protecting Confidentiality in Archival Data Resources

Christopher S. Dunn
Erik W. Austin
Inter-university Consortium for Political and Social Research

Data sharing is a disputed norm in scientific affairs (Fienberg et al., 1985; Weil and Hollander, 1990; Fienberg, 1994; Mishkin, 1995). On the one hand, principal investigators argue that they and their research teams are the most competent analysts of originally collected data and best able to safeguard the data against release of confidential information. They know the details and nuances of the sampling procedures, instrumentation, data reduction, and missing data. They have an investment in the original research that should be repaid by first rights of publication. They also argue that for certain kinds of complex studies—for example, observational, organizational, longitudinal, and clinical research, and studies involving geocoded data or administrative records linked to survey data—they themselves are the only or the principal safeguard against violations of confidentiality.

On the other hand, researchers also argue that publicly supported data collections should be available to the public, or at least to competent researchers.Datasets can be purged of identifying information. Competent researchers can perform responsible secondary analyses of the data while simultaneously upholding the protection of confidentiality. The investment of public funds in data supersedes ownership rights, at least with respect to access to the data, as do the norms of science as an activity open to and dependent upon the scrutiny and review of other scientists.

Since 1962, ICPSR has been responsible for many of the technical and normative developments in social science data sharing. As an archive that acquires data from many principal investigators, ICPSR has undertaken the development and implementation of procedures that assure original investigators that the distribution and subsequent analyses of their data will not compromise the protection of confidentiality.

Over the past 36 years, ICPSR has responded to new technical challenges in preserving the confidentiality of data, while simultaneously charting a course that satisfies both proponents and opponents of data sharing, both data producers and data users.

Ways that confidentiality can be violated

There are four major ways that confidentiality can be violated, resulting in the release or deduction of individual identities and/or identifying characteristics:

- Accidental release may occur as a result of sloppy data management procedures, ignorance or errors on the part of staff, or failure to follow standard procedures. The idea of accidental release suggests that confidentiality is preserved by educating staff about the need for confidentiality protection procedures, training and monitoring staff in the application of those procedures, performing quality control checks on data files that are developed for restricted use or public release, and maintaining adequate security for confidential information.

- Malicious release may occur through theft or unauthorized transfer of data by disgruntled staff, by staff or others seeking financial gain, or through breaches of computer systems security. The central feature of malicious release is the idea that information (and hence data) has value and that there are people, whatever their motives, who may attempt to translate that value into financial gain or otherwise use the information inappropriately. Disgruntled staff, for example, may satisfy a symbolic urge for retaliation or retribution by unauthorized transfer or release of information. Regardless of whether the motive is instrumental or symbolic, the inappropriate, illegal behavior can be countered by deterrence and punishment. These dynamics suggest that organizations should have and employ policies that prohibit the unauthorized use, transfer, or release of data. In the case of public release, even though there is no restriction on who can access the available data, there should be use restrictions consistent with the research and educational purposes of the organization.

- Compulsory release may occur as the result of legal action or court order. This matter is too complicated and uncertain to be dealt with in an encapsulated discussion here. It is sufficient to note that the ethics of research are not the only requirements that researchers face and that the legal protection accorded the confidentiality of research data is not absolute or uniform across states or in different legal matters. Researchers have been ordered to release confidential data. Some have complied, others have refused and been penalized, still others have had initial orders overturned or modified on appeal. Again, the focus with this type of release seems to be on organizational policy against compulsory release that has as its basis the necessity of confidential-

The present emphasis on the biomedical roots of modern human subjects protection regulations and their original implementation in the Department of Health and Human Services obscures some important origins of the protection of confidentiality of records and data. Foundational documents of ethical principles of biomedical research are largely silent on issues of privacy and confidentiality. In contrast, mid-19th-century U.S. Census legislation required enumerators to keep information they collected in the course of conducting the Census confidential, principally as an instrumental means to promote subject cooperation and truthful response. The practice of maintaining confidentiality of Census data was extended to all Census employees in 1900. Gradually, professional associations adopted policies for the protection of confidentiality based on ethical principles such as safeguarding the privacy of individuals and minimizing potential harm to subjects through disclosure of sensitive information to third parties.

The confidentiality practices of the Census Bureau and the Bureau of Labor Statistics initiated in the late 19th and early 20th centuries anticipated two of the four major possibilities for failure to maintain confidentiality. The early statements about treating information in the Census as confidential recognized that individuals with legitimate access to microdata could also behave illegitimately by selling or transferring data to third parties. Industrial Census guidelines in 1910 and population and agriculture Census guidelines in 1930 recognized that individual or microlevel identities could be deduced from macrolevel tabular data with small cell sizes, thereby reflecting the first concerns about statistical disclosure.
ity in social research. Where possible, such policies should be backed up by regulatory or statutory nondisclosure protections, such as the DHHS certificates of confidentiality or the U.S. Department of Justice statutes (42 USC 3789g) and regulations (28 CFR 22) prohibiting evidentiary or other non-research uses of justice research data.

