

Preliminary unreviewed draft for comment.
10/3/03; minor corrections 10/14/03
Not for quotation.

**Government Statistics and Individual Safety: Revisiting the Historical Record of
Disclosure, Harm, and Risk**

by

WILLIAM SELTZER
(Fordham University)

and

MARGO ANDERSON
(University of Wisconsin Milwaukee)

Contact Information

William Seltzer
Department of Sociology and Anthropology
Dealy 407
Fordham University
441 East Fordham Road
Bronx, NY 10458
email: seltzer@fordham.edu

Margo Anderson
Department of History
University of Wisconsin - Milwaukee
Milwaukee, WI 53201
email: margo@uwm.edu

The purpose of this paper is to revisit the issue of the disclosure of information about individuals and families collected by statistical agencies and programs under government auspices. We also assemble in one place the results of several recent studies of such disclosures, or of actions that might lead to such disclosures, actions placing respondents and other members of the public at added risk of death, prosecution, or the loss of liberty. In the broadest terms, the paper has two interrelated objectives: (1) the presentation of a body of facts and (2) the presentation of a reconceptualization of a number of the issues related to disclosure and statistical confidentiality needed to understand the implications of the facts assembled. A reconceptualization that is rooted in the ethical, statistical policy, and statutory origins of the idea of statistical confidentiality. In our view both parts are essential. Indeed, as we will illustrate, in the absence of this reconceptualization many of the facts themselves remained invisible for decades. Accordingly, any sound discussion of disclosure must be based on a clear understanding of the history of the concept of statistical confidentiality as it has evolved in its ethical, policy, and legal dimensions. Such a discussion must also address issues related to the concepts of harm and risk. While risk is a term frequently used in discussions of disclosure, it is most often interpreted to mean the risk of disclosure. We emphasize in this paper two quite different kinds of risk: risks to data providers occasioned by different sorts of disclosures and risks to government statistical agencies posed by such disclosures. Each of these risks entails some consideration of the harm that individuals, vulnerable population subgroups, and statistical agencies may be exposed to through such disclosures.

As the title of this paper indicates, our focus is on one aspect of a complex set of issues: the interface between the actions of government statistical programs, their staff, and the agencies in which they are located, on the one hand, and the extent to which individuals and vulnerable population subgroups may be threatened or protected by such actions, on the other hand. One's understanding of the requirements of statistical confidentiality and the safeguards needed to prevent disclosure in turn define how to protect respondents from harm. The threats we refer to relate to the personal safety of such individuals and groups from governments themselves, and, in the extreme, have placed people at increased risk of major human rights abuses. Because of their potential seriousness, we believe that such risks deserve special attention, just as potentially catastrophic events in other fields are studied so that they may be avoided or otherwise protected against.

As the title of this paper indicates, our focus is on one aspect of a complex set of issues: the interface between the actions of government statistical programs, their staff, and the agencies in which they are located, on the one hand, and the extent to which individuals and vulnerable population subgroups may be threatened or protected by such actions, on the other hand. One's understanding of the requirements of statistical confidentiality and the safeguards needed to prevent disclosure in turn define how to protect respondents from harm. The threats we refer to relate to the personal safety of such individuals and groups from governments themselves, and, in the extreme, have placed people at increased risk of major human rights abuses. Because of their potential seriousness, we believe that such risks deserve special attention, just as potentially catastrophic events in other fields are studied so that they may be avoided or otherwise protected against.

Our focus on the more serious, but hopefully rare, cases of possible disclosures related to a government's own misuse of the statistical system does not imply that we think disclosures arising from other sources should be ignored. In fact, many past studies of disclosure have concentrated instances of someone "finding themselves" or "finding someone else," in a file, perhaps by using special information known about an individual (for example, efforts to find out about the financial condition of a former spouse) or concatenating search criteria. Such cases of disclosure or disclosure risk are important to study and protect against both from the perspective of the reputation of a statistical agency and the privacy of the affected individuals. These disclosures, however, may be likened to "mom and pop" retail efforts. By contrast, the present paper addresses efforts that may be likened to state-sponsored wholesale enterprises.

It is sometimes suggested that, given the scope of potential private sector threats to privacy, activities of the state, particularly those involving the misuse of data gathered or held by statistical agencies, no longer pose any special threat to the individual. Such suggestions ignore the special status accorded various outputs of many federal statistical agencies by federal, state, and local courts and administrative entities in the United States.¹ Under long-standing rules of evidence, courts and administrators may take "judicial note" of the results obtained by statistical agencies, such as the Census Bureau (see, for example, *United States v. Esquivel*, 88 F.3d 722 (9th Cir.1996) in which the court stated that "census documents meet the requirements of Rule 201(b), Fed. R. Evid., in that they are "not subject to reasonable dispute" because they are "capable of accurate and ready determination by resort to sources whose accuracy cannot reasonably be questioned.""). In short, courts and government officials can and have taken decisions relating to life, liberty, rights, and property relying on the official status of the information gathered by federal statistical agencies.

In discussing issues related to data disclosure and harm to individuals, three different types of data are usefully distinguished: *macro* data, *meso* data, and *micro* data. Since the terms are used frequently in this paper, we will define them at the outset. *Macro* data refer to tabulated aggregates for national or large geographic areas, *meso* data to tabulated data for sufficiently small geographic areas that the results can be used operationally to identify and target a vulnerable population subgroup, and *micro* data to identifiable records for each individual. While the relative protection offered by the statistical aggregates of *macro* data and the relative vulnerability of individual records that constitute *micro* data have long been recognized, the special risks posed by *meso* data have only been explicitly examined in the past few years. Accordingly, a few more words about the nature and role of *meso* data may be useful.

As used here, *meso* data are statistical results presented at such a fine level of geographic disaggregation, whether in tabular or graphic form, that the results may be used in conducting field operations at the local level. Thus the borderline between *macro* data and *meso* data will depend in part on the size of the geographic units, the distribution of the target population among these units, and the intended operational uses. For example, census aggregates showing the number of persons in a target population for an individual small village maybe operationally useful, while similar data for a large city, would need to be further broken down by tract, ward, or even block to be operationally useful.

It should be noted that over the years there have at times been some differences in the degree of protection accorded to information obtained by statistical agencies about business enterprises compared to information relating to persons, at least in the United States. Although this paper concentrates almost exclusively on issues related to information about individuals, some reference is made to important interactions between developments related to individuals and enterprises. Even as we exclude from the paper any detailed consideration of the disclosure of information collected about enterprises and establishments, we acknowledge that this is an important area of research in its own right.

Finally, we would emphasize, as we have previously, that while addressing instances of the misuse or the diversion of government statistical systems, we do so as strong supporters these systems, their outputs, and their staffs. Government statistical systems, and the United States federal statistical system is a prime example, gather a broad range of valuable data for policy, administration, and research. We take that as a given. We believe, however, that only by also documenting and studying things that went wrong, or nearly went wrong, are we in a position avoid, minimize, or control for such mistakes in the future. This is a proposition that underlies all studies of error by data users and producers. The present study simply deals with a different dimension of error.

After an introductory section reviewing the history of the concept of statistical confidentiality (Section I), we discuss the related issues of disclosure, harm, and risk that have emerged from the research on the use of government statistical agencies or programs to assist in the nonstatistical task of targeting individuals or vulnerable population subgroups (Section II). Section III of the paper sets out the available evidence concerning such governmental efforts, some of which have led to serious human rights abuses. The section draws on previously reported research supplemented by more recent research by ourselves and others, some of it ongoing. It is divided into two subsections, the first dealing with examples from the United States (Section III.A) and the second with examples from other parts of the world (Section III.B). In both subsections we also identify some areas where additional research appears warranted. In Section IV we describe a number of barriers to the study of disclosures, harms, and risks associated with the activities of government statistical programs. We conclude by offering a few recommendations that appear to us to flow from the material presented in the paper (Section V).

I. Introduction: the concept of statistical confidentiality

The concept of the confidentiality of information about individuals obtained in the course of statistical investigations has foundations in ethics, statistical policy, and law. As with many aspects of official statistics, the modern concept, while over a century old, has evolved over time. This evolution has proceeded somewhat differently in terms of ethics, statistical policy, and law. Initially, these three strands evolved independently of one another. More recently they have interacted, but each retains its distinctive character. We shall address each in turn in so far as possible, given the interactions. With respect to statutory protections, we shall confine ourselves exclusively to developments in the United States.

(a) Ethics. The normative foundations of statistical confidentiality have roots deep in at least two ethical principles. First is the moral imperative that we must be alert to our capacity to cause harm in our eagerness to do good. Second is the caution that personal information provided in various professional contexts should not be shared with others. The rationale for the first is the recognition that in an aura of beneficence or zeal we are sometimes blind to the harm that our actions may cause. The rationale for the second is twofold: (1) the recognition that unless those we must interact with trust us, we cannot do our job, and (2) that in doing our job we are very likely to obtain information that can be used against those we must interact with. Well before statisticians considered such issues, those in the clergy, law, and medicine addressed them. For example, both these norms are reflected in the Hippocratic Oath that dates to back to about 400 BCE.² Of course, ethical issues as they present themselves in official statistics are also rooted in developments in the field itself as well as in developments related to experimentation and research involving human subjects.

One major development in this process was the so-called Nuremberg code, formulated by two expert witnesses during the course of one of the trials conducted at the conclusion of World War II. The trial dealt with some of those involved in carrying out the so-called medical experiments on concentration camp inmates.³ This Code, which was accorded the force of international humanitarian law by virtue of its adoption by the Court, stressed the absolute duty of those carrying out experiments (and research) on human subjects to avoid serious harm and to obtain voluntary informed consent from their research subjects. This Code, along with various revisions and extensions, also served as the basis for a number of normative, policy, and statutory efforts in the United States and abroad, to guide and regulate experiments and research involving human subjects during the second half of the twentieth century.

Nevertheless, initial efforts to develop a set of normative guidelines for statistics undertaken by the American Statistical Association [ASA] at the national level in the late 1940s and early 1950s concentrated primarily on issues of statistical standards, rather than non-technical matters, and confidentiality did not feature as one of the topics addressed [Clausen, et al., 1954]. The only reference to the issue seems to have been in a 1952 published comment listing among the provisions that might be included in a normative code for statisticians:

3. If source material is furnished him on condition that the respondent should not be specifically identified, he preserves this anonymity. [Freeman, 1952: 20]

The process by which the ASA reached a normative statement was very slow indeed, see Gibbons [1973] and Ellenberg [1983] for a review of the process. However, when it was finally approved and published the initial ASA Ethical Guidelines for Statistical Practice contained a clear and detailed statement concerning statistical confidentiality [Ireland et al., 1983: 5].

The Ethics Declaration adopted by the International Statistical Institute (ISI) in 1985 also contains extensive guidance on issues related to obligations to human research subjects [International Statistical Institute, 1985, Part two, section 4], particularly section 4.4, “Protecting the interests of subjects,” which states

Neither consent from subjects nor the legal requirement to participate absolves the statistician from an obligation to protect the subject as far as possible against potentially harmful effects of participating. The statistician should try to minimize disturbance both to subjects themselves and to the subjects' relationships with their environment.

section 4.5, “Maintaining confidentiality of records,” which states

Statistical data are unconcerned with individual identities. They are collected to answer questions such as 'how many?' or 'what proportion?', not 'who?'. The identities and records of co-operating (or non-cooperating) subjects should therefore be kept confidential, whether or not confidentiality has been explicitly pledged.

and section 4.6, “Inhibiting disclosure of identities,” which states

Statisticians should take appropriate measures to prevent their data from being published or otherwise released in a form that would allow any subject's identity to be disclosed or inferred.

Building on these earlier efforts, and informed by the research on the misuse of national population data systems to assist in some major human rights abuses, the most recent version of the ASA's ethical guidelines for statistical practice, approved in 1999, addressed the subject in several ways [American Statistical Association, 1999]. Thus, in a section of the preamble dealing with “shared values,” the guidelines state

5. Adherence to all applicable laws and regulations, as well as applicable international covenants, while also seeking to change any of those that are ethically inappropriate.

6. Preservation of data archives in a manner consistent with responsible protection of the safety and confidentiality of any human beings and organizations involved. [1999:4]

Within the main body of the ASA guidelines, Section D, “Responsibilities to Research Subjects (including census or survey respondents and persons and organizations supplying data from administrative records, as well as subjects of physically or psychologically invasive research),” contains eight provisions. Among them, three seem to be particularly relevant to the issues of confidentiality, disclosure, and harm associated with government statistical programs:

1. Know about and adhere to appropriate rules for the protection of human subjects, including particularly vulnerable or other special populations who may be subject to special risks or who may not be fully able to protect their own interests. Assure adequate planning to support the practical value of the research, the validity of expected results, the ability to provide the protection promised, and

consideration of all other ethical issues involved. Some pertinent guidance is provided . . . at the end of this document for U.S. law, the U.N. Statistical Commission, and the International Statistical Institute. Laws of other countries and their subdivisions and ethical principles of other professional organizations may provide other guidance.

4. Protect the privacy and confidentiality of research subjects and data concerning them, whether obtained directly from the subjects, from other persons, or from administrative records. Anticipate secondary and indirect uses of the data when obtaining approvals from research subjects; obtain approvals appropriate for peer review and for independent replication of analyses.

5. Be aware of legal limitations on privacy and confidentiality assurances. Do not, for example, imply protection of privacy and confidentiality from legal processes of discovery unless explicitly authorized to do so. [1999: 7-8]

(b) Statistical Policy. The idea of the confidentiality of individual level statistical information first arose as a matter of statistical policy in the United States in the latter part of the 19th Century in connection with the population census. Initially, population census returns were posted in public places for the local citizenry to review to make sure they were complete and accurate. By the middle of the 19th century, Congress and census officials began to be concerned about such public posting, and the practice was discontinued. Instead census officials concerned with the integrity of the enumeration became increasingly worried about an enumerator revealing information improperly, or making use of census information for private benefit. As a result, those responsible for conducting the 1850 Census were cautioned against such abuses by the Secretary of the Interior [Bohme and Pemberton, 1991: 3].

By the time of the Ninth Census in 1870, Census Superintendent Francis A. Walker instructed the assistant marshals responsible for the enumeration that

No graver offense can be committed by assistant marshals than to divulge information acquired in the discharge of their duty. All [information obtained] should be treated as strictly confidential ... The [Department of Interior] is determined to protect the citizen in all his rights in the present census. [quoted in Bohme and Pemberton, 1991: 4]

The value of offering respondents the assurance that their responses in statistical inquiries would not be used against them by the Government seems to have first been articulated in explicit terms by the first Commissioner of Labor (subsequently the Commissioner of Labor Statistics), Carroll Wright. From the outset, he seemed to be cognizant of the need to maintain the confidentiality of responses from individuals and firms. At the time Wright's investigations into labor conditions were exceedingly controversial. He relied on voluntary responses. He knew he would not receive cooperation from respondents, or confidence in his analyses, without objective and complete data. He made it a point to reassure respondents, for example, telegraphing a San Francisco businessman in 1898 that "I pledge my word as a government officer that names of

your plants and of city and State in which located shall be concealed. This will be done for all plants. If senator or representative should ask for these names, he should not have them” [quoted in Goldberg and Moye, 1985: 13].

Along the same lines, Herman Byer, Assistant Commissioner of Labor Statistics in the late 1940s described how a subsequent Commissioner of Labor Statistics, Ethelbert Stewart, was said to have responded to congressional pressure in the 1920s to reveal identifiable *micro* data. According to Byers, Stewart was asked at a Congressional hearing to reveal the data on individual automobile manufacturers to Congress, and he refused on grounds of confidentiality. When the committee chair threatened Stewart, “Mr. Stewart, our committee will subpoena those records,” Stewart responded, “You do, and I’ll burn them first” [quoted in Duncan and Shelton, 1978: 168]. Like Walker, both Wright and Stewart grounded their insistence on statistical confidentiality on the requirements of statistical policy, rather than on any specific statute.

The same policy was clearly reflected in the census proclamation issued by President Taft in 1910 in connection with the Thirteenth Census,

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; quoted in full in Bohme and Pemberton, 1991: 6].

With few notable exceptions, this view has also been reflected in statistical policy in the United States since then and in virtually all open societies since the end of World War II. For example, writing in 1967 about the U.S. federal statistical system, Edgar Dunn Jr was emphatic,

legal and procedural protections against revealing information about individuals have been a very basic part of the operation of the Federal statistical programs for many many years! ... The protection of personal privacy has long been an obsession with the directors of federal statistical programs ... [Dunn, 1967: 23]

Although in retrospect, he overstated the situation somewhat, his statement certainly conveyed the prevailing ethos.

