diverted into a consideration of whether or not a given action is legal or in conformity with permitted administrative arrangements, or simply in accord with the policies and priorities of the current government. In addition, because government statistical personnel also properly perceive themselves as engaged in scientific work, they sometimes consider that the beneficence of their scientific work, particularly when apparently legal, excuses further ethical examination.

Such practices and views, however, are short-sighted and in the extreme contrary to the laws of many countries as well as international law (for example, national and international laws relating to genocide and crimes against humanity). It may be recalled that among the defenses offered by those charged with medical experiments, and the related collection of information and anthropological materials, in the so-called doctors' trial at Nuremberg after World War II were that (a) their actions were legal under current national law and (b) they were engaged in scientific work for an important and beneficent purpose. These defenses were explicitly rejected by the Nuremberg Tribunal when it found the defendants guilty of crimes against humanity. Subsequently many national and international ethical statements, research policies and regulations as well as related laws have adopted the same perspective.

Ethical shortcomings in official statistics have also at times been associated with very dire events, including genocide (see for example, Seltzer (1998) and Seltzer and Anderson (2001)). However, more commonly, unaddressed ethical issues simply threaten the credibility of a statistical agency, undermining the trust that data providers, the public, and policy makers place in the agency or government statistical work generally. Such a backlash against a government statistical agency may be immediate or long-delayed.

In any event, neither the governmental or scientific character of official statistics absolves statistical office staff and management from their individual and corporate ethical responsibilities. In seeking to carry out these ethical responsibilities, a variety of ethical codes, standards, and guidelines adopted by national statistical societies and the ISI are available (see, for example, ASA (1999) or ISI (1986)). In addition, the United Nations Fundamental Principles of Official Statistics (United Nations, 1994) was developed to explicitly address the institutional responsibilities of governments and their statistical agencies. However, as useful as this document is, it does not address the individual responsibilities of staff and management in these agencies.

2. What are the main ethical challenges that arise in official statistics?

(a) Use of sound methodology

Although technically sound science is not necessarily ethical science, the failure to use sound technical methods can be so flagrant or long-standing as to present serious ethical issues. In official statistics, this sort of situation may arise when a seriously flawed or outdated methodology continues to be used by an agency long after its shortcomings have been identified and alternative approaches explored, but for inappropriate reasons (for example, political pressures or a false sense of institutional pride), the flawed approach remains in use, sometimes for decades. Recently, Powers (2002) characterized the continued use by the U.S. Census Bureau of a long-outdated methodology for estimating poverty thresholds as an example of this type of ethical challenge.

It should be understood that not every error or flawed procedure is an indicator of an ethical lapse. Official statistics is carried out in a world of deadlines and limited resources. Also, as in any scientific endeavor, it uses imperfect methods and data. Indeed, a sound statistical system benefits from studying its mistakes, examining ways of overcoming them, and over time introducing needed improvements. But in the end there is both an ethical and scientific imperative to make the needed improvements. Not to do so, at some point, can become both a scientific and ethical failure.

(b) Protection of confidentiality

The roots of the concept of statistical confidentiality and the protection from harm attributable to cooperating with statistical inquiries can be traced back to the Hippocratic oath where physicians agree not to cause harm to their patients and not to gossip about information obtained in the course of their professional work. The modern concept of statistical confidentiality, at least in the United States, evolved in the 1890s as a means of encouraging businesses to report accurately by assuring them that business rivals, muckraking journalists and populist members of Congress would not have access to the information they provided, except as statistical aggregates.

These assurances were first extended to population data in the proclamation issued by President Taft in connection with the 1910 Decennial Census, using the following language, "The census has nothing to do with taxation, with army or jury service...or with the enforcement of any national, State, or local law or ordinance, nor can any person be harmed in any way by furnishing the information required." (Barabba, 1975: 27)

Effective legal protections against disclosure were first introduced in U.S. laws relating to data collected by the US Census Bureau concerning businesses. By the time of the 1920 and 1930 Censuses effective protections were extended to cover population data obtained by the Census Bureau as well. Despite periodic set backs, these protections have by now been extended to cover all data collected solely for statistical purposes by most U.S. federal statistical agencies (see, for example, Seltzer and Anderson (2003)). However, legal protections relating to statistical confidentiality in the United States only pertain to identifiable microdata. Unfortunately, both in the United States and elsewhere direct harms have arisen through the use of mesodata (tabulated data for very small geographic units, such as blocks or enumeration districts) to operationally assist in targeting vulnerable population subgroups for internment or worse. Given this experience, and in

the absence of clear legal protections relating to mesodata, statistical agencies, together with their leadership and staff, are under heavy ethical obligations to provide as wide a protective net as possible over mesodata pertaining to such vulnerable populations.

c. Integrity of the statistical agencies and the national statistical system

An important focus of the Fundamental Principles of Official Statistics is the maintenance and enhancement of the integrity of the national statistical system. As discussed in an earlier study (Seltzer, 1994) threats to integrity can arise in a number of ways, including, among others, arbitrary political manipulation of concepts, definitions, and the extent and timing of the release of data, doctoring the actual data released, using the agency for political analysis or other political work, and politicizing agency technical staff. Certainly, the public at large, data users, and political leaders in and out of the government have a long-term interest in resisting such threats. For agency leadership and staff, individual statisticians and national statistical societies, the IAOS, and the ISI that interest in resisting such threats becomes an ethical responsibility.