Statistical disclosure results from logical use or analysis of data to identify cases or events that are infrequent or rare, or unique patterns of characteristics which, when associated with data from other sources, lead to subject identification. The topic of statistical disclosure is also too complex to be discussed in detail here, but fortunately, there is more information available on this topic than on the others. Statistical disclosure has been the focus of both professional and academic attention. There are a variety of established methods for preventing disclosure (Cox et al., 1985; OMB, 1994). There is also developmental work in progress for devising and testing new methods (e.g., Duncan, undated; Dutta Chowdhury et al., undated). But the primary means for preventing statistical disclosure have in common a technical focus on the data themselves.

Assuring protection of confidentiality at ICPSR

The following sections borrow heavily from the ICPSR Guide to Social Science Data Preparation and Archiving.

Two kinds of variables often found in social science datasets present problems that could endanger the confidentiality of research subjects. Most familiar are the direct identifiers that may have been obtained in the process of data collection. These include items such as names, addresses (including ZIP codes), telephone numbers (including exchanges), Social Security numbers, and other linkable identification numbers such as driver license numbers, certification numbers, etc. Data collectors typically remove all such identifiers when preparing public use datasets. If datasets are received with such variables, ICPSR removes them as part of the lowest level of study processing. Increasingly, consideration is being given to returning to investigators datasets that are received with direct identifiers. This is because ICPSR practice is to preserve originally submitted data; legal action could be taken should it be known that ICPSR maintains a copy of such a dataset.

Another category of variables can often become problematic depending on the content of the data collection and the nature of the research subjects included in the dataset. These are indirect identifiers that might be used (in combination or in conjunction with publicly-available information) to identify individual respondents. This category is harder to deal with, since it includes items that are often the focus of or useful for statistical analysis. That is probably why such information was collected in the first place. Some examples of these indirect identifiers are detailed geography (e.g., state, county, or census tract of residence), organizations to which the respondent belongs, educational institution from which the respondent graduated (and year of graduation), exact occupations held, place where the respondent grew up, exact dates of events, detailed income, and offices or posts held by the respondent. Such indicators should be reviewed by the principal investigator/data collector and a judgment made about the effect of retaining such items upon the anonymity of the research subjects before depositing the data in a public archive.

Sometimes, variables usually considered to be indirect identifiers can become direct identifiers depending upon features of the research design. Job title or occupational role can directly identify a respondent when there is only one such position in an organization, one such organization in a town (or department in an organization), and the town (or organization) is identified, as well as the date of the data collection. For example, the police chief, Presbyterian minister, high school principal, or any other unique figure in a community or organization might identify his or her job title or occupational role, and the community or organization may be identified. If the date of the data collection is also known, it is easy to determine exactly who that person was at that time.

Handling indirect identifiers

If, in the judgment of the principal investigator, a variable might act as an indirect identifier (and thus could be used to compromise the privacy of a research subject), the investigator should "treat" that variable when preparing a public use dataset. Modifications commonly used are:

- **Removal**—eliminating the variable from the dataset entirely
- **Bracketing**—combining the categories of a variable
- **Top-coding**—grouping the upper range of a variable to eliminate outliers
- **Collapsing and/or combining variables**—merging the concepts embodied in two or more variables by creating a new summary variable

An example from a national survey of physicians (containing many details of each doctor's practice patterns, background, and personal characteristics) may help to illustrate each of these categories of treatment of variables to protect confidentiality. Variables identifying the school from which the medical degree was obtained and the year graduated should probably be removed entirely, due to the ubiquity of publicly available rosters of college and univer-
Some types of studies include variables that pose complex threats to confidentiality but are also difficult to modify because of their central importance to the study. One such type is the multilevel study consisting of hierarchical files with linkage variables between files. Another type is the study that contains exact event dates and birth dates. A third type is the study with geocoded information. A fourth is the qualitative narrative interview study. A fifth type, the longitudinal panel study, is not especially problematic when ready for archiving, but the need to maintain linkage and locator identifiers from one round to the next makes the study vulnerable to threats to confidentiality during its operational phases.

University graduates. The state of residence of the physician could be bracketed into a new “Region” variable (substituting more general geographic categories such as “East,” “South,” “Midwest,” and “West.”) The upper end of the range of the “physician’s income” variable could be top-coded (e.g., “$150,000 or more”) to avoid identifying the most highly paid individuals. Finally, a series of variables documenting the responding physician’s certification in several medical specialties could be collapsed to a summary indicator (with new categories such as “Surgery,” “Pediatrics,” “Internal Medicine,” “Two or more specialties,” etc.).