The most notable exceptions in the United States to Dunn’s optimistic assessment were: (1) the successful effort in May 1917 by Census Director Samuel Rogers to gain permission to use identifiable *micro* data from the 1910 census to assist in the investigation and prosecution those suspected of violating World War I draft registration laws (see section III.A.4 below); (2) the decision in 1939 by President Roosevelt to permit the FBI and the Army and Navy intelligence services to have access to identifiable census *micro* data on an individual, case by case basis (see Seltzer and Anderson [2002: 39-40]); (3) the activities of Census Director J.C. Capt in 1941 and 1942 to make identifiable *micro* data in the hands of the Census Bureau available to assist in the war effort (see section II.c below); and (4) the view held shortly after September 11, 2001 by at least some in National for Education Statistics (NCES) and the Department of Education that identifiable NCES *micro* data provided an effective and appropriate means to locate or track

suspected terrorists (see Seltzer and Anderson [2002]).

Despite these detours, current statements of federal and international statistical policy with respect to statistical confidentiality remain unaltered. Thus the Federal Statistical Confidentiality Order issued by the Office of Management and Budget in 1997, the Fundamental Principles of Official Statistics adopted by the United Nations Statistical Commission in 1994, and the Committee on National Statistics in its 2001 publication, *Principles and Practices for a Federal Statistical Agency* express a similar policy view to the effect that statistical confidentiality means that the information so obtained must be used for exclusively statistical purposes. The longer OMB and CNSTAT texts also spell out that such information should not be used by the government against those to whom the information pertains. In the words of the OMB order

Consistent government policy protecting the privacy and confidentiality interests of persons who provide information for Federal statistical programs serves both the interests of the public and the needs of the government and society. The integrity and credibility of confidentiality pledges provides assurance to the public that information about persons or provided by persons for exclusively statistical purposes will be held in confidence and will not be used against them in any government action. Public confidence and willingness to cooperate in statistical programs substantially affects both the accuracy and completeness of statistical information and the efficiency of statistical programs. Fair information practices and functional separation of purely statistical activities from other government activities are both essential to continued public cooperation in statistical programs. [OMB, 1997: 35047]

The relevant United Nations Statistical Commission policy states

Individual data collected by statistical agencies for statistical compilation ... are to be strictly confidential and used exclusively for statistical purposes [United Nations Economic and Social Council, 1994: Principle 6].

Finally, in the words of the CNSTAT report,

Data providers must be able to rely on the word of a statistical agency. An agency receives credibility with its providers by ensuring appropriate confidentiality of responses. Maintaining confidentiality, in particular, precludes the use of individually identifiable information for any administrative, regulatory, or enforcement purpose. [National Research Council, 2001: 5]

(c) Statutory Protections. With respect to the census, and probably to statistical operations generally, the first statutory reference to confidentiality in the United States appeared in the Tenth Census Act, adopted in 1879 and required that the enumerator to swear under penalty that "... [I] will not disclose any information contained in the schedules, lists, or statements except to my superior officers" [Holley, 1938: 1], although as Bohme and Pemberton [1991: 5] point out the penalty clause of this statute pertained only to "statistics of property or business."⁴ This

anomaly was corrected in the act adopted in 1889 for the next census [Bohme and Pemberton, 1991: 6].

The acts adopted in 1899 for the Twelfth Census and in 1909 for the Thirteenth Census left the statutory concept of census confidentiality essentially unchanged [Holley, 1938: 2-4]. That is, these census acts were designed to protect the Government from unauthorized disclosure of census *micro* data by census staff, not to safeguard individuals from the Government's use of such data to the detriment of the respondent. (Indeed the acts adopted for the 1900 and 1910 Censuses specifically authorized the Census Director to provide to governors or courts copies of individual census returns.⁵) Nevertheless, the modern concept of census confidentiality appeared to make a major advancement in 1910 based on the language used by President Taft in the first presidential census proclamation quoted above in connection with the review of statistical policy. Despite the broad blanket of protection seemingly provided by President Taft's official assurance that no "person [can] be harmed in any way by furnishing the information required," as more fully discussed in section III.4 below, seven years later the Census Director was advised that nothing in the Census Act itself prevented him from providing information from the 1910 Census schedules pertaining to the names and ages of suspected draft evaders to registration authorities.

The Fourteenth Census Act, adopted in 1919, marked a key advance transforming the policy visions of Francis Walker, Carroll Wright, and President Taft's census proclamation into statutory language. For the first time the census statute provided that "... in no case shall information furnished under the authority of this act be used to the detriment of the person or persons to whom such information relates ... [Holley, 1938: 4]."

Title 13 was first enacted in 1929 as part of the Fifteenth Census Act. This act repealed all prior acts. It also carried forward the modern concept of statistical confidentiality, building on and further elaborating the statutory language of the previous act [Mitchell, 1930: 2-3]. Section 11 of the 1929 Act stated,

That the information furnished under the provisions of this Act shall be used only for the statistical purposes for which it is supplied. No publication shall be made by the Census Office whereby the data furnished by any particular establishment or individual can be identified, nor shall the Director of the Census permit anyone other than the sworn employees of the Census Office to examine the individual reports. [Public Law 13, 71st Congress, June 18, 1929]

Similarly, Section 18 of Title 13, authorized the Census Director, to provide individuals with their own returns, undertake special tabulations, and

at his discretion, upon the written request of the governor of any State or Territory or of a court of record, to furnish such governor or court of record with certified copies of so much of the population or agricultural returns as may be requested....*Provided, however,* that in no case shall information furnished under the authority of this Act be used to the detriment of the person or persons to whom such information relates.

In 1930, Census Director Steuart wrote to the Secretary of Commerce indicating it was his view that providing unprotected *micro* from the 1930 and subsequent census would both violate the new law and be “inadvisable as a matter of policy” and requested legal advice on the matter in light of some pending requests from from other federal and local agencies [Steuart, 1930]. The matter was referred to the Justice Department, and the Attorney General confirmed that it was the “duty of the Director of the Census to decline” to provide such unprotected *micro* data [Mitchell, 1930].

Thus, by the advent of the Roosevelt administration in 1933, statistical confidentiality was embedded in the practice of major statistical agencies such the Census Bureau and the Bureau of Labor Statistics, and with respect to data collected by the Census Bureau under the provisions of Title 13, in the law as well.

Nevertheless, as war broke out in Europe in fall of 1939, the Roosevelt administration found the constraints of Title 13 burdensome and sought to amend it to support national defense. As recounted elsewhere [Seltzer and Anderson, 2002: 37-38], the Departments of Justice, War, and Navy attempted to amend section 11 of Title 13 by inserting the following language

Provided, however, that the records of the Bureau of the Census, including the individual reports, shall be available to the Federal Bureau of Investigation of the Department of Justice, Office of Naval Intelligence of the Department of the Navy, and the Intelligence Division of the Department of War in connection with violations of the laws against espionage and other matters relating to the national defense whenever, in the opinion of the Attorney General, the Secretary of War or the Secretary of the Navy, the public welfare would be served by according such access to said records. [FDR Library, President’s Official File: 3b-3c, Department of Commerce, Box 6, Folder: Commerce Department, 1939-1940, Census Bureau]

The Census Bureau, under the leadership of Director William Lane Austin, and with strong support from the Department of Labor was eventually able to convince the President, in the weeks before the 1940 Census was to go into the field, that the proposed legislation not be introduced [FDR Library, Presidents Official File (POF) 3b-3c: Folder: “Commerce Dept., 1939-40 Census Bureau”].

While Roosevelt was convinced that a legislative battle over Title 13 during a census and an election year was politically dangerous, there were strong signals that he was not going to let the issue die. In 1941, secure in their win from the 1940 election, New Dealers again looked at strengthening national defense and again looked to gain access to the individual level information in census data. William Austin, Census Director who led the battle against the 1939 legislation, was involuntarily retired at the end of January 1941.⁶

Ignoring the universal recommendation from the statistical, social science, and user communities, that a person with professional knowledge be appointed as Austin’s replacement,

the nomination of J.C. Capt as Director was announced on April 22 and confirmed by the Senate, without debate, on May 13.⁷ Capt took office on May 22 and, almost immediately, on June 6, 1941, he initiated and obtained the support of Commerce Secretary Jesse Jones for legislation to eliminate the 1941 Census of Manufactures, to provide authority for periodic surveys for national defense needs, and to make individual level census reports available for use in the “national defense program.” Section 3 of the S1627 provided:

That notwithstanding any other provision of law, any individual census report or any information contained therein may be used in connection with the national defense program under such rules and regulations as may be prescribed, with the approval of the President, by the Secretary of Commerce. No person shall disclose or make use of any individual census report or any information contained therein contrary to such rules and regulations; and anyone violating this provision shall be guilty of a misdemeanor and upon conviction thereof shall be fined not exceeding \$500 or be imprisoned not exceeding six months or both.
[Congressional Record (77th Cong., 1st Session), volume 87, pt. 6, p. 6969, August 11, 1941]

The Senate report accompanying the bill was explicit about the goals of the legislation:

The needs of the defense program are of such a character as to require full and direct information about specific individuals and business establishments. It is clearly the intent of Congress and of the administration to implement in every possible way the defense program. An essential part of this implementation must be through the proper use of statistical data to speed production and to provide the detailed knowledge needed for the planning of total defense. To continue to impose the rigid provisions of the present confidential use law of the Census Bureau on data now in the possession of the Bureau and that to be gathered and used for national defense would defeat the primary objects of the legislation here proposed. As a safeguard against the possible misuse of information submitted in individual census reports, the Secretary of Commerce is given the authority to prescribe with the approval of the President the terms and conditions for use in the defense program only of such information. A penalty is provided for the disclosure of any individual census report or any information contained therein contrary to the terms and conditions prescribed by the Secretary. [77th Congress 1st session, Senate Rept 495, June 26, 1941, to accompany S 1627]

The bill passed the Senate in August 1941 and went to the House, where it remained until the attack on Pearl Harbor brought the United States into the war.

Four days later, on December 11, 1941, Census Director Capt reminded the Commerce Secretary that the Census Bureau was still constrained by the confidentiality requirements of Title 13, and proposed to write the language of S1627 into an executive order to get around the ban. As Capt wrote,

[T]he Bureau of the Census has no authority at the present time to permit other governmental agencies to obtain from Census records information about individuals or business establishments that may be indispensable to the defense of the nation. Authority therefore is needed for the Bureau of the Census to make available for war purposes any record of information in the possession of the Bureau of the Census when directed to do so by the Secretary of Commerce.

In my judgment, it is necessary to have these powers vested in the Secretary of Commerce at once to make possible the flexible, efficient, and economical war-time operation of the Bureau of the Census in obtaining and making available statistics for planning and directing war efforts. [NARA, RG40, General Records of the Department of Commerce, Office of the General Council, General Counsel's Subject and Index File, 1903-1946, Box 152, File 5706 -33. Memorandum from J.C. Capt to the Secretary of Commerce, 12/11/41]

This time, it was the Justice Department, now under a new Attorney General, that objected to the proposal to void the confidentiality provisions of Title 13. The Attorney General's Office decided that there was no legal authority for such an Executive Order [NARA, RG40, General Records of the Department of Commerce, Office of the General Council, General Counsel's Subject and Index File, 1903-1946, Box 152, File 5706 -33. Memorandum to the file, E.T. Quigley, 12/23/41].

Instead, the matter was handled legislatively and incorporated into section 1402 of the Second War Powers Act enacted on March 27, 1942, which in part read,

That notwithstanding any other provision of law, any record, schedule, report, or return, or any information or data contained therein, now or hereafter in the possession of the Department of Commerce, or any bureau or division thereof, may be made available by the Secretary of Commerce to any branch or agency of the Government, the head of which shall have made written request therefor for use in connection with the conduct of the war... [U.S. Code Congressional Service, 1943, P.L. 507, 77th Congress, 2d Session (S2208)]

Although there were slight differences of wording between Capt's initial legislative effort of the previous year and section 1402, with the former explicitly referring to "any individual report," section 1402 effectively implemented Capt's and the administration's proposals. Section 1402 of Title XIV of the Second War Powers Act was repealed as part of the First Decontrol Act of 1947 (S931), Chapter 29, Public Law 29, passed March 31, 1947, (50 U.S.C.A. Appendix, Section 644a).⁸

During the last 50 years of the 20th Century there was a general tendency to strengthen the statutory protections governing statistical confidentiality and extend them to a broader range of statistical programs. For example, such newly-established agencies as the Bureau of Justice Statistics, the National Center for Health Statistics and the National Center for Education Statistics had, like the Census Bureau, strong agency-specific statistical laws prohibiting

information about individuals gathered for statistical purposes to be used against them. Furthermore, even where courts found the statutory protections inadequate, Congress moved quickly to remedy the situation [Corcoran, 1963].

The past three years have seen two important statutory developments with respect to statistical confidentiality – the first a major backward step, the second a strong legislative reaffirmation of the principle of statistical confidentiality. As discussed below (see section III.3) and described in detail in Seltzer and Anderson [2002], in late October 2001, in response to the terrorist attacks on the United States of September 11, the USA Patriot Act became law. Title V of that multifaceted legislation was labeled “Removing Obstacles to Investigating Terrorism” and included section 508, “Disclosure of Information from NCES Surveys.” This section allowed the Justice Department to obtain and use “reports, records, and information (including individually identifiable information)” relevant to an authorized investigation or prosecution of domestic or international terrorism or terrorism that transcends international boundaries in the possession of the National Center for Education Statistics (NCES), notwithstanding the strong confidentiality provisions of the 1994 National Center for Education Statistics Act. As part of a major over-haul of federal education research, the NCES Act of 1994 was effectively repealed in late 2002, but the confidentiality protections of that Act, as amended by the Patriot Act, were made part of the Education Sciences Reform Act of 2002 (107th Congress, Public Law 279, November 5, 2002), the new legislation governing education research, including NCES.⁹

Even as Congress approved section 508 of the Patriot Act, relaxing the confidentiality protections for education statistics, virtually without debate, in 2001, and the recodification of this relaxation in November 2002, again apparently without debate, Congress also enacted in late 2002 the first government-wide law specifically protecting statistical confidentiality. This legislation, “Title V - Confidential Information Protection and Statistical Efficiency” of the E-Government Act of 2002 (107th Congress, Public Law 347, December 17, 2002), is a clear and unambiguous statement of policy and law.¹⁰ Because this law is both recent and important, we describe it in some detail.

Subtitle A of Title V, “Confidential Information Protection,” consists of three sections, section 511, which contains a statement of findings and purpose, section 512, which spells out the limitations on use and disclosure of data and information under the act, and section 513, which sets out the fines and penalties for the disclosure of such confidential data and information.¹¹ Four of the provisions of section 511 are particularly relevant to the present paper:

The Congress finds the following:

...

(2) Pledges of confidentiality by agencies provide assurances to the public that information about individuals or organizations or provided by individuals or organizations for exclusively statistical purposes will be held in confidence and will not be used against such individuals or organizations in any agency action.

(3) Protecting the confidentiality interests of individuals or organizations who provide information under a pledge of confidentiality for Federal statistical

programs serves both the interests of the public and the needs of society.

(4) Declining trust of the public in the protection of information provided under a pledge of confidentiality to the agencies adversely affects both the accuracy and completeness of statistical analyses.

(5) Ensuring that information provided under a pledge of confidentiality for statistical purposes receives protection is essential in continuing public cooperation in statistical programs.

The statements of legislative purpose set forth in section 511 are equally relevant:

(1) To ensure that information supplied by individuals or organizations to an agency for statistical purposes under a pledge of confidentiality is used exclusively for statistical purposes.

(2) To ensure that individuals or organizations who supply information under a pledge of confidentiality to agencies for statistical purposes will neither have that information disclosed in identifiable form to anyone not authorized by his title nor have that information used for any purpose other than a statistical purpose.

(3) To safeguard the confidentiality of individually identifiable information acquired under a pledge of confidentiality for statistical purposes by controlling access to, and uses made of, such information.

When read in context with the definition of “non-statistical activities” that includes “any administrative, regulatory, law enforcement, adjudicatory, or other purpose that affects the rights, privileges, or benefits of a particular identifiable respondent” [section 502 (5) (A)] or “is the subject of information requested or required to be supplied to an agency” [section 502 (6)], the language of section 511 sets out a policy standard that is fully consonant with current ethical norms. This standard recognizes that statistical confidentiality is not simply about disclosure, but also about how the government uses data and information collected for a statistical purpose.