3. How may one deal with the ethical problems that arise in official statistics?

During the course of work in official statistics one may encounter or anticipate a range of ethical challenges. The issue then arises over what to do. I shall attempt to briefly answer this question by outlining a series of coping and prevention strategies.

(a) Coping strategies

A number of options are available in dealing with what one perceives as an ethical problem in government statistical work. The basic elements of an initial response are: speaking up about the perceived problem, consulting with other colleagues informally, establishing a written record, and explicitly informing one's supervisor about these concerns. Additional avenues of recourse may be available through an agency ombudsman, where one exists, or through the mechanism of legislative oversight. When documenting these concerns in writing it is useful to indicate how the proposed action or existing practice violates established professional, agency, or government-wide norms for statistical work. In thinking about what to do at any stage, it is important to remember the principle of proportionality of response to threat or harm. Minor harms do not usually justify major action and we should also avoid escalating differences over scientific methods into an ethical controversy.

Further options for action are also possible. The next step beyond those just outlined is to "go public" with one's concerns. The act of "going public" with an ethical concern may raise legal and personal issues and it may be wise to obtain legal and other advice. It also may involve an ethical choice of obligations to one's own agency versus other obligations. Some of the ethical issues involved are addressed in the ASA guidelines in the following terms

In cases of conflict, statistical practitioners and those employing them are encouraged to resolve issues of ethical practice privately. If private resolution is not possible, recognize that statistical practitioners have an ethical obligation to expose incompetent or corrupt practice before it can cause avoidable harm to research subjects or society at large. [ASA, 1999: 10]

Going public may take several forms: verbal or written statements made at professional conferences or elsewhere, bringing the matter to the attention of the press, resignation, or non-violent civil disobedience. All of these avenues have been used by those working in the U.S. federal statistical system in the past (Seltzer, 2001).

(b) Prevention strategies

A robust set of prevention strategies is perhaps the best means of avoiding serious ethical problems and of ensuring that any that do arise can be dealt with responsibly and expeditiously.

Possible preventative actions include: (1) studying and documenting previous problems; (2) developing and disseminating case studies that illustrate ways of addressing ethical issues based on real or hypothetical examples; (3) developing enhanced models of disclosure risk that take into account both the probabilities of disclosure and the possible harms that may arise from such disclosures; (4) providing education and training on ethics in university and statistical agency training programs, including those specifically tailored for mathematical statisticians and computer methods staff who often have central responsibility for work on disclosure safeguards; (5) developing agency-specific plans for fostering discussions of agency ethical issues and agency-specific mechanisms for responding to ethical concerns; (6) developing, with external input and public comment, statements articulating the ethical standards that agency staff and management are expected to follow; and (7) further developing and applying a range of substantive, methodological, operational, and legal safeguards. Statistical agencies that are actively pursuing such preventive strategies will also have a management and staff alert to the ethical issues that arise in their work. Such a sense of ethical alertness is perhaps the single best defense against major ethical tragedies.

REFERENCES

ASA [American Statistical Association]. 1999. "Ethical Guidelines for Statistical Practice." Also available at <<http://www.amstat.org/profession/index.cfm?fuseaction=ethicalstatistics>>.

Barabba, Vincent. 1975. "The Right of Privacy and the Need to Know." In U.S. Census Bureau, *The Census Bureau: A Numerator and Denominator for Measuring Change*, Technical Paper 37. Washington, D.C.: US Government Printing Office.

International Statistical Institute. 1986. "Declaration of professional ethics for statisticians." *International Statistical Review* 227-247. Also available at <<http://www.cbs.nl/isi/ethics.htm>>.

Powers, Mary. 2002. "Concepts, Definitions, and Classifications: Some Ethical Issues in Demographic Research and Policy." Paper presented at the PAA Annual Meeting, May 2002.

Seltzer, William. 1994. "Politics and Statistics: Independence, Dependence, or Interaction?" Working Paper No. 6, U N Dept. of Econ. and Soc. Information and Policy Analysis: New York.

_____. 1998. "Population Statistics, the Holocaust, and the Nuremberg Trials." *Population and Development Review*, 24 (3): 511-552.

_____. 2001. "U.S. Federal Statistics and Statistical Ethics: The Role of the American Statistical Association's Ethical Guidelines for Statistical Practice." Based on a presentation to the Washington Statistical Society, February 2001. Available at <<http://www.uwm.edu/~margo/govstat/integrity.htm>>.

Seltzer, William and Margo Anderson. 2001. "The Dark Side of Numbers: The Role of Population Data Systems in Human Rights Abuses." *Social Research* 68 (2): 481-513.

_____. 2003. "Government Statistics and Individual Safety: Revisiting the Historical Record of Disclosure, Harm, and Risk." (with Margo Anderson). Paper prepared for presentation at a workshop, organized by the Panel on Confidential Data Access for Research Purposes, Committee on National Statistics, The National Academies, Washington, DC, October 16-17, 2003. Available at <http://www7.nationalacademies.org/cnstat/Seltzer_Anderson.pdf>.

United Nations Economic and Social Council. 1994. *Report of the Special Session of the Statistical Commission, New York, April 11-15*. E/1994/29. Also available at <<http://unstats.un.org/unsd/methods/statorg/FP-English.htm>>.

RESUMÉ

Cet article adresse trois questions au sujet de l'éthique et des statistiques officielles: (1) l'éthique a-t-elle un rôle dans des statistiques officielles? (2) quels sont les défis moraux principaux qui se produisent dans des statistiques officielles? et (3) comment on peut-il traiter les problèmes moraux qui se produisent dans des statistiques officielles?