ICPSR staff consult with principal investigators to help them design or modify a public use dataset that maintains (to the maximum degree possible) the anonymity of respondents. The staff will additionally perform an independent confidentiality review of datasets submitted to the archive and will work with the investigators to resolve any remaining problems of confidentiality. The goal of this cooperative approach is to ensure that all reasonable steps have been taken to protect the privacy of research respondents whose information is contained in ICPSR’s public use datasets.

Research trends that pose problems for confidentiality

Multilevel studies, in which data are collected about places, organizations, households, persons, and events simultaneously are especially difficult to handle with the usual means of modifying variables. Often, information in the linked files makes it easy to identify individual subjects, but the linkage variables between files are essential to maintaining the multilevel value of the study. Where identification risks are high because the multiple levels of information might permit an analyst to narrow the focus on individuals, ICPSR will consider making the study a restricted-use dataset.

Studies with many precisely dated events and birth dates also pose risks to confidentiality, especially if the event information has been publicized in the media or recorded in publicly available administrative records (e.g., court dockets). Exact dates in the study information and event characteristics can be matched against media or administrative record data, allowing subjects to be easily identified. Nevertheless, the exact date information is often useful for various forms of time-dependent analyses like survival analysis or event history analysis. Removing exact dates reduces the value of the information. Once again, the solution may be to create a restricted dataset rather than removing information.

Studies with geocoded information are also problematic. Depending upon the nature of other information in the study and the degree of area resolution, geocoded studies may permit identification of individuals, especially when public information is available. For example, it would be inappropriate, unethical, and potentially dangerous to release a dataset with the address locations of rape victims.

With qualitative narrative interview studies, the level of detail provided is extensive and often contains many references to people, places, events, associations, organizations, family relationships, and so forth. Someone with intimate knowledge of these patterns of information may be able to identify the individuals involved. The very richness of the detailed information is simultaneously the value of the study and the threat to confidentiality. Original investigators are loathe to restrict the richness of the narratives, yet are unwilling to release such detailed information because of the potential threat to confidentiality.
Providing access to original indirect identifiers

It is rarely the case that variables removed or modified to maintain confidentiality are without value for research purposes. Archives and other data providers, therefore, frequently field requests for some form of access to original data values. Three of these forms of access will be discussed here:

**Customized data analysis** performed by the archive/data provider, often at cost to the researcher, affords the opportunity of obtaining analytic results from restricted variables. Typically, researchers will be asked to provide detailed analytic instructions—usually in the form of software commands—and the requested analyses are performed at the archive, with analytic output sent to the requesting party. At ICPSR and elsewhere, the output is examined by staff to ensure that the analysis results will not endanger the confidentiality of respondents.

Delivery of a **private-use dataset** allows original data values to be provided to a researcher, with the requestor explicitly assuming responsibility for maintaining confidentiality of those data. Most organizations that provide private-use datasets require a transaction form, replete with both researcher and official signatures certifying that such data will be securely held and will be used only by the requesting party in ways that protect the privacy of respondents.

A third mechanism bundles an entire dataset in an analytic software package that prevents examination of discrete values/cases while allowing statistical access to all variables. This **front-end software** alternative usually prevents extracting or downloading of original values on some or all variables. The National Center for Education Statistics’ Data Analysis System (DAS) is one example of such a software-based solution for protecting the confidentiality of research subjects. Other such front-ends are actively being explored, including at ICPSR.

Each of the mechanisms described above has advantages as well as drawbacks. None is completely satisfactory to both the research community and the repository/holder of original data. Tightest control of original data values is an attraction of the customized data analysis option, but this is the least popular with active researchers. It is typically costly (in terms of both time and money), and frequently thwarts the iterative analytic style most common in the social sciences. Private-use datasets permit the most researcher control of the analytic process, at the expense of certainty of protection of respondent confidentiality. Enforcement of private-use dataset provisions agreed to by requestors is difficult to effect, and sanctions against violators of promised assurances would inevitably involve a litigious voyage on mostly uncharted waters. Possibly the most secure yet flexible alternative is the front-end software option. Yet from the archive’s standpoint, this is probably the most expensive of the three alternatives; putting data into such a package is a time-consuming that it can practically be utilized on very few data collections. Furthermore, it is doubtful that front-end software is wholly impervious to hacking by a skilled and determined violator. Finally, the learning of “yet another” software package and its guaranteed limitations raises the bar over which interested researchers must jump to access needed research data.