The use and disclosure limitations contained in section 512 deal separately with statistical data and information and data or information for nonstatistical purposes. With regard to the former, the statutory protection, unless subsequently set aside by future legislation, is unequivocal

(a) **USE OF STATISTICAL DATA OR INFORMATION-** Data or information acquired by an agency under a pledge of confidentiality and for exclusively statistical purposes shall be used by officers, employees, or agents of the agency exclusively for statistical purposes.

(b) **DISCLOSURE OF STATISTICAL DATA OR INFORMATION-**

(1) Data or information acquired by an agency under a pledge of confidentiality for exclusively statistical purposes shall not be disclosed by an agency in

identifiable form, for any use other than an exclusively statistical purpose, except with the informed consent of the respondent.

(2) A disclosure pursuant to paragraph (1) is authorized only when the head of the agency approves such disclosure and the disclosure is not prohibited by any other law.

(3) This section does not restrict or diminish any confidentiality protections in law that otherwise apply to data or information acquired by an agency under a pledge of confidentiality for exclusively statistical purposes.

The last provision insures that strong confidentiality protections of existing agency-specific laws, such as Title 13, are not deemed to have been weakened by Title V of the E-Government Act. (Several other provisions Title V also reiterate the priority to be accorded to the protections contained in Title 13, particularly with regarding the demographic data or information collected by the Census Bureau.)

With respect to the use of data or information for nonstatistical purposes, section 512 (c) provides the following rule

A statistical agency or unit shall clearly distinguish any data or information it collects for nonstatistical purposes (as authorized by law) and provide notice to the public, before the data or information is collected, that the data or information could be used for nonstatistical purposes.

Despite its strong and clear language, the effectiveness of Title V will depend in part on the rules each agency develops for its implementation. This is particularly true with respect to section 512 (c) where the phrases “clearly distinguish” and “provide notice to the public” could be subject to a range of interpretations. Since the degree of clarity and extent of notice may be considered by some agencies to be inversely correlated with expected response rates, agency staff may find themselves torn between full compliance with the intent of Title V and more utilitarian considerations. In these circumstances, the Office of Statistical Policy within the Office of Management and Budget (OMB) has an important role to play as Title V specifically accords coordination and oversight responsibilities for the implementation Title V to OMB.

In closing this discussion of the statutory protections accorded to statistical confidentiality, it is important to remember, and the events of the past three years underscore the point, that the book on statutory protections is always open. Laws related to statistical confidentiality can be and have been changed. However, Title V does provide a strong base for mobilizing efforts to protect against future efforts to weaken the concept.

II. Understanding disclosure, harm, and risk in the context of statistical confidentiality

The term ‘disclosure’ as frequently used in the context of discussions of privacy and

confidentiality is narrow one. This narrow concept focuses on the *identification* of individual data providers, by individual curious, profit-seeking, or malignant persons or entities, in sets of *micro*-records maintained by statistical agencies. This narrow focus by statisticians is understandable, since many issues of identification, and disclosure so defined, can be readily addressed by a variety statistical approaches and methods. However, such a focus frames the issue far too narrowly as the discussion of ethics, statistical policy, and U.S. statutory protections just presented makes clear. Unfortunately, this narrow focus on the identification issue may cause us to neglect other, equally important concerns. Although identification is an important issue, two other questions are at least as critical: (1) who is attempting the identification and (2) why is the information being sought.

As noted in Section I, the concept of statistical confidentiality explicitly takes into account the “who” and the “why” of disclosure as well as issues related to identification. In particular, the concept of statistical confidentiality as it has evolved places a special emphasis on protecting against disclosure to non-statistical government agencies and programs (the “who” issue) for their use to adversely affect individual “rights, privileges, or benefits” (the “why” issue). Related to the “why” issue and, sometimes explicitly referred to in ethics and policy statements is the topic of harm. (Recall that President Taft used the concept in his 1910 Census proclamation.) While we are right to be concerned about harm, we need to be aware that it often raises issues also arise in the study of multiple causation. As noted elsewhere,

the determination of whether or not a specific activity gives rise to harm may itself be an ethical trap that can catch those who are unprepared. For example, one of the defenses offered by the doctors who carried out medical experiments on concentration camp inmates during WWII was that the research subjects "were doomed to die anyway" [Caplan, 1992: 266]. Indeed, no harm is often interpreted as "no added harm" or "no added risk of harm." Along the same lines, it has been said that the Dutch Jews would have been rounded up for deportation and the Japanese Americans would have been interned, regardless of the actions of those involved with the data systems that assisted in these endeavors. Implicit in these statements is the view that the statisticians and the others involved in these data systems didn't really cause harm and so did nothing really wrong. Using almost the same language, Jones [1981: 207] observed that a number of health officials associated with the Tuskegee syphilis study seemed to justify the continuation of that study, even after penicillin became available, on the grounds that the research subjects “were now beyond medical help, thus strongly implying that no real harm had been done...” [Seltzer, 2001]

Governmental statistical programs, given their long-term interest in promoting public confidence, are probably aware that such strained interpretations are unlikely to be productive over the long run. It may also be recalled that this “they would have died anyway” defense was rejected by the judges at the Nuremberg Trial. Moreover, the ethical principles that were developed in the course of that Trial placed a number of responsibilities on future experimenters. Among these were a duty to take into account the serious harm they can cause research subjects. Given the inherent power of governments and the special legal status accorded to the outputs of

federal statistical programs, it is only prudent that examinations of disclosure risks associated with these programs also consider potential harms flowing from such disclosures.

A special set of disclosure issues arise when the target of adverse governmental action is not specifically named individuals, but all members of a vulnerable population subgroup. Historically, in the United States most of the leadership in the federal statistical system has taken the view that the statistical system has no special obligation to protect against sub-group harm in contrast with individual harm. As a result statistical confidentiality and disclosure issues have been addressed largely in terms of individually identifiable persons or entities (i.e., *micro* data). Defenses of the Census Bureau's role in the forced expulsion of the Japanese American population from the West Coast were often expressed in these terms (see, for example, Barabba [1980] and Dedrick [1981]).

Some realization that this narrow view, strictly interpreted, is inadequate seems to be emerging, perhaps influenced by a greater awareness of the direct operational use of census *meso* data to, for example, help target Jewish neighborhoods in Amsterdam for attack in early 1941 [Seltzer, 1998: 515; 525] and assist the U.S. Army in early 1942 in locating Japanese American individuals and families living in the West Coast (see, Seltzer and Anderson [2000] and section III.A.2 below). Indeed, the concept *meso* data was introduced to identify a specific class of aggregated data that was effectively indistinguishable from identifiable *micro* data in terms of the harm which can be inflicted on those targeted.

However, even before the *meso* data concept was introduced, statisticians in the United States and elsewhere had begun to note the problem presented by the use of small-area statistical data to target vulnerable population subgroups. For example, Beeger and de Vries [1987], writing from a Dutch perspective, observed that the use aggregate census data to target vulnerable population subgroups, while perhaps not a technical violation of traditional census confidentiality provisions, represented a highly suspect use of population census data. Similarly, former Census Director Barbara Bryant, while continuing to cast the use of the 1940 Census tabulations to target Japanese American largely in terms of the traditional, narrow view of statistical confidentiality, concluded by observing that

The Japanese-Americans who lost property and were interned simply because of their ancestry see things differently, indeed. Their devastating fate has been cited repeatedly by critics as a breach in spirit, if not in fact, of census confidentiality. [Bryant and Dunn, 1995: 32-33]

Both Beeger and Bryant (and their co-authors) were merely reflecting the general proposition that simply following the letter of the law will not always lead to a result that is either consistent with the spirit of the law or one that can be easily reconciled with related policy statements or our ethical responsibilities.

In these circumstances, it is important that policies and practices aimed at promoting statistical confidentiality take into account the possible harms arising from *meso* data in addition to those associated with identifiable *micro* data. Accordingly, in our analysis we cover targeting

involving either *meso* or identifiable *micro* data. The recent increased caution of the Census Bureau with respect to the release of small-area statistics would also seem to be a reflection of a similar concern.

III. The Use of Population Data Systems to Target Individuals and Population Subgroups

Table 1 presents, in highly summarized form, a listing of instances in the United States and in other countries where efforts were made by national states to use a population census, a registration system, or a related activity to target individuals or vulnerable population subgroups for adverse action. The definition used is a broad one and includes examples, at one extreme, where the efforts lead directly to the use of identifiable *micro* data to assist in a grave human rights abuse to, at the other extreme, efforts that were never fully implemented. Furthermore, given the range of time periods and countries covered by Table 1, there is wide variation in the extent to which each data gathering activity listed was subject to statistical confidentiality legislation. Our justification for using a broad definition is simple. In view of the gravity of some of the examples, both for those targeted and for the statistical programs, agencies, and staffs involved, we consider that full exploration of the historical record is important so that we can assure that we have done all we can to avoid any new misuse by national or local governments.

[Table 1 about here]

Indeed, we view Table 1 as a sort of “incident report” similar in nature to ones used by those responsible for avoiding other types of public disasters in such fields as nuclear energy production, space exploration, or commercial aviation. In the study of rare events, with highly adverse consequences, it is important to examine not only the gravest disasters, but also the minor disasters, near misses, and problems avoided, to develop sound policies and practices. The scope of Table 1 is, however, limited by the exclusion of examples that are based solely on suspicion or suggestion, without some evidence that efforts were initiated to use a population data system for targeting or to otherwise significantly increased risks of vulnerable individuals or population subgroups to such targeting. Lists of these suggestive or speculative instances are presented toward the end of sections III.A and III.B below. (We should also clarify that though Table 1 refers to *macro* data in the body of the table, applications involving only the use of *macro* data, without some intention to also make use of *meso* or *micro* data, were omitted from Table 1.¹²) Also excluded from Table 1 are what we suspect are numerous unreported incidents where government statistical agencies in the United States and elsewhere have quietly and successfully resisted explicit efforts to use the statistical system to target individuals and vulnerable population subgroups. It would seem useful for such statistical agencies to keep a systematic record of these thwarted efforts and to periodically report on them publicly.

At this point Table 1 contains 15 “incidents.” Underscoring the rapidly evolving nature of this line of research, we note that in 2001 the last time we compiled the equivalent of Table 1, 10 incidents were listed [Seltzer and Anderson, 2001: 487]. The additional cases now included relate to: (a) the Australian aborigines, (b) Norwegian population censuses in the 19th and early

20th centuries, (c) the South African 1951 population census and that country's population registration system, (d) the 1910 U.S. population census and (e) the recent effort made to use information collected by the National Center for Education Statistics under a pledge of statistical confidentiality to investigate and prosecute terrorism. Because the number of cases listed remains small and the specific cases included is largely a function of the focus of the research carried out over the past five years, we do not attempt a quantitative meta-analysis of the experience summarized in Table 1. Instead, we will discuss these 15 examples individually in sections III.A and III.B below. Toward the end of these two sections, additional incidents not yet included in Table 1 – some speculative, some probable – are listed. We do note that two of the new cases added to Table 1 (the Australian aborigines and South Africa) were among the suggestive and speculative cases identified in our 2001 assessment [Seltzer and Anderson, 2001: 501-503].

A. The US record

Three U.S. examples where efforts were made to use or permit the use of data gathered through government statistical programs to assist in targeting individuals or population subgroups for criminal prosecution or forced migration have been studied in some depth so far. These are: (a) the use of regular decennial population censuses and various federally mandated special censuses in connection with policies and programs directed against Native Americans in the 19th century [Seltzer, 1999], (b) the actions of the Census Bureau in support of the forced migration and internment of Japanese Americans after the entry of the United States in World War II [Seltzer and Anderson, 2000; 2001], and (c) the use of the USA Patriot Act to permit the use of data collected by the National Center for Education Statistics under assurances of confidentiality to be used for the investigation and prosecution of terrorism [Seltzer and Anderson, 2002]¹³.

1. The 19th Century Native American Experience

The 19th century experience involving Native Americans, more fully described in Seltzer [1999] and summarized in Seltzer and Anderson [2001], did not primarily involve issues of disclosure. However, a number of treaty and administrative censuses carried out by the federal government were used for mixed statistical and administrative uses, and specifically to assist in planning the forced migration and the identification of individual . For example, according to DeRosier [1970: 137] a “careful census” of the Choctaw nation “was conducted to determine the exact number of Indians who were eligible for removal.” He noted that the results, which indicated that the nation consisted of 17,963 Indians, 151 white persons, and 521 slaves, were “important because the War Department wanted to move approximately one-third of the nation in the first group” and were used in ordering supplies to be used in the forced migration of the Choctaws in the early 1830s from their lands in the Southeastern United States. (This effort closely parallels the use of local area data and maps from the 1940 Census by the head of the Census Bureau's Statistical Methods Division where “a major problem in the [Japanese American] evacuation program was that of ... describing areas for a predesignated number of evacuees. The desired unit of evacuation was considered to be approximately 1,000 persons.” [Wartime Civil Control Administration, 1943, p.137]¹⁴).

The extent to which microdata from some special censuses were used to establish a system of

controls is illustrated by a 1901 memoir describing a census carried in the San Carlos Indian Reservation in what is now Arizona [Elliot, 1948: 98].

In 1884 a complete census had been made, the tribes being enumerated under their head chiefs and each camp of Indians of the same tribe under its head man. Brass tags of different shapes with one shape for each tribe had been provided. The band or subdivision of a tribe was designated by a letter of the alphabet, and each [member of a band had his number, stamped by the provost officer on the tag of the proper shape and given to each Indian whose name was recorded in books kept for the purpose. Each man was required to wear his tag at all times and to produce it when called upon Any failure to comply with these regulations was severely punished, and in a short time the system worked to the perfection I found it on my arrival.

Elliot also goes on to describe how these census-based metal tags were used to identify suspects in criminal identifications (Elliot, 1948: 100-101). (A similar uses of metal tags dating from the same period have been described by Hochschild [1998: 163] in connection with workers in the Belgian Congo, by Mamdani [2001: 12] in connection with the Hereros in German South West Africa after their release from concentration camps in 1908, and by Kraly [Personal communication, 2002] in connection with the aboriginal population of Australia.)

2. The Japanese American Experience

The first results of our ongoing research into the use of federal statistical outputs and staff in the forced migration and internment of the Japanese American population, the majority born in the United States and U.S. citizens, were given in Seltzer and Anderson [2000]. Some additional findings, based on further archival research, were included in Seltzer and Anderson [2002]. Since this research is continuing, the present report cannot in any way be viewed as definitive. Rather we consider it to be an interim progress report that focuses on the correction of earlier misinterpretations and the presentation of major new findings. Given this focus, much of the detail previously presented will be omitted, although some must necessarily be repeated for understanding. We present these results under four headings (a) the role of Census Bureau management and staff, (b) *macro* data, (c) *meso* data, and (d) unprotected *micro* data.

(a) The role of Census Bureau management and staff

Perhaps our most serious error of misinterpretation in our original 2000 paper was our treatment of Census Director J.C. Capt. This also distorted to some degree our presentation of the activities of Dr. Calvert L. Dedrick, who was reassigned from Chief Statistician in the Census Bureau's Division of Statistical Research to be the Bureau's "West Coast representative" and provided direct assistance to the military in the round-up and internment operations.

In our earlier portrayal, based on the evidence then in hand, we focused on Dedrick as the pivotal actor in the Bureau's involvement in assisting the Army in the internment program and treated Capt as playing a more passive role, at most providing enthusiastic, if naive support, to the activities of others. Based on a number of new materials since located, it is clear that Capt, as Director, provided leadership and direction to the Bureau's involvement first in defense and then

in war efforts, including the Bureau's assistance to the Japanese American internment program. Three examples Director Capt's leadership and initiative in these matters may be cited. First, as we have already described above in our review of the history of statutory protections accorded to statistical confidentiality (section I.c), it is clear that Director Capt was persistent and ultimately successful in his efforts to relax the confidentiality provisions of Title 13, an effort that he began within days of becoming Director.

Second, we recall that in our earlier study after recounting an exchange in the January 1942 Census Advisory Committee meeting between Dr. Leon Truesdell, the Bureau's chief population statistician and Director Capt that included the following colloquy,

Dr. Truesdell: ... We got a request yesterday, for example, from one of the Navy officers in Los Angeles, wanting figures in more or less geographic detail for the Japanese residents in Los Angeles, and we are getting that out....

...

Dr. Truesdell: That Los Angeles request I just referred to asked for census tracts.

...

The Director: We think it is pretty valuable. Those who got it thought they were pretty valuable. That is, if they knew there were 801 Japs in a community and only found 800 of them, then they have something to check up on....[Census Advisory Committee, January 1942: 20-21]

We commented on the apparent naivete of Capt's remark [Seltzer and Anderson, 2000, endnote 1]. Capt's reference to the number 801 was far less naive than we realized as it appears that this was the Japanese American alien population for Terminal Island according to the 1940 Census based on unpublished census tract tabulations.¹⁵ Terminal Island in Los Angeles harbor was the site of major Navy installations. It was also home to a thriving Japanese American fishing community. After the attack on Pearl Harbor, local officials and some naval officials came to perceive the fishing community as a major security threat. The Navy forcibly evacuated all the Japanese American residents of Terminal Island late February 1942.

Finally, we note two of Capt's telegraphic instructions sent to the West Coast during initial stages of the Bureau's co-operation with the Army. The first was sent on February 26, 1942 to Col. Karl Bendetsen, GHQ Fourth Army, San Francisco, who had the responsibility for planning and executing the actual expulsion operations. (In his work on the West coast, Dedrick reported to Bendetsen.) This message from Capt read

Dr. Calvert L. Dedrick leaving Washington 6:30 pm today ... He has authority to act in helping you find solution of the problems discussed with you and General Gullion yesterday." [NARA, RG40, Commerce, Entry 1, Box 144, File no. 67104]

Capt's second telegram was sent March 21, 1942 Capt and addressed to Dedrick, Hotel Whitcomb, San Francisco, and reads

Reference 9th paragraph your letter of March 19, phrase “Confidential until released” is habitual label long used by the Bureau and need not be taken too literally under emergency conditions prevailing in wartime. Your instructions confirmed by wire to Colonel Bendetsen gave you authority to act for this Bureau. Any information available to you should be used in accordance with good judgement as dictated by the needs of the national war effort. I am relying on you to be prompt, practical and effective in the performance of your duties without being hampered by old Bureau habits, precedents, and practices that are not in complete accord with the urgent, rapidly shifting necessities of the times as they develop from hour to hour. Of course you understand that it is important to the Bureau that it receive official and public recognition for all its work. J. C. Capt – Director – Census. [NARA, RG40, Commerce, Entry 1, Box 144, File no. 67104]

None of this detracts from Dedrick’s outstanding technical and organizational skills, which he employed in full measure in the support he provided to Col. Bendetsen during the planning and operational phases of the exclusion operations. But it is now clear that Dedrick’s work and the other aspects of the Bureau’s cooperation with the Army in this matter were carried out in line with the policies of Director Capt and in response to his over-all direction.

(b) The role of *macro* data

In our initial report on our research into the Japanese American internment in 2000, we provided some information on the rapid production and the general dissemination of *macro* data from the 1940 Census on Japanese Americans in the days immediately after Pearl Harbor (see, for example, Seltzer and Anderson [2000: 5-6]). It is now clear that in addition to the public dissemination previously described, Director Capt was also providing these data directly to the Army and Naval intelligence services as well as to Sumner Wells in the State Department. (Wells and some of his staff played an important coordinating role on behalf of the President with respect certain war-related intelligence operations.)

On December 10, 1941, for example, Capt provided the December 9 release to Harold B. Hoskins, Executive Assistant at the State Department. On December 13, Hoskins in turn requested Capt to send additional copies of these data to officials in the Office of Naval Intelligence, Army Military Intelligence and the FBI. On December 11, Henry Field, the intelligence official, who had requested the tabulations from the Census Bureau in November, transmitted them to Undersecretary of State Sumner Welles; on December 17, Capt forwarded additional tabulations to Welles at Field’s request.¹⁶ [NARA, Record Group 59, Records of the Department of State, Decimal File, 811.5011/261-1/2; 811.5011/2662 PS/LIC; 740/00115 Pacific War/419]

Along the same lines, on December 18, 1941, the Census Bureau sent a telegraphic transmission of Japanese native born, foreign born and total population figures to the Western Defense Command in San Francisco. The tabulations listed the Japanese population in Western Defense Command by race for minor civil division in California, Oregon and Washington, and for the total Japanese American population for Utah, Nevada, Idaho, Montana and Arizona. [NARA, Record Group 338, Western Defense Command and Fourth Army, Wartime Civil Control

There is, of course, nothing illegal or wrong in such actions. However, the Census Bureau defined “Japanese” as racial category from the start, while the other agencies involved -- the INS, the FBI, and the military – tended to see the issue, initially at least, as one relating to Japanese aliens. During January 1942 the military began to switch its view of the problem from dealing with enemy aliens to one that defined the problem in racial terms (i.e., the Census concept). Certainly, a variety of factors contributed to this change in view, which ultimately shaped the entire exclusion and internment program. The issue we continue to ask is what role did the use of the “race” concept by the Bureau in defining “Japanese” contribute to this change, coupled as it was with the energetic efforts of the Census Bureau to tabulate and disseminate data on the Japanese American population?

(c) The role of *meso* data

It was our understanding that the evidence presented in our 2000 study relating to the use of *meso* data from the 1940 Census to provide direct assistance to the Army operations directed against the Japanese American population on the West Coast, was generally accepted. We have recently learned that some have questioned whether such assistance in the form of *meso* data from the 1940 Census was provided by Dedrick and the Census Bureau. These doubts seem to be attributable to the fact that the Bureau has been unable to document from its own internal records some of the assertions relating to the use of *meso* data contained in Seltzer and Anderson [2000]. Accordingly, we will summarize here the what we consider the strongest evidence contained in our 2000 paper related to the use of *meso* data, as well as present some of further material documenting such use.

As we noted earlier [Seltzer and Anderson, 2000: 7], Dedrick himself confirmed the use of detailed *meso* data from the 1940 Census to assist the Western Defense Command in the process, stating that at the end of February 1942 when he was in San Francisco, the Census Bureau was asked to provide him,

a detailed cross-tabulation for even the most minute areas, the smallest areas for which data were collected. In other words, enumeration districts and in some instances cities by blocks.... Sheets of paper from the tabulation machines were sent out to WCCA at the hotel in Market Street in San Francisco, and became the basis for the WCCA statistical activities. [Dedrick, 1981: 172-173]

We also noted [2000:7]

These recollections [of Dedrick’s] appear to be confirmed by contemporary Bureau records that included in one of its listings of war related activities for 1942, the provision of “photographic and photostatic copies of block maps for 10 cities” to the Western Defense Command [US Census Bureau, 1942a, Exhibit A:12].

Finally, several reports prepared by Census Bureau staff which we cited [2000: 13-15] included references to the Bureau providing various types of *meso* data to the Army as part of the Bureau's defense that identifiable *micro* were not provided (see, for example, Bohme [1975], U.S. Census Bureau [1982], and Bohme and Pemberton [1991], although there is considerable variation among these sources concerning the geographical detail involved [Seltzer and Anderson, 2000: Table 1].

Although we have not located the urban block maps the Bureau reported it provided to the Western Defense Command in 1942 [US Census Bureau, 1942a, Exhibit A:12], and which Tom Clark apparently recalls being used in planning the round-up operations [Clark, 1972], we have found examples of their rural counterparts. (As discussed below in section IV, "Barriers to the study of disclosure, harm, and risk," there may have been an active effort to suppress certain details of these operations.)

In terms of new material we will cite two examples of the availability and use of detailed 1940 Census *meso* data in planning the round-up of Japanese Americans. Dedrick's Statistical Section, working under Col. Bendetsen, produced four statistical bulletins, "Japanese Population and Family Heads, by Minor Civil Division: 1940," each dated 3/16/1942, for Arizona (Bulletin 2), California (Bulletin 3), Washington (Bulletin 4), and Oregon (Bulletin 5), respectively. Each report consists of (a) a brief introductory text stating that "The Bureau of the Census ... furnished the data for this bulletin by special request ..."; (b) a state map with county boundaries, showing for each county the number of "Japanese" persons and family heads; and (c) a table showing the same data by county and minor civil division [Wartime Civil Control Administration, 1942a-d]¹⁷

That the contents of these bulletins, coupled with more detailed cartographic information than contained in the bulletins themselves, had direct operational consequences is clear from a one page memorandum that Dedrick, as Chief Statistics Division, wrote to Col. Bendetsen. This memorandum, dated 3/21/1942, provided in tabular form 1940 Census data, by minor civil division, relating to the 281 Japanese and 48 Japanese family heads, by citizenship status, and the 40 Japanese farm operators, together with a map of Bainbridge Island with minor civil division boundaries. Dedrick's memorandum, a copy of which is reproduced as Figure 1, constitutes one element of the Western Defense Command's "Evacuation Memorandum No. 1 [Dedrick, 1942]. March 21, 1942 was a Saturday. On Sunday, the Army "ordered the forcible evacuation of all Japanese from Bainbridge Island in Puget Sound ... effective March 30." [NY Times, "Army Ousts Japanese Near Bremerton Yard," 3/23/1942, p. 1]

[Figure 1 about here]

(d) The role of *micro* data

The issue of whether identifiable *micro* data from the 1940 Census pertaining to the Japanese American population were actually used in efforts directed against them remains an open one. There seems no doubt at this time, however, that in early 1942 such a use was intended by Census Director Capt and the Congress. In 2000, we had noted [Seltzer and Anderson, 2000: 9] that

at the same January 1942 Census Advisory Committee Meeting, in which Director Capt and his senior staff indicated that by January 10 that they were

already providing tract level data on the Japanese Americans to the military, Director Capt expressed his willingness to take the next step: “We’re by law required to keep confidential information by individuals. But in the end, [i]f the defense authorities found 200 Japs missing and they wanted the names of the Japs in that area, I would give them further means of checking individuals. [Census Advisory Committee, January 1942: 21]”

Again, Capt’s observation was not an idle one: we now know that he had been working for months to change the law barring such uses of confidential census data. Indeed, he renewed his efforts several weeks after the January 1942 Advisory Committee meeting by appearing before the House Judiciary Committee and presenting an “amendment to the omnibus war-powers bill ...” to among other things, “make this and other information now obtained by the Department of Commerce under the seal of confidence available to other war agencies.” [Letter from Commerce Secretary Jones to Hatton Sumners, Chairman, Judiciary Committee, dated 2/4/1942.¹⁸ Reproduced in U.S. Congress. House. Committee on the Judiciary. “Second War Powers Act, 1942.” Report No. 1765. 77th Congress, 2nd Session, February 9, 1942.] The Judiciary Committee quickly approved this amendment, and at least according to a NY Times story dated 2/6/1942, Capt’s intentions in introducing the amendment were clear,

... The House Committee on the Judiciary before giving its approval today to the Senate adopted Second War Powers Bill, amended the measure to take back the promise it made in 1940 to that all data obtained by the Census takers would be held strictly confidential, even from other bureaus and agencies ... Some agencies of the government want data now as a matter of national safety. They seek some of the information obtained particularly from Japanese and others who since have become enemy aliens, especially about those in costal areas from which they have been ordered evacuated by the Department of Justice ... [Such] data, now a secret under law, government officers believe, would be of material aid in mopping up those who had eluded the general evacuation orders.” [NY Times, “SPY DATA SOUGHT FROM 1940 CENSUS.” 2/7/1941, p. 9]

It is also clear that the motivation cited in the NY Times story was not simply journalistic speculation. Some weeks later, Senator Warren Austin (R-Vermont), who was a member of the House-Senate conference committee dealing with this legislation, remarked in the course of a closed March 13, 1942 hearing of the Senate Committee on Military Affairs

we are now considering ... legislation that would empower the Secretary of State to requisition from the Commerce Department, Bureau of the Census, what the census shows about these people. That would give an enumeration of the Japanese and it would also give names and residences, so that, when the Army makes its evacuation it can ... compare its list of evacuees against the census and have some knowledge of whether this has been an effective protection or not.” [U.S. Congress. Senate. Committee on Military Affairs. 1942: 8-9]

Although Senator Austin misstated the administrative path by which the formerly confidential

census information would be obtained under the Second War Powers Act (the law, as enacted, provided that the head of “any branch or agency of the Government” could request such information if it was needed “for use in connection with the conduct of the war,” not merely the Secretary of State), the planned use, “mopping up” as the NY Times termed it, seems quite consistent with other available evidence. Indeed, we had previously speculated about such a selective use of *micro* data [Seltzer and Anderson, 2000: 10].

Despite the very strong evidence that, in early 1942, both the Census Bureau and the Congress had every intension that 1940 Census identifiable *micro* data be used to assist in the expulsion and internment of the Japanese American population, the evidence on whether such *micro* data were actually so used is still conflicting. Table 2 presents the evidence located to date that seems to support or contradict actual use. Since a number of important elements listed in Table 2 were located as recently as two months before this paper was written, we consider that the specific question posed by the table needs further investigation. However, given that we now know such a use was intended by the Bureau and made lawful by Congress, the importance of resolving the issue of whether or not there was actual use seems somewhat diminished.

[Table 2 about here]

3. Education statistics and the USA Patriot Act

As we recounted in some detail relatively recently [Seltzer and Anderson, 2002], apparently in response to an idea generated internally within the National Center for Education Statistics (NCES) or one of its contractors, a serious belief emerged in the Department of Education and the Justice Department in the days after 9/11 that NCES data files might provide useful leads in the investigation or prosecution of terrorism. As a result, as described above in our review of the legal protection provided to statistical confidentiality above (section I.c), the strong confidentiality laws that hitherto assured that NCES identifiable *micro* data could not be used to target individuals were relaxed by the Patriot Act to permit such targeting in connection with terrorism investigations and prosecutions. There is little new to add at this point beyond the possible impact of subsequent adoption of the Education Research Reform Act of 2002 and Title V of the E-Government Act of 2002, both discussed above in section I.c. The fact that NCES staff appear to have been severely restricted in what they can say publicly about the matter also has hampered further understanding of how and why things evolved the way they did. In retrospect, we consider the NCES incident, a prime example of what has sometimes happened to the legal status of statistical confidentiality during times of national stress.¹⁹ The fact that, as far as we know, there has been no use of NCES *micro* data for such investigative purposes, may be less an indication of investigative restraint than that the most useful NCES data sets from an investigative perspective, are based on samples.

4. Identified, but largely unstudied examples

In addition to these cases which have been subject to extensive recent active research, other examples beginning to be examined more carefully. Principal among these are the extent of disclosures, if any, flowing from the opinion of the Solicitor of the Department of Commerce, issued within three months of the country’s entry into World War I, permitting the Director of the Census at his discretion to provide local draft boards and the Department of Justice with the names and ages of suspected selective service law violators taken from the 1910 Census *micro*

records [Bohme and Pemberton, 1991:10-11]²⁰.

This opinion was issued in response to a request from the Census Bureau Director made in June 1917 [Rogers, 1917a] that began

I have received numerous requests from registration officials in various parts of the United States to furnish them with information from the census records, showing the ages of men who they believe have failed to registered, although between the ages of 21 and 30 ...²¹

The request goes on to quote from President Taft's 1910 proclamation issued in connection with the 1910 Census²²

The census has nothing to do with taxation, with army or jury service, with the compulsion of school attendance, with the regulation of immigration, 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.

as well as from the statement of Census Director Durand "to each head of a family"

Read the President's proclamation on this schedule, which emphatically states that you can not be harmed in any way by giving information to the Census.

Observing that the proclamation would seem to prevent the Census Bureau from furnishing the information desired by the registration officials, Census Director Rogers nevertheless states in explicit terms his objective,

I believe that every branch of the Government, including this bureau, should assist at the present time, so far as possible, in securing a full registration. Accordingly, it is recommended that the matter be taken up with the President, with the view to having an order issued waiving the rigid rule laid down in Ex-President Taft's proclamation, and authorizing this Bureau to supply the proper officials (both registration and Federal) who are in control of the registration and prosecution of individuals who have failed to register, with data from the census schedules, which may show the ages of such individuals.

Within four days, the Acting Solicitor issued the requested opinion.²³ It gave the Census Director the authority to provide names and ages to the registration authorities based on such considerations as,

It does not appear that any person will be harmed by the furnishing of the information desired and for the purpose which it is desired ... Other provisions of the law prohibit the Director of the Census from giving certain information in regard to the business of individuals, firms, and corporations, but these provisions do not, in my opinion, apply to information in regard to names and the ages of

individuals ... There is nothing in the law or the proclamation which manifests an intent to restrict the Government, through the head of the Department, from furnishing names and ages of individuals as recorded in the Census office. [Department of Commerce, 1917]²⁴

Although Bohme and Pemberton [1991:10] refer, in general terms, to the actual use such of information in individual prosecutions, we are unaware of any research dealing with such cases. Bohme and Pemberton [1991: 11] also refer to the use of *micro* data from the 1920 Census in connection with deportation cases. Again, the matter has not been further studied to our knowledge.

Finally, in 1930 a series of legal opinions gave greater status to the laws relating to census confidentiality, permitting or requiring the Census Director to turn down requests for *micro* data. In retrospect, some of the requests seem benign, such as the request from the Woman's Bureau for individual *micro* level data relating to women's employment [Bohme and Pemberton, 1991:12].

Other requests, also refused, dealing with lists of illiterates in several parts of the country, may well have been far less benign.²⁵ Illiteracy at that time was one of the factors considered by national and local eugenic activists as possible grounds for involuntary sterilization and similar negative eugenic efforts. Such a census-based list, as was also true of the lists of non-registrants or aliens, could be directly admitted as evidence in most legal and administrative proceedings.

B. The international record

1. Holocaust (*Shoah* and *Porajmos*) related examples

Six of the fifteen examples listed in Table 1 relate to the targeting of Jews and Gypsies (more properly referred to as *Roma*) for segregation, forced migration, and extermination by the Nazi authorities and some of their allies during World War II. Five of these examples were discussed in detail by Seltzer [1998] and all six of them in more summary form in Seltzer and Anderson [2001]. While the activities in each country are listed as one "incident" in Table 1, several are compound in nature. For example, while Seltzer [1998] describes the use of a 1942 special census to identify Jews in Norway as a preparatory step to their expulsion from the country, but one not carried out by the Norwegian statistical office, Sjøbye [1998] describes the persistent efforts of the Director-General of that office to take advantage of the situation to establish a population registration system.

We will not repeat here the detailed descriptions of these six cases contained in Black [2001], Seltzer [1998], and Sjøbye [1998] and the sources cited therein. As we observed in Seltzer and Anderson [2001: 486]

although these six cases were Nazi-inspired crimes, in only two cases, Germany itself and Poland, could the misuse of the data systems be attributed solely to Nazi initiatives. In France, Henri Buhle and René Camille, and in Norway, Gunnar

Jahn, the heads of the statistical agencies, took advantage of the political climate of German occupation or influence, to expose vulnerable target populations to further risks by proposals to undertake major new data-gathering efforts to serve both statistical and administrative purposes (Rémond, 1996; Søybye, 1998).

With respect to the Netherlands we noted [2001: 486; 488]

the effort at establishing a comprehensive population registration system for administrative and statistical purposes was completed even before the Nazi-occupation [Methorst, 1936; Thomas, 1937]. In 1938 Methorst, who was then the director-general of the Dutch Central Bureau of Statistics and formerly also head of the Dutch office of population registration, reported on the rapid progress being made in the Netherlands in implementing a new comprehensive system of population registration that would follow each person “from cradle to grave” and would open up “wide perspectives for simplification of municipal administration and at the same time social research” [1938: 33: 713-714]. By early 1941 Methorst’s successor as head of the population registration office, J.L. Lentz, had quickly adapted this general “cradle to grave” system to create special registration systems covering the Jewish and the [Roma] populations of the Netherlands. These registration systems and the related identity cards played an important role in the apprehension of Dutch Jews and [Roma] prior to their eventual deportation to the death camps ... The critical role of the registration system in the over-all process has been stressed by such diverse observers as the German *Generalkommissar* for administration and justice in the Netherlands in September 1941 [Presser, 1969: 38] and the British historian Bob Moore [1997].

However, some discussion of subsequent developments related to events in France, Germany, and the Netherlands seems called for.

France. The dispute over the precise role of René Carmille, head of the French statistical service under the Pétain, referred to in Seltzer [1988: 522-523], appears to continue. Led by his son, Robert Carmille, there is an ongoing effort to rehabilitate Carmille’s reputation by, among other things, alleging that some of his communications to his superiors in the Vichy government were only written to deceive the government while he was deliberately using his post as head of the statistical service to sabotage the government’s anti-Jewish measures. However, René Carmille’s role, and the role of the French statistical service itself, appear to remain ambiguous at best, pending further independent research.

Germany. Our understanding of the role of the German statistical service in the Holocaust continues to benefit greatly from ongoing research by a large number of German scholars. This research, some carried out by independent scholars and some working under commission from the German statistical service, has examined operations of individual population data system and individual statisticians and demographers under the Third Reich. Most of this extensive body of research is available only in German, see for example, Aly and Roth [1984] and Weitog [2001], both previously cited in Seltzer and Anderson [2001]. Fortunately, the German

Association for Demography is planning to hold a a bilingual seminar on the subject in December 2003.

The Netherlands. Despite the indisputable role that the ID cards linked to the Dutch population registration system played in the Jewish and Roma Holocausts in the Netherlands and the role that *meso* data from the 1930 Census played in the dot maps used to target Jewish neighborhoods, senior officials of Statistics Netherlands have persisted in the view that the statistical service, its leadership, and its outputs were uninvolved. For reasons set out in more detail previously [Seltzer, 1998: 523-526; Seltzer and Anderson, 2001: 486, 488], we think their seemingly uncategorical denials are unjustified. We would simply note that after the end of the war, Statistics Netherlands dropped religion as a census item and the Dutch population registration introduced a series of operational safeguards, such as decentralization, designed to make its use for targeting purposes more difficult. Despite these measures, it appears that the Dutch population had such reservations about the statistical system that they effectively sabotaged the planned 1971 population census. We are also unaware of any knowledgeable, independent research on the issue that confirms the view that in the 1930s and early 1940s the Dutch statistical service, its leadership, and its outputs were had nothing to do with ensuing events. One issue that such research might attempt to address is the extent to which the Dutch central statistical agency and its Director-General used the mantel of that agency's influence and prestige in the 1930s to secure the passage of the legislation needed to strengthen the county's population registration system.

2. Other European examples

Norway (19th Century)

A recent study by lie [2002] describes the activities of the Norwegian Central Bureau of Statistics (CBS) in defining and targeting two ethnic minorities, the Samis (an indigenous people from Northern Norway, some times referred to as Lapps, a pejorative term) and the Kven (immigrants from Finland and their descendants) beginning in the latter half of the 19th century. With respect to, *macro* data, it appears that the Norwegian CBS shaped the definitions and classifications related to these two ethnic minorities to further government policies related to "Norwegianization" and boundary disputes with Finland, and related security concerns. With respect to identifiable *micro* data, Lies [2002: 805] reports that the CBS archives contain a number of examples during this period "where transcripts of the personal [census] forms of certain individuals and groups were made at the request of other governmental agencies ... But the only examples found of transcripts of information from whole groups, concerns the minorities." He also cites [2002: 813] a specific example where a transcript was provided, on request, to the chair of the so-called Lapp Commission, who was also a district attorney, containing individually-identified personal data for "all Sami in and south of ... Søndre Trondhejn" from the 1891 population census, including data on occupation, nationality, family ties, mental health, and ownership of property and lives stock.

The former USSR

At this point the best summary source that we are aware of on this experience remains Blum [2000]]. As we indicated previously [Seltzer and Anderson, 2001: 492-493], that study "reconfirmed several instances when census *micro* data were used to target minority population

groups for forced migration and other human rights violations. It also found that the evidence seemed to indicate that by the time of the 1937 census, Stalin was relying on other data systems for *micro* data and that census was primarily used as a source of *macro* data to evaluate policies, including those of forced migration and other programs with human rights consequences.”

3. Elsewhere

Australia (Aborigines)

As part of an ongoing research effort aimed at documenting and better understanding the use of population data systems by national and state governments in Australia to monitor and control the Aborigines, Kraly and McQuilton [2002], have produced an initial report dealing with Victoria, first as a colony and then as a state in the early Federal period. In line with the conceptualization, used here and in our previous studies, Kraly and McQuilton present some there results under the rubric of *macro*, *meso*, and *micro* data. Through the early Federal period, in Victoria at least, there seemed to be relatively little interest *macro* data relating to the Aborigines, although unlike several other parts of the country population censuses after 1857 made attempts to cover Aborigines. By contrast, there was great interest in *micro* and *meso* data in the administration of this population, including an early emphasis on securing *micro* records with identifiable names. A population register was established to help control the population required to live in designated reserve areas and to exclude persons of mixed race from living in these reserves. The population register was also used to facilitate forced migration.

Since it is known that practices varied widely across Australia during the period covered, it is important to await further research, most of it archival in nature, before attempting to draw more general conclusions.

Rwanda (1994 Genocide)

In Seltzer and Anderson [2001: 493], we briefly summarized the population registration system established in the colonial period, noting that

in the 1930s this registration system was used to help fix the identity of the population in terms of the hitherto somewhat amorphous categories ‘Hutu’ and ‘Tutsi,’ primarily to assist a pro-Tutsi policy by the Belgian colonial administration based on pseudo-scientific racial grounds.

The registration system continued to function well as a statistical and administrative activity right up to the out break of the genocide in 1994, which made use of the system to target the Tutsi population in some areas. We also commented in connection with our discussion of needed research [2001: 500]

Moreover, although the use of the population register in the 1994 Rwanda genocide has been established, its role has yet to be spelled out in any detail.

Unfortunately, as far as we are aware no further research along these lines has been undertaken.

South Africa (Apartheid)²⁶

In our previous study, we also identified South Africa between the 1930s and 1993 as a possible area for further research, noting [2001: 501]

Although de Klerk [1998: 74] identified the Population Registration Act as the “cornerstone” of *apartheid*, given the way that so many parts of the state administration were used in furthering systematic abuses directed at the nonwhite population and the long gestation period of the system, a full review of activities related to all population data systems, including the population registration system and regular and special censuses, seems in order.

Such a study is now underway focusing initially on the period 1950-1959, and in particular, on 1950 Population Registration Act, on its early implementation, and on the role played therein by the Department of Census and Statistics, one of its Directors, J.I. Raats, and the 1951 Census of South Africa.

In brief, in May 1948, the Nationalist party won the Parliamentary elections in South Africa, and their leader, Dr. Daniel F. Malan became Prime Minister. They were elected on platform that emphasized rigid racial separation, *apartheid*, although the term meant quite different things to different members of the government and its supporters. It appears that in one or two of the influential Nationalist documents published between 1946 and 1948, reference was made to the importance of establishing a population registration system. In any event, about six months after the election, the Prime Minister indicated that his government might establish a system of “racial registration” [NY Times, “South African Hints Racial Registration,” 11/17/1948, p. 7]. After some delay, the government introduced legislation establishing such a population registration system in February 1950. (During the course of the Parliamentary debate the government attributed the delay to the need to send J.I. Raats, Director of the Department of Census and Statistics, and another civil servant on a study tour to Europe and North America to collect experience on population registration in other countries.)

Among the features of this legislation of special relevance to our paper were that: (1) it provided for the establishment of a population registration system and a related system of personal identification cards, (2) each person above a minimum age was assigned to one of three main racial groups (white, “Colored”, and Native) with some additional ethnic/tribal information obtained for those assigned to the latter two groups, (3) the initial assignment of “race” was to be made primarily on the basis of information to be reported in the 1951 Population Census, (4) the responsibility for assigning race to each person was given to the Director of Census and Statistics, (5) the responsibility for implementing the Act, including the registration system and the ID cards, was given to the Department of Census and Statistics, and (5) the oversight responsibility for the implementation of the Act to assigned the national Statistical Council [South Africa (Union of), 1950].²⁷

The motives for and many features of the legislation were hotly debated in both in the Assembly and in the Senate and the final measure was not adopted until June, after the rules of debate were altered and closure invoked [South Africa. Parliament. House of Assembly, 1950a; South

Africa. Parliament. The Senate of South Africa, 1950]. In terms of its initial objectives it appears that the Act was serving a twofold purpose: first, was an important early step in developing the infra-structure of the *apartheid* system, and second, it was specifically aimed in the short run at the “Colored” population (ie, the mixed- race and Indian populations), who the government perceived as a threat in racial and political terms. Hence, initially at least, targeting actually attributable to the 1950 Population Registration Act seemed to be directed against the “Colored” population.

During the course of the debates Director Raats’ views and expertise were frequently cited by the Interior Minister and other members of the majority coalition in support of the bill and he was a major witness before the Select Committee created to finalize the bill in the middle of the legislative session [South Africa. Parliament. House of Assembly, 1950b]. Indeed, on the basis of his statements at the Select Committee, it is clear that Raats, himself, was the major author of the legislation as it was introduced and that on his own initiative he designed an exemplar ID card circulated to Select Committee members.

Moreover, Director Raats and the whole South African statistical system were often invoked explicitly to provide legitimacy to the proposed registration system. For example, the Minister of Interior, T. E. Dönges, at the end of his lengthy opening statement introducing the bill stated

In conclusion, I want to say that the this whole system will be under the supervision of the Statistical Council. This is a council brought into being under the Statistics Act ... I may just add that the Statistics Council has expressed itself very strongly in favor of the introduction of a population register. The fact that the Statistical Council is entrusted with the task of supervising the functioning of this scheme in general is the best guarantee of the scientific character of this system and of its bona fides ... [South Africa. Parliament. House of Assembly, 1950a: cols. 2523-2534]

Despite Director Raats’ role, it is clear that the information he obtained during the course of his study tour was not always correct and the legislation as drafted was flawed in a number of respects. Nevertheless, the legislation was passed, the 1951 Census was conducted, the Department of Census and Statistics received added posts and accommodations to support its new implementation responsibilities, and the national Statistical Council continued to approve the work being carried out. Throughout Mr. Raats’ directorship, his annual reports reflect great optimism, slow progress, and the continued need for more resources. That things were far more complicated can be seen, for example, in a brief story that appeared in the NY Times in 1955, dealing with the implementation of the Act, that noted that “the director of the census, J.I. Raats ... personally classified 700 persons after separate interviews.” [New York Times, 8/21/1955, “KIMBERLEY BEGINS TO CLASSIFY RACES; Some Who Had Considered Themselves Coloreds Are Placed in Negro Bracket,” p. 26]. A more extended documentation of some of this classification work may be found in the report of the human rights worker, Muriel Horrell [1959].

In 1956, Director Raats left his post as director and was replaced by his deputy, Dr. H.M. Stoker [South Africa. Bureau of Census and Statistics, 1960]. The next annual report, prepared by Director Stoker, was far more gloomy. The 1959 annual report of the national Statistical Council contains the following entry under “Population Registration”

as from 1st October 1959, separate establishments were created in respect the Bureau of Census and Statistics on the one hand, and the Population Register on the other. It takes note of the proposed amending legislation before Parliament in this respect and that it is the intention to terminate the supervisory function of Council regarding the Population Registration Act. It takes cognizance of the verbal information supplied by the Director, that substantial progress has been made in the work of the Population Register. [South Africa. Bureau of Census and Statistics, No date: 2-3]

With these changes, the added posts and office space quickly vanished and for a number of years thereafter the South African Department of Census and Statistics appears to have fallen into the shadows.

This account is only a summary of the information gathered to date on the South African experience and many lacuna remain. However, even from this abbreviated and partial report, the deep and extended involvement of the South African statistical service and its leader in diverting many of the core functions of the statistical system, including the population census, to harm individuals and population subgroups is clear. A more detailed report will be presented in Seltzer and McNeil [In progress].

4. Not yet studied, includes probable and speculative

In Seltzer and Anderson [2001: 501-503], we listed a number of counties and situations, outside the United States, where further research might be productive based indications that ranged from the suggestive to the speculative. That listing included: Colonial Africa (the late nineteenth century to the 1950s), South Africa (1930s-1993), Namibia (particularly 1960s-1980s), Korea and Taiwan (1890s-1990s), Countries of Eastern Europe (1948-1990s), the treatment of the *Roma* (Gypsy) population in many European countries (since 1900), persons of Japanese ancestry in several Latin American countries (1942-1945), the treatment of minority populations in China (since 1950s), the treatment of minority population groups in Finland, Sweden, and Norway, Aborigines in Australia, and suspected Irish “terrorists” in the United Kingdom, Arab “terrorists” in Israel, and Israeli “terrorists” in a number of Islamic countries. Studies dealing with three of these topics are already underway and preliminary results are reflected in Table 1 (the Australian Aborigines, the Samis and Kvens in Norway, and South Africa in the 1950s).

A few additional comments may be useful. First, we add to the list of areas where further research on targeting might be productive the several countries in Central and South America that were under severely repressive regimes, particularly during the 1970s and 1980s. Second, while the phrase “Arab ‘terrorists’ in Israel” was used generically, we note that the Palestinian Authority’s Statistical Office was attacked by Israeli armed forces in 2002, and it is alleged that a full set of records, with individual identifiers, from the Palestinian population census was

removed by Israeli security forces. Whether this allegation is true, and the use, if any, being made of these records needs further investigation. Third, the generic use of the phrases, “Irish ‘terrorists’ in the United Kingdom” and “Israeli ‘terrorists’ in a number of Islamic countries,” also could well be pursued. In this connection, we would suggest that practices in Northern Ireland, particularly when it was under military control be examined. Finally, we would stress again the value of promptly undertaking studies of “targeting” by the statistical services of the countries of Eastern Europe when they were under the control of the former USSR and the Communist Party. The countries all had strong central statistical services, active programs of population data collection, and a strong potential interest in targeting. But what actually happened? Where there differences among countries? If so, what accounts for these differences. The answers to these questions may well help statistical services elsewhere cope with future threats. It appears that the time is now ripe for such research. On the one hand, state security archives have opened in a number of these countries and, on the other hand, many individuals who may have been involved with both statistical and security operations may still be alive.

IV. Barriers to the study of disclosure, harm, and risk

This paper focuses on substantial harms or threats of substantial harm that have been identified as associated with national statistics agencies and their operations, principally through the use of *meso* and *micro* data to target individuals or members of vulnerable population subgroups. Although some of these harms have been of the gravest sort, the number of such instances identified and documented is not large in absolute terms. How much solace can be drawn from the fact that so few cases have been so far identified? We believe, at this point, only a limited sense of security is justified.

Three broad groups of factors, in our view, account for this state of affairs. First, the subject has not been studied in depth until comparatively recently. The first study in the United States appeared only in 1998 [Seltzer, 1998]. Thus, the full dimensions of the problem has not been determined. So far, moreover, further study has identified further confirmatory instances.

Second, research on the topic is inherently difficult, often involving material deposited in diverse archives²⁸ and requiring multidisciplinary skills that include knowledge of governmental statistical operations, methods of historical research, and familiarity with the general history of the era involved. In addition, pertinent material may not be referenced in finding aids or may be discarded altogether as of no historical interest.

Moreover, the fact that the topic has only recently been identified as a distinct subject of research has in the past combined with these inherent difficulties to apparently render the issue invisible to otherwise well-qualified researchers. For example, in the course of his well-known study of the forced evacuation of Japanese Americans from the West Coast, Mortin Grodzins [1949: 332-337] quotes and discusses at some length the unpublished March 13, 1942 hearing of the Senate Committee on Military Affairs without even noting that this hearing contains an explicit

statement of Congressional intent concerning the use of 1940 Census *micro* data to pursue Japanese Americans missed in the initial Army round-up [U.S. Congress. Senate. Committee on Military Affairs, 1942: 8-9] (see section III.A.2.d above, for details).

Far more recently, and even where national statistical systems and statistical confidentiality were the focus, both Duncan et al. [1993] and Seltzer [1994] underplayed or ignored the complicity of statistical systems and personnel in targeting and gross human rights abuses.

Over and above these problems of focus, effort, and expertise, other factors of a less neutral sort also operate to thwart the discovery and documentation of efforts to divert a national statistical system from its traditional mission to one of targeting individuals or members vulnerable subpopulations. In particular, statistical agencies and their supporters, together with the involved civilian, military, and intelligence agencies, often have little interest in discussing such incidents and sometimes take active measures to suppress any knowledge about them altogether.

At times, statistical agencies and some of their supporters may be reluctant to discuss the subject over fears that any public airing of the issue will adversely impact future response rates. (In these circumstances, those who cite questionable past practices are sometimes labeled “irresponsible.”) Moreover, in many countries, those non-statistical agencies most commonly associated with the use or attempted misuse of population data systems (Ministries of the Interior or Justice, the military, and intelligence agencies and security police, etc.) have routine policies and practices that limit the availability and accessibility of pertinent archival material.

As difficult as it is to identify and document some cases of disclosure and misuse, it is often equally difficult to identify efforts aimed at the deliberate suppression of information about the involvement statistical agencies, their data files, and their staff in such misuse. This is particularly true when major human rights abuses take place since perpetrators frequently take wide-ranging overt actions to hide or disguise their operations.

The case of the forced migration and internment of Japanese Americans on the West Coast, discussed earlier, represents a very mixed picture when it comes to barriers to the study of disclosure. Over-all, the treatment of the Japanese American after Pearl Harbor is one of the most thoroughly studied human rights abuses based, in part, on contemporary and continuing research and a very rich archival record. However, with respect to the activities and outputs of the Census Bureau there are several major lacuna. For example, the archival files maintained by the National Archives and Records Administration for the Census Bureau (Record Group 29) contain relatively little information on Dedrick’s operations on the West Coast and the Census Bureau’s supporting activities, although numerous materials can be found elsewhere. Interestingly, among the relevant materials that Record Group 29 does contain are two notes by Dedrick one indicating with respect to Box 219, “nothing of permanent archival value in this box. To be destroyed” [Dedrick, no date] and the other “I perceive *no need* [emphasis in the original] to keep any of the materials in Box 222” [Dedrick, 1973]. While both of these notes were later over-ridden by others, they do indicate a state of mind and might well help to explain why almost nothing pertaining to his work as the West Coast representative for the Census

Bureau between February and June 1942 were retained among Dedrick's other papers stored in Record Group 29.

A similar situation seemed to play out on the West Coast. As already described, Dedrick established and headed the Statistical Branch of the Wartime Civil Control Administration (WCCA). The WCCA was under direct control of the military and was operationally responsible for the actual expulsion of the Japanese American population from their homes and farms and maintaining them in temporary assembly centers pending their further relocation to more permanent camps. The latter camps were administered by the War Relocation Authority (WRA), an authority under civilian control. The WRA also had a statistical unit and in the words of the WRA official who at the end of the war had over-all responsibility for this unit,

The WCCA maintained an elaborate statistical division which played an important part not only in planning and executing the evacuation but also in directing the subsequent movement to relocation centers. Population accounting procedures established ... [by the WCCA] were ... continued by ... [the WRA]. The San Francisco offices of the two organizations were on adjacent floors of the same building, and information from the Statistical Division of the WCCA was available to the WRA. There appears to have been a more or less implicit understanding by the WRA that many, if not all, of the basic evacuee records of the WCCA, particularly those relating to number, location, identification, and movement of evacuees would be transferred to the WRA ... [Stauber, 1946: 15]

For what ever reason, Dedrick's Statistics Division decided to withhold these materials as Stauber indicates:

In the fall of 1942, about the time the transfer of evacuees from [temporary] Assembly Centers to [more permanent] Relocation Centers was approaching completion, it became apparent, that contrary to expectations, the WCCA basic records and files would not be transferred to the WRA [Stauber, 1946:17].

In any case, as previously mentioned, a number of key documents related to Dedrick's and the Census Bureau's involvement in the exclusion and internment have not yet been found. These include, for example, a long letter Dedrick wrote to Capt about his work on the West Coast dated 3/19/42. (A copy of Capt's telegraphed response was recently located in Commerce Department files [Capt, 1942].) The detailed urban block level 1940 Census tabulations and block maps with information from these tabulations shown on the maps have also not been located.

V. Conclusions and Recommendations

As stated at the outset, this paper has attempted to do two things. First, it presents a reconceptualization of a number of the issues related to the consideration of data access and statistical confidentiality, particularly with respect to data and information gathered by, or in the name of, government statistical programs. This reconceptualization, is not an arbitrary one,

rather it is derived from the detailed historical review of the ethical, statistical policy, and statutory background of the concept of statistical confidentiality presented in section I of the paper. The implications of this revised perspective might be summarized in these terms: any consideration of data access needs to take into account the nature of the consequences that such access might bring, particularly consequences that might threaten the life, liberty, and property of individuals and vulnerable population subgroups. Further, our responsibilities in this regard are defined not only by law, but also by a body of well-articulated statistical policy and ethics.

The paper also presents a review of some 15 “incidents” where population data systems have been used to target individuals or population subgroups, where such efforts were initiated, or where such targeting has been seriously contemplated as well as a brief discussion of other instances of possible targeting, ranging from the probable to the speculative. Collectively, they suggest that such incidents are comparatively rare, but when they have occurred, their impact has sometimes been catastrophic. Moreover, until the topic is more broadly and thoroughly researched, any attempt to quantify the expression “comparatively rare” would be unjustified.

A number of qualitative or policy conclusions do seem warranted, which we shall express in terms of four summary recommendations. In one form or another, the first three of these recommendations have been made previously [Seltzer, 1998: 544-545; Seltzer and Anderson, 2000: 40-43; 2001: 500-506], so it will be unnecessary to elaborate on them in detail. The final one is new.

1. Safeguards and prevention strategies

We recommend, as we have done previously, that active attention be given to developing and maintaining a robust and diverse system of safeguards against misuse together with a related prevention strategy. As we set out in more detail in Seltzer and Anderson [2001: 495-500], such a system includes attention to substantive safeguards, methodological and technological safeguards, organizational and operational safeguards, legal safeguards, and ethical safeguards. Over the short run, one or another of these safeguards may weaken or fail. However, over the longer-term and collectively, they appear to minimize the harm that is likely to flow from efforts designed to use population data systems to target individuals and vulnerable population subgroups. Another way of considering prevention and safeguards strategies is to look for those factors that appear in the past to have been associated with misuse or near misuse. Table 3 presents such an assessment, which lists in terms of critical and aggregating factors those situations that seem to increase the possibility for respondent or group harm.

[Table 3 about here]

2. Training

Although not specifically mentioned in the listing of safeguards, training is certainly an important element of any prevention strategy. However, such training will need to go beyond simply the technical aspects of disclosure prevention or the related legal and regulatory requirements for disclosure, as important as such training may be, to cover a broader range of issues, prevention strategies, and safeguards.

3. Further Research

Related to training is the need for further research. This includes further research on known incidents of targeting or attempted to clarify key uncertainties, research into possible additional incidents, several of which have been mentioned earlier in the paper, meta research into the phenomenon of targeting, and finally research into specific methods and approaches for improving safeguards.

4. The Establishment of an Incident Register

Our final recommendation takes up an idea mentioned earlier in the paper when we referred to one shortcoming of Table 1, noting that it excluded

what we suspect are numerous unreported incidents were government statistical agencies in the United States and elsewhere have quietly and successfully resisted explicit efforts to use the statistical system to target individuals and vulnerable population subgroups

and suggested that a registry of such incidents be developed. Such a register could serve both as a deterrent against efforts aimed at undermining statistical confidentiality and as a record of the success of federal statistical agencies in resisting efforts to compromise their responsibilities to data providers. Such a register might be organized on an agency basis, with a master register maintained by the Office of Statistical Policy, and with some provision for periodic public accountability, either through an independent intermediary, such as CNSTAT or the GAO, or by releasing the information directly to the public. The existence of such a register might not have prevented the gravest of the incidents listed in Table 1, but it might have helped to deter some of those that took place in democratic and open societies.

REFERENCES

- Aly, Götz and Karl Heinz Roth. 1984. *Die restlose Erfassung: Volkszählen, Identifizieren, Aussondern im Nationalsozialismus* [Complete Registration: Population Census, Identity, and Separation under National Socialism]. Berlin: Rotbuch Verlag.
- 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.
- _____. 1980. Letter to Raymond Okamura. December 16, 1980. Quoted in part in Okamura, Raymond Y. 1981. "The myth of census confidentiality." *Amerasia Journal* 8(2): 111-120.
- Barron, William G., Jr. 2002. Email to W. Seltzer, 1/8/02.
- Begeer, W. and W.F.M. de Vries. 1987. "Public confidence in official statistics." *Netherlands Official Statistics* 2(3).
- Black, Edwin. 2001. *IBM and the Holocaust: The Strategic Alliance Between Nazi Germany and America's Most Powerful Corporation*. New York: Crown, 2001.
- Blum, Alain. 2000. "Stalinism and Population Statistics." Paper presented at the Annual Meeting of the Population Association of America, Los Angeles, CA.
- Bohme, Frederick G. 1975. Letter to Roger Daniels, dated December 3, 1975. Records relating to the War History Project (item 156); US Census Bureau, Record Group 29; National Archives Building, Washington, DC.
- Bohme, Frederick G. and David N. Pemberton. 1991. "Privacy and Confidentiality in the US Censuses -- A History." Paper prepared for presentation at the annual meeting of the American Statistical Association, August 18-22, Atlanta, GA.
- Bryant, Barbara Everitt and William Dunn. 1995. *Moving Power and Money: The Politics of Census Taking*. Ithaca (NY): New Strategist Publications.
- Caplan, Arthur L. 1992. "The Doctor's Trial and Analogies to the Holocaust in Contemporary Bioethical Debates." In Annas, George J., and Michael Grodin, Eds. *The Nazi Doctors and the Nuremberg Code: Human Rights in Human Experimentation*. New York: Oxford University Press, pp. 258-275.
- Capt, J.C. 1942. Telegram to Dedrick, Hotel Whitcomb, San Francisco, 3/21/1942. NARA, RG40, Commerce, Entry 1, Box 144, File no. 67104.
- Census Advisory Committee (CAC). January 1942. Full transcript of meeting of January 9-10. Census Advisory Committee Meeting, January 9 and 10, 1942; General Records of the Census

Advisory Committee (item 148); US Census Bureau, Record Group 29; National Archives Building, Washington, DC.

Clark, Tom C. 1972. Interview for the Earl Warren Oral History Project. The Bancroft Library, University of California, Berkeley.

Clausen, John A, Herbert S. Conrad, William W. K. Freeman, Nicholas Hobbs, John W. Hopkins, Alfred N. Watson, and Rensis Likert. 1954. "Report of the Ad Hoc Committee on Statistical Standards." *The American Statistician*, Vol. 8, No. 3. (Jun. - Jul), pp. 19-23.

Corcoran, Thomas F. 1963. "On the Confidential Status of Census Reports." *The American Statistician*, Vol. 17, No. 3. (Jun.), pp. 33-40.

Dedrick, Calvert L. 1942. "Japanese Population of Bainbridge Island." Memorandum to Col. Bendetsen, 3/21/1942. Bancroft Library, Banc MSS 67/14c FILM, Reel 12, frames 213 and 215; also in NARA, RG338, WCCA, Unclassified correspondence, Box 15, File 093 "Evacuation of Bainbridge Island."

_____. 1973. Hand written note, dated 11/29/1973. General records maintained by Calvert Dedrick, 1935-42 (item 210); US Census Bureau, Record Group 29; National Archives Building, Washington, DC.

_____. 1981. Testimony of Calvert Dedrick before the Commission on Wartime Relocation and Internment of Civilians, November 3, 1981, NARA, CWRIC Hearings Transcripts, pp. 180-181.

_____. No date. Hand written note, undated. General records maintained by Calvert Dedrick, 1935-42 (item 210); US Census Bureau, Record Group 29; National Archives Building, Washington, DC.

DeRosier, Arthur H. Jr. 1970. *The Removal of the Choctaw Indians*. Knoxville: The University of Tennessee Press.

Duncan, G. T., T. B. Jabine, and V. A. de Wolf, eds. 1993. *Private Lives and Public Policies: Confidentiality and Access of Government Statistics*. Panel on Confidentiality and Data Access, Committee on National Statistics, National Research Council and the Social Science Research Council. Washington, DC: National Academy Press.

Duncan, Joseph W., and William C. Shelton. 1978. *Revolution in United States Government Statistics, 1926-1976*. Office of Federal Statistical Policy and Standards. Washington, DC: U.S. Department of Commerce.

Dunn, Edgar S., Jr. 1967. "The Idea of a National Data Center and the Issue of Personal Privacy." *The American Statistician*, Vol. 21, No. 1. (February), pp. 21-27.

- Ellenberg, Jonas H. 1983. "Ethical Guidelines for Statistical Practice: A Historical Perspective." *The American Statistician*, Vol. 37, No. 1. (Feb.), pp. 1-4.
- Elliot, Charles P. 1948. "An Indian reservation under General George Crook." *Military Affairs* 12 (2, Summer): 91-102.
- Fay, Robert G. 2001. "Evidence of Additional Erroneous Enumerations from the Person Duplication Study." Executive Steering Committee on Accuracy and Coverage Policy II (ESCAP II), Report 9. Preliminary version, Oct. 26, 2001
- des Forges, A. *Leave None to Tell the Story*. New York: Human Rights Watch, 1999.
- Freeman, William W. K. . 1952. ["The Statistician and His Conscience"]: Discussion (in Standards of Statistical Conduct in Business and Government). *The American Statistician*, Vol. 6, No. 1. (Feb., 1952), pp. 18-20.
- Goldberg, Joseph P., and William T. Moyer. 1985. *The First Hundred Years of the Bureau of Labor Statistics*. Washington, DC: US Government Printing Office.
- Grodzins, Morton. 1949. *Americans betrayed: politics and the Japanese evacuation*. Chicago: University of Chicago Press.
- Holley, Robert H. 1938. "Confidential nature of individual data returned on population schedules." Memorandum to the Director, dated 9/8/1938. NARA, RG29, Entry 274, Box 56 (Holley), Folder: Austin, W.L.
- Horrell, Muriel. 1959. *Race Classification in South Africa -- Its Effects on Human Beings*. Johannesburg : South African Institute of Race Relations.
- 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>.
- Ireland, C. Terrence, Lee-Ann Hayek, Tore Dalenius, Sidney Hollander, Charles R. Mann, Eli S. Marks, Janace S. Pierce, Frederick J. Scheuren, and William Seltzer. 1983. "Ethical Guidelines for Statistical Practice: Report of the Ad Hoc Committee on Professional Ethics." *The American Statistician*, Vol. 37, No. 1. (Feb.), pp. 5-6.
- Kraly, Ellen Percy, and John McQuilton. 2002. "The 'Protection' of Aborigines in Colonial and Early Federation Australia: The Role of Population Data Systems." Paper prepared for the First International Conference on Population Geographies, University of St. Andrews, St. Andrews, Scotland
- Jones, James H. 1981. *Bad Blood: The Tuskegee Syphilis Experiment*. New York: The Free Press.

Lie, Einar. 2002. "Numbering the Nationalities: Ethnic Minorities in Norwegian Population Censuses 1845-1930." *Ethnic and Racial Studies*, Vol. 25, No. 5 (September): 802-822

Mamdani, Mahmood. 2001. *When Victims Become Killers: Colonialism, Nativism, and the Genocide in Rwanda*. Princeton (NJ): Princeton University Press.

Methorst, H.W. 1936. "The New System of Population Accounting in the Netherlands." *Journal of the American Statistical Association* 31 (196): 719-722.

_____. 1938. "The New System of Population Accounting in the Netherlands." *Journal of the American Statistical Association* 33 (204): 713-714.

Mitchell, William D. 1930. Letter from the Attorney General to the Secretary of Commerce, dated 9/29/1930. NARA, RG40, General Records of the Department of Commerce, Office of the General Council, Subject and Index Files, 1903-1947, Box 157 and RG29, Entry 142, Box 3 Austin, Folder: Attorney General, Department of Justice.

Moore, Bob. 1997. *Victims and Survivors: The Nazi Persecution of the Jews in the Netherlands, 1940-1945*. London: Arnold.

Mule, Thomas. 2001. "ESCAP II: Person Duplication in Census 2000." Executive Steering Committee on Accuracy and Coverage Policy II (ESCAP II), Report No. 20. Washington, DC: U.S. Census Bureau, October 11, 2001.

National Research Council. 2001. *Principles and Practices for a Federal Statistical Agency*. Second Edition. Margaret E. Martin, Miron L. Straf, and Constance F. Citro, eds. Committee on National Statistics. Washington, DC: National Academy Press.

Office of Management and Budget. 1997. "Order Providing for the Confidentiality of Statistical Information." *Federal Register* 62:124 (June 27) 35044-35050, accessed 5/30/02 from <www.whitehouse.gov/OMB/inforeg/statpol.html>.*

Presser, J. 1969. *The Destruction of the Dutch Jews*. New York: E.P. Dutton and Co. Originally published in Dutch in 1965 as *Ondergang*.

Rémond, René, et al. *Le 'Fichier Juif' : Rapport de la Commission présidée par René Rémond au Premier Ministre*. Paris: Plon éditeur, 1996.

Rice, Stuart. 1941. Letter to W.F. Ogburn, dated 2/6/1941. University of Chicago Library, Special Collections, W.F. Ogburn papers. Box 27, Census Advisory Committee; folder 5, Bureau of the Census, correspondence, 1941.

Rogers, Samuel. 1917a. Letter to the Secretary of Commerce, dated 6/22/1917. NARA, RG40, General Records of the Department of Commerce, Office of the General Council, Subject and Index Files, 1903-1947, Box 157.

_____. 1917b. Letter to the Solicitor, Department of Commerce, dated 6/28/1917. NARA, RG40, General Records of the Department of Commerce, Office of the General Council, Subject and Index Files, 1903-1947, Box 157, file 3400-124.

Rowe, James, Jr. 1941. Memorandum to FDR, dated FDR Library, James Rowe Jr. Collection, Box #5, folder "Census Bureau."

Seltzer, William. 1994. "Politics and Statistics: Independence, dependence, or interaction?" Working Paper No. 6, United Nations Department of Economic and Social Information and Policy Analysis: New York.

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

_____. 1999. "Excluding Indians Not Taxed: Federal Censuses and Native-American in the 19th Century." Paper presented at the 1999 Joint Statistical Meetings, Baltimore, MD, August 1999. A shorter version appears in *1999 Proceedings of the Government and Social Statistics Section*. Alexandria, VA: American Statistical Association, 161-166.*

_____. 2001. "U.S. Federal Statistics and Statistical Ethics: The Role of the American Statistical Association's Ethical Guidelines for Statistical Practice." Paper based on presentation to the Washington Statistical Society, February 2001. Draft of September 17, 2001.*

Seltzer, William and Margo Anderson. 2000. "After Pearl Harbor: The Proper Role of Population Statistics in Time of War." Paper presented at the Annual Meeting of the Population Association of America, Los Angeles, CA, March 2000.*

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

_____. 2002. "NCES and the Patriot Act: An Early Appraisal of Facts and Issues." Paper prepared for presentation at the annual Joint Statistical Meetings, New York, August 10-15, 2002.*

Seltzer, William and Melissa McNeil. In progress. "Institutionalizing Apartheid: the Role of the South African Department of Census and Statistics, 1950-1959."

South Africa (Union of). 1950. "Population Registration Act." Pp. Sections 1-24 in Act No. 30 of 1950.

South Africa (Union of). Bureau of Census and Statistics. [1960]. *Union Statistics for Fifty Years, 1910-1960: Jubilee Issue*. Pretoria: Government Printer.

South Africa (Union of) . Bureau of Census and Statistics. No date. "Annual report of the Statistics Council, 1959 incorporating the annual report of the Director of Census and Statistics." Pretoria: the Government Printer.

South Africa (Union of). Parliament. House of Assembly. 1950a. *Debates of the House of Assembly. (Hansard). Third Session - Tenth Parliament. 20th January to 24th June 1950.* Cape Town: Unie-Volkspers Bpk., Parliamentary Printers.

_____. 1950b. *Report. Select Committee on the Population Registration Bill.* Chairman T. E. Donges. Cape Town: Cape Times Ltd.

South Africa (Union of). Parliament. The Senate of South Africa. 1950. *Debates (Official Report). Third Session - Tenth Parliament (Fifth Senate). 1950.* Cape Town: Unie-Volkspers Bpk.

Søbye, Espen. 1998. "Et moerkt kapittel i statistikkens historie?" [A dark chapter in the history of statistics?]. *Samfunnspeilet* 8, no.4 (10 September). Oslo: Statistics Norway.

Stauber, Ralph. 1946. "Final Report of the Relocation Planning Division." War Relocation Authority. Typewritten report, dated April 26, 1946. Bancroft Library, Banc MSS 67/14c FILM, Reel 26, frames 50ff.

Steuart, William. 1930. Memorandum to the Secretary of Commerce, dated 7/8/1930. NARA, RG40, General Records of the Department of Commerce, Office of the General Council, Subject and Index Files, 1903-1947, Box 157.

Thomas, Dorothy S. 1937. Review of *De Bevolkingsboekhouding* by J.L. Lentz. *American Sociological Review* 2 (1) : 117-118.

U.S. Census Bureau. 1942a. "Exhibit A - Work done on reimbursable basis by Bureau of the Census for war or other agencies charged with execution of some phase of the war program. December 1941 through November 1942." December 4, 1942. Mimeographed. War Work Reports folder; National Defense; Papers of Philip Hauser (item 146); US Census Bureau, Record Group 29; National Archives Building, Washington, DC.

_____. 1942b. U.S Bureau of the Census, Sixteenth Census of the United States, 1940, Population and Housing. Statistics for Census Tracts, Los Angeles-Long Beach, Calif. Washington DC, GPO: 1942.

_____. 1982. "Statement on Census Bureau Actions at the Outset of World War II as Reported in Infamy: Pearl Harbor and its Aftermath by John Toland." October 1982. [Attachment to letter from Bruce Chapman to Raymond Okamura, dated October 21, 1982.] CWRIC microfilm.

_____. 2001. "The Public's Privacy – No idle pledge!." Washington, DC: U.S. Census Bureau, November 2001, downloaded from www.census.gov on 10/25/01.

U.S. Congress. Senate. Committee on Military Affairs. 1942. "S. 2352." Unprinted hearing report prepared by Ward and Paul, official reporters, 77th Congress, 2nd Session, March 13 1942, pp.8-9. Bancroft Library, Banc MSS 67/14c FILM, Reel 4, frame 69.

U.S. Department of Commerce, 1917. "IN RE authority of the Director of the Census to furnish registration officials with names and ages of certain persons from census records." Opinion of the Acting Solicitor, Office of the Solicitor, Department of Commerce, dated 6/25/1917. NARA, RG40, General Records of the Department of Commerce, Office of the General Council, Subject and Index Files, 1903-1947, Box 157, file 3400-124.

Wartime Civil Control Administration. 1942a. "Japanese Population and Number of Japanese Family Heads in Arizona, by Minor Civil Division: 1940." Bulletin No. 2. Western Defense Command and Fourth Army, Wartime Civil Control Administration, Statistical Section, 3/16/1942. Bancroft Library, Banc MSS 67/14c FILM, Reel 12, frames 543ff.

Wartime Civil Control Administration. 1942b. "Japanese Population and Number of Japanese Family Heads in California, by Minor Civil Division: 1940." Bulletin No. 3. Western Defense Command and Fourth Army, Wartime Civil Control Administration, Statistical Section, 3/16/1942. Bancroft Library, Banc MSS 67/14c FILM, Reel 12, frames 544ff.

Wartime Civil Control Administration. 1942c. "Japanese Population and Number of Japanese Family Heads in Washington, by Minor Civil Division: 1940." Bulletin No. 4. Western Defense Command and Fourth Army, Wartime Civil Control Administration, Statistical Section, 3/16/1942. Bancroft Library, Banc MSS 67/14c FILM, Reel 12, frames 552ff.

Wartime Civil Control Administration. 1942d. "Japanese Population and Number of Japanese Family Heads in Oregon, by Minor Civil Division: 1940." Bulletin No. 5. Western Defense Command and Fourth Army, Wartime Civil Control Administration, Statistical Section, 3/16/1942. Bancroft Library, Banc MSS 67/14c FILM, Reel 12, frames 557ff.

_____. 1943. "Section III Miscellaneous Statistics." In Bulletin 12. Western Defense Command and Fourth Army, Wartime Civil Control Administration, Statistical Division. 3/15/1943. Bancroft Library, Banc MSS 67/14c FILM, Reel 13, frames 151ff.

Wietog, Jutta. 2001. *Volkszählungen unter dem Nationalsozialismus. Eine Dokumentation zur Bevölkerungsstatistik im Dritten Reich*. Berlin: Duncker and Humblot.

* Paper or link available at <<http://www.uwm.edu/~margo/govstat/integrity.htm>>.

Table 1. List of Cases Where Population Data Systems Have Been Used to Target Individuals or Population Subgroups, Where Such Efforts Were Initiated, or Where Such Targeting Has Been Seriously Contemplated

[The time periods and intended targets specified refer only to those studied in the sources cited.]

<i>Place</i>	<i>Time period</i>	<i>Targeted individuals or groups</i>	<i>Data systems involved</i>	<i>Type of data</i>	<i>State objective</i>	<i>Source</i>
Australia	19 th and early 20 th centuries	Aborigines	Population registration	Micro	Forced migration, elements of genocide	Kraly and McQuilton, 2002
France	1940 - 1944	Jews	Population registration, special censuses	Macro, micro	Forced migration, genocide	Rémond, 1996; Seltzer, 1998
Germany	1933 - 1945	Jews, Roma, and others	Numerous	Macro, micro	Forced migration, genocide	Seltzer, 1998
Netherlands	1940 - 1944	Jews and Roma	Population registration system	Macro, meso, micro	Forced migration, genocide	Seltzer, 1998
Norway	1845-1930	Samis and Kvens	Population censuses	Macro, micro	Ethnic cleansing	Lie, 2002
Norway	1942 - 1944	Jews	Special census and proposed population registration	Macro, micro	Genocide	Seltzer, 1998; Søbbye, 1998
Poland	1939 - 1943	Jews	Primarily special censuses	Macro, micro	Genocide	Seltzer, 1998
Romania	1941-1943	Jews and Roma	1941 Population Census	Macro, micro	Forced migration, genocide	Black, 2001
Rwanda	1994	Tutsi	Population registration	Macro, micro	Genocide	des Forges, 1999

South Africa	1950-1993	African and "Colored" populations	1951 Population Census and population register	Micro	Apartheid, voter disenfranchisement	Seltzer and McNeil, in progress
United States	19 th century	Native Americans	Special censuses, population registers	Macro, micro	Forced migration	Seltzer, 1999
United States	1917	Suspected draft law violators	1910 Census	Micro	Investigation and prosecution of those avoiding registration	See section, II.A.4 below
United States	1941 - 1945	Japanese Americans	1940 Census	Macro, meso, micro	Forced migration and internment	Seltzer and Anderson, 2000; in progress
United States	2001-2002 (?)	Suspected terrorists	Surveys and administrative data gathered by the National Center for Education Statistics	Micro	Investigation and prosecution of domestic and international terrorists	Seltzer and Anderson, 2002
USSR	1919 - 1939	Minority populations	Various population censuses	Macro, micro	Forced migration, punishment of other serious crimes	Blum, 2000

Table 2. *Were Micro-records from the 1940 Census Actually Used in the Forced Evacuation and Internment Process? A Summary of the Conflicting Evidence*

Yes, 1940 Census *micro* data were so used

FDR's willingness in 1939 to ignore the confidentiality provisions of Title 13 on a selective basis.

Legislative efforts in 1939 by the Departments of Justice, War and Navy to relax the confidentiality provisions of Title 13.

Census Director Capt's willingness, expressed in the January 1942 Census Advisory Committee meeting that "If the defense authorities ... wanted the names of [missing] Japs ... I would give them further means of checking individuals"

Capt's persistent efforts, beginning in May 1941 just after becoming Census Director and continuing through November 1941, to revive the 1939 legislative effort; his efforts in December 1941 just after Pearl Harbor to achieve the same ends by Executive Order; and finally his successful effort, in early 1942, to have an amendment inserted in the Second War Powers bill to permit relaxing of the confidentiality provisions of Title 13.

NY Times story of 2/7/1941 about amending the Second War Powers bill that includes very suggestive language, such as, "Some agencies of the government want data now as a matter of national safety. They seek some of the information obtained particularly from Japanese and others who since have become enemy aliens, especially about those in costal areas from which they have been ordered evacuated by the Department of Justice ... [Such] data, now a secret under law, government officers believe, would be of material aid in mopping up those who had eluded the general evacuation orders."

Senator Warren Austin's (R- Vermont) statement, during a March 13, 1942 hearing of the Senate Committee on Military Affairs, that "we are now considering ... legislation that would empower the Secretary of State to requisition from the Commerce Department, Bureau of the Census, what the census shows about these people. That would give an enumeration of the Japanese and it would also give names and residences, so that, when the Army makes its evacuation it can ... compare its list of evacuees against the census and have some knowledge of whether this has been an effective protection or not."

Capt's March 21 telegram to Dedrick in San Francisco instructing him to carry out his "duties without being hampered by old Bureau habits, precedents, and practices that are not in complete accord with the urgent, rapidly shifting necessities of the times as they develop from hour to hour."

Western Defense Command (WDC) statements that Dedrick's office in San Francisco had "access to the census records of 1940" and specifically mentioning other kinds of identifiable *micro* data

Several references to special name searches in WDC documents addressed to Dedrick in the second half of 1942, for example a memorandum from Bendetsen to Dedrick, dated 7/3/1942, "Location of certain persons of Japanese ancestry," (although none of these documents specifically mentions the 1940 census).

No, 1940 Census *micro* data were not so used

No explicit evidence yet found of actual use in the various NARA record groups, Presidential libraries, and other archives so far examined (although there are several suggestive pieces of evidence that efforts were made by Dedrick and possibly others to suppress information about what actually happened).

A number of Western Defense Command documents refer to the role of Dedrick's Statistics Section as providing statistics and maps to assist in planning the evacuation and not to tracking individuals for control purposes (although there was reference to furnishing agencies concerned with resettlement with "data pertaining to the needs of individuals and families, occupations and locations").

Given the *meso* data and maps provided, not clear the additional value of 1940 census *micro* data in the evacuation process (although the use of such *micro* data could have been used later in 1942 as part of an effort to account for "missing" Japanese Americans, as the NY Times story and the comments of Senator Austin indicate was the intended use of such census *micro* data).

Dedrick's strong opposition in 1939 to the proposed legislation designed to relax the confidentiality provisions of Title 13.

No record of any applications to waive the confidentiality provisions of Title 13 for individuals, as required by the 1942 law and the related Executive Order, has been found, (although if the proposed use was deemed not to be statistical, it is not clear that the waiver provision would have been invoked).

Table 3. Factors Contributing to Higher Risk of Population Data Collection Effort based on Potential for Respondent or Group Harm

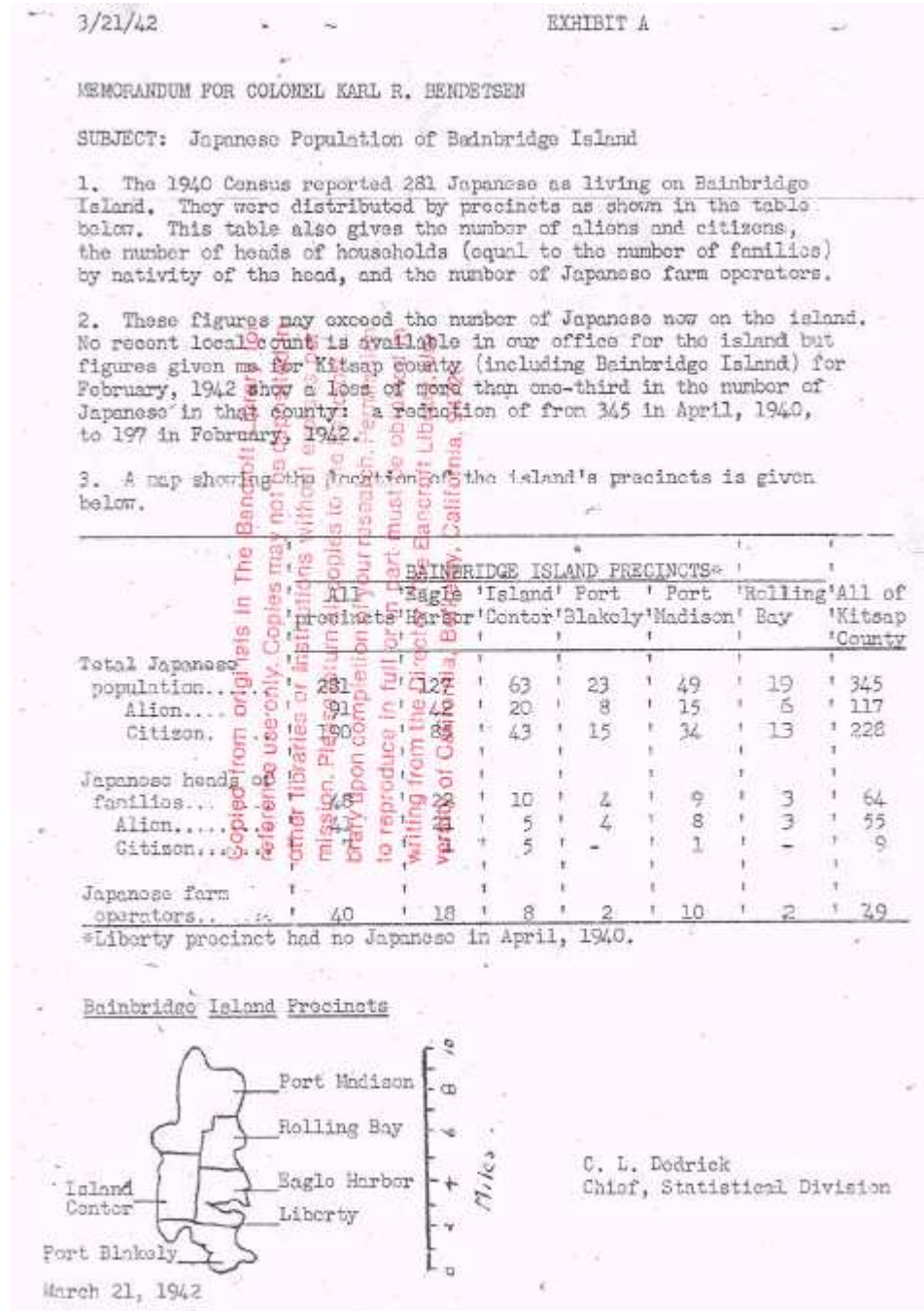
<p><u>A. Critical factors</u></p> <ol style="list-style-type: none">1. Population studied is weak or otherwise vulnerable.2. Data gathering or research involves variables that are on “sensitive” topics, typically topics that are or can be used to identify or stigmatize one or more vulnerable groups, or use classifications that permit the identification or stigmatization of such groups. <p><u>B. Aggravating factors</u></p> <ol style="list-style-type: none">1. All or substantially all of population is covered, i.e., sampling is not used.2. Longitudinal data gathering is involved, or the activity can be linked to a longitudinal system.3. Participation is mandatory or is effectively coerced.4. Little or no input from the subject population in planning the data gathering or research activities. (The risk potential is further enhanced if there are substantial inputs in terms of expertise, staff, or funds from foreign persons or institutions.)5. The data gathering or research is carried out in a war, a period of civil disruption, or during or shortly after a similar emergency.6. Little or no attention given to organizational, operational, methodological, and technological safeguards against the misuse of information obtained for non-statistical purposes.7. Confidentiality assurances provided to respondents have limited or no legal basis.8. Ethical reviews are not carried out, are perfunctory, or are heavily influenced by utilitarian considerations.

Note: The presence of either or both critical factors gives rise to a presumption of risk and each additional aggravating factor present further augments such risk. On the other hand, it should be emphasized that the presence of critical and aggravating factors does not mean that actual harm has occurred.

Source: Seltzer, William. 2003. “Data collection, Ethics Issues in.” In *Encyclopedia of Population*. Paul Demeny and Geoffrey McNicoll, eds. New York: Macmillan Reference USA, pp. 195-197.

Figure 1. Memo from C. L. Dedrick to Col. Bendetsen, 3/21/1942

[Bancroft Library, Banc MSS 67/14c FILM, Reel 12, frame 215. Reproduced with permission from the Bancroft Library, University of California, Berkeley, California. May not be further distributed or copied without written permission.]



ENDNOTES

1. Analogous arrangements seem to be in place in a number of other countries.
2. As expressed by Hippocrates, “abstain [with respect to treatment] from whatever is deleterious and mischievous” and “whatever, in connection with my professional practice ... I see or hear, in the life of men, which ought not to be spoken of abroad, I will not divulge, as reckoning that all such should be kept secret” [Accessed at <<http://classics.mit.edu/Hippocrates/hippooath.html>> on 9/16/2003]. While some other translations portray Hippocrates as using more direct language, they do not alter the substance of these two principles.
3. Indeed, the War Crimes Tribunal that tried the “Doctors’ Case” at Nuremberg moved beyond medical experimentation to include a nonmedical crime involving data collection, admittedly of an extreme form [Seltzer, 2001: 3-4].
4. Lie [2002: 805] makes a similar point about statistical confidentiality practices in 19th century Norway, “it appears that it was only individual data for taxation purposes that was not to be released ... In practice, those who did not pay taxes or who had a lower social standing were more poorly protected.”
5. According to Bohme and Pemberton [1991: 6], the provision of permitting the release of individual census records to governors and courts was not removed from the census legislation until 1976. However, after the 1919 Act, the provision of such information was not permitted if it was to be used to the detriment of the person to whom it related.
6. Austin reached 70 near the end of January and it required a Presidential Executive Order for him to remain as Director beyond January 31, 1941. Such Executive Orders were by no means exceptional. Indeed, both he and the federal statistical community had expected that he would be extended for an additional year so he could complete the main work of the 1940 Census [Rice, 1941].
7. Capt was an able administrator and very-well connected politically. As described in the words of one White House adviser to the President, “he is the man Harry Hopkins is to speak to you about, Harry brought him over [to the Census Bureau from the WPA] ... to handle the political patronage [as a Confidential Assistant to Austin] ... did a splendid job ... [so that] the Senators and the Congressmen are all for him ...”, but as the same adviser adds, “[u]nfortunately, he has no professional background or standing in his profession. He has absolutely no statistical background” [Rowe, 1941].
8. The act ended governmental control of the economy except in the case of commodities and products that were still in short supply or needed control during reconversion at home or abroad (e.g., rubber and sugar). The opening section of that law noted that “The Congress hereby declares that it is vital to a free economy and full production in the United States that all emergency controls and war powers under the Second War Powers Act be removed except in certain limited instances.” The language continued by detailing the limited circumstances in

which control could continue, with no further mention of Section 1402. [Seltzer and Anderson, 2002: 44]

9. This was accomplished by Section 401 (a) (6) of the Education Sciences Reform Act, which transfers

all the subsections (including subsections (a) through (c) [of the 1994 NCES Act]) and redesignating such subsections as subsections (c) through (e), respectively, at the end of section 183 of this Act.

Although the Patriot Act went unmentioned in PL 107-279, section (c) of the 1994 NCES Act was the provision inserted by the Patriot Act. As a result, section 183 (e) of the new law contained the substantive content of section 508 of the Patriot Act. PL 107-279 also extended the confidentiality protections accorded to information collected for statistical purposes by the 1994 NCES Act to research and evaluation purposes.

10. There is a degree of uncertainty as to whether the provisions of the E-Government Act prevail over the provisions of the Patriot Act now embedded in the Education Sciences Reform Act. The E-Government Act became law in mid-December 2002, while the Education Sciences Reform Act became law in early November 2002, so that under the normal rules of legislative construction, PL 107-347 (the E-Government Act), as a later expression of legislative intent, would prevail over PL 107-279. Moreover, while section 504 of Title V of PL 107-347 does discuss the effect of this law on other laws, there is no mention of the Patriot Act, PL 107-279, or any other law specifically related to the Department of Education or NCES. This would seem to imply, therefore, that Congress intended that Title V of the E-Government Act should prevail over all previous laws not specifically mentioned in section 504, including the Patriot Act and the Education Sciences Reform Act. Nevertheless, the final decision on the issue, if it ever came to a test, would be up to first the Attorney General and then the federal courts.

11. Subtitle B of Title V permits the sharing of identifiable *micro* data pertaining to businesses among designated federal statistical agencies exclusively for statistical purposes.

12. Although we exclude misuse or potential misuse attributable solely to *macro* data from the scope of this paper, it should be understood that several European countries have legal or regulatory barriers that prevent government programs, including statistical programs, from compiling or storing even *macro* data using definitions or classifications that permit the identification of potentially vulnerable population subgroups. Many of these laws and regulations seem to have been put in place as a reaction to the events of World War II.

13. The accounts presented in this paper concerning these three examples have been largely drawn from the cited sources. However, in each case some additional material is presented here. This is particularly so with respect to the role of the Census Bureau in the forcible removal of Japanese Americans from the West Coast and the related efforts of Census Bureau Director J.C. Capt to set aside the confidentiality protections of Title 13.

14. Although Dedrick's name is not cited in Bulletin 12, the double log quality control chart comparing the estimated and evacuated populations for each of the 108 exclusion orders [Wartime Civil Control Administration, 1943, p. 149] reproduced in Bulletin 12 is a clear indication that these estimates were his work.

15. On January 28, 1942, the Los Angeles Times reported that "slightly more than 800" of the "2200 Japanese living on Terminal Island" were aliens. ["Eviction of Jap Aliens Sought," Los Angeles Times, 1/28/1942, p. 1A. See also Los Angeles Times, 1/29/1942, p. 6A]. Census Tract no. 294 encompassed the portion of Terminal Island in the City of Los Angeles. The other half of the island is in the City of Long Beach. Table A-1 reported a total population of Tract 294 of 3,831 people, of these 2,253 were "Nonwhite, Other races." The white population was reported in native born and foreign born categories, but the Nonwhite was not [U.S. Census Bureau, 1942b]. To our knowledge, the Census Bureau never published a breakdown of the detailed subcategories of "Nonwhite, Other races" or the Nonwhite population by nativity.

16. Our further research has provided additional support to Field's assertion that he was trying to obtain data from the 1940 Census on the Japanese American population prior to December 7, 1941 as part of an intelligence operation conducted for the White House. However, none of his efforts appeared to relate to identified *micro* data and we continue to remain highly skeptical of his later published accounts that he obtained such *micro* data.

17. In the 1940 Census, outside of major cities, most minor civil divisions fully met the criteria used in defining *meso* data, that is a unit small enough that it was operationally useful in targeting individuals or population sub-groups.

18. Capt was well-known to Chairman Sumners. Both were from Texas, and a year earlier, Sumners had himself written to Jones, urging that Capt be appointed as Census Director. [Letter to Commerce Secretary Jones, dated 2/8/1941, cited in acknowledgment from Baxter to Sumners, 2/12/41, NARA, RG 40, Entry 12, Subject files of Undersecretary of Commerce Wayne C. Taylor, folder "Census recommendations (for Director)"]

19. Given this lesson, we remain concerned about the continued existence of the Census Bureau's ESCAPII data files with machine-readable personal identifiers. As we previously noted [Seltzer and Anderson, 2002: 45-46]

as part of Census 2000 processing "for the first time personal identifiers have been captured from the forms" [Barron 2002; Fay, 2001: 2]. These computer files permitted the Bureau, as part of the 2000 Census evaluation program, to carry out a series of sophisticated case by case matches to all 280 million or so individuals enumerated in the 2000 Census [Mule, 2001].

While we understand the important role these files played in the evaluation program for the 2000 Census, as time passes the benefit to risk ratio associated with the further maintenance of these files, at least with 100 percent population coverage, continues to decline. Traditionally, it has

been Census Bureau policy that “names and addresses are separated from the electronic files that contain an individual's answers when no longer needed, to protect the respondent's confidentiality” [U.S. Census Bureau, 2001]. Disposing of these files would add to the difficulty of using identifiable *micro* data from the 2000 Census to target individuals and vulnerable population subgroups. Such an operational safeguard would by now seem a prudent complement to the important, but sometimes uncertain protection, provided by legal safeguards.

20. In fact, Bohme and Pemberton [1991: 10] indicate that the opinion involved was issued by the U.S. Solicitor General. We take the more conservative approach of referring to this as a Departmental opinion based on the archival records we have located.

21. The United States entered World War I on 4/6/1917, the Selective Service act became law on 5/18/1917, all men born 1886-1896 were required to register on 6/5/1917 (a special public holiday set aside for this purpose), and Census Director Rogers' request to the Secretary of Commerce was dated 6/22/1917.

22. A more extended version of the quote from President Taft in support of census confidentiality is provided by Bohme and Pemberton, 1991: 8. They note that same quote appeared in several subsequent census proclamations.

23. Bohme and Pemberton [1991: 10] give the 6/26/1917 as the date of this opinion; a file reference in NARA, RG40, General Records of the Department of Commerce, Office of the General Council, Subject and Index Files, 1903-1947, Box 157, file 3400-124 gives 6/25/1917 as the date. It is possible that the original opinion of the Commerce Department's Solicitor was confirmed the next day by the U.S. Solicitor General, although Census Director Rogers' letter of thanks only refers to the opinion issued by the Departmental Solicitor [Rogers, 1917b].

24. Holley [1938: 4] notes that one section of the 1910 Census Act that states “That no information furnished under the provision of the next proceeding section shall be used only for the statistical purposes for which it is supplied,” actually refers to data provided by various kinds of manufacturing and other establishments, and not population data.

25. Although the opinion of the Attorney General stated that the *micro* data wanted was to assist in literacy programs [Mitchell, 1930], Census Director Steuart in asking for a legal opinion on the matter did not cite any reason why the information was sought [Steuart, 1930]. Steuart did indicate that the information was sought by many local organizations and the Secretary of the Interior. The latter was both the senior government official responsible for federal education work and a former president of the American Medical Association. Steuart also indicated that it had been previously agreed to supply the requested micro data on illiteracy.

26. The material presented in this sub-section on South Africa is taken from an ongoing research study by the senior author and Melissa McNeil, formerly a graduate student at Fordham University.

27. During the course of its passage through Parliament the legislation was amended several times by the government in response to many criticisms that arose in the debate. However, the features listed here were in the measure when it was first introduced and remained after it was adopted.

28. For example, in research we have done relating the involvement of the Census Bureau in the internment of Japanese Americans and related efforts to relax the confidentiality provisions of title 13, important materials were found not only in the National Archives holdings for the Census Bureau and the Department of Commerce, but also among the legislative and executive branch and military records maintained by the National Archives, two presidential libraries and several academic institutions. While some material was found in multiple archives, much of it survived only in a single archive.