**Licensing** a researcher to use a dataset containing indirect identifiers is a variant on the private-use dataset arrangement described above. A licensed use is agreed to after completion of a transaction form. Unlike private-use dataset agreements, however, most licenses impose an up-front fee in the form of a security bond as surety for maintaining confidentiality. The fee has been known to range from a few hundred to many thousands of dollars. Several license mechanisms also require the researcher and her/his institution to assume all legal liability in any instance of breaching confidentiality. Needless to say, the popularity of this form of “access-with-assurance” is quite low in the research community (not to mention in the college/university legal offices).

A second alternative method is the “perturbing” of original data values to break the certain bond between any given data value and the (possibly identifiable) individual who may have provided the initial information. Since the essence of this technique is the altering of original data values, it remains suspect in the minds of several generations of social scientists. These individuals find it difficult to overcome one legacy of their training—getting error out of research data collections—which clashes with the practice of introducing error into a dataset (however noble the purpose underlying that activity).

Perhaps more promising is the concept of **secure data analysis laboratories**. In such facilities, original data would be available for data analysis in a controlled setting, precluding such things as making copies of original data, investigating single cases, or transmitting the data offsite. Scholars would apply to visit the site to perform data analysis in the laboratory under secure conditions. An experiment using this form of access can be found at Carnegie Mellon University, for its National Consortium on Violence Research supported by the National Science Foundation and...
the National Institute of Justice. Data from the National Crime Victimization Survey, which have long been distributed without geographic sector information, are available with geographic information at Carnegie Mellon to the Violence Research Consortium members. This mechanism represents, for social scientists, a departure from a long-term trend of facilitating the export of research data from an archive or producer site directly to the institution (or desktop) of the interested scholar. It should be noted parenthetically that many research materials utilized by both historians and social scientists are available only by visiting the site where the research materials are housed. Included among such facilities are traditional archives and other repositories, including some fine social science collections like those of the Henry Murray Center at Radcliffe College. Undoubtedly more costly for the individual researcher (and perhaps for the archive as well), this mode of access to confidential data may become more common with heightened concern for preserving confidentiality.

The search for suitable mechanisms for protecting confidential microdata promises to become a high-stakes venture. At risk is the “big kahuna” of post-WWII social scientific research practice—readily available, empirical microdata. Some in the statistical and social science communities, as well as in government, are beginning to worry about the release of any microdata, with a few even predicting its demise.

**Conclusion**

The very progress of social science research methodology has made it more difficult to safeguard the confidentiality of the research data. Removing direct identifiers is a foundational requirement for public use datasets, but that is a comparatively easy task. More difficult tasks involve investigating which variables could be used as indirect identifiers and modifying them without significantly reducing the value of the data collection. Careful attention must be paid to interactions among the context of the study, the nature of the sample, and the characteristics of respondents to prevent ordinarily unrevealing information from becoming the pointer to an individual.

**References**


Additions to holdings

ABC News AIDS Poll, January 1996 — ABC News (ICPSR 6833)
ABC News Abortion/Cuba Poll, January 1998 — ABC News (ICPSR 2506)
ABC News Between-Debates Horserace Poll, October 1996 — ABC News (ICPSR 2165)
ABC News Bill Bradley Poll, August 1995 — ABC News (ICPSR 6675)
ABC News Clinton/Jones Poll, April 1998 — ABC News (ICPSR 2485)
ABC News Colin Powell Poll, September 1995 — ABC News (ICPSR 6676)
ABC News Japan Poll, April 1996 — ABC News (ICPSR 6819)
ABC News Listening to America Poll, May 1996 — ABC News (ICPSR 6820)
ABC News Marijuana Poll, May 1997 — ABC News (ICPSR 2489)
ABC News Media Poll, January 1997 — ABC News (ICPSR 2171)
ABC News New Jersey Governor Poll, October 1997 — ABC News (ICPSR 2503)
ABC News New York City Mayor Poll, Spring 1997 — ABC News (ICPSR 2498)
ABC News New York City Rent Control Poll, June 1997 — ABC News (ICPSR 2497)

ABC News “Nightline” Hillary Clinton Poll #1, January 1996 — ABC News (ICPSR 6829)
ABC News “Nightline” Hillary Clinton Poll #2, January 1996 — ABC News (ICPSR 6831)
ABC News “Nightline” Newt Gingrich and Ethics Poll, January 1997 — ABC News (ICPSR 2172)
ABC News Powell “No Go” Poll, November 1995 — ABC News (ICPSR 6682)
ABC News State of the Union Poll, January 1996 — ABC News (ICPSR 6832)
ABC News Timothy McVeigh Verdict Poll, June 1997 — ABC News (ICPSR 2491)
ABC News “20/20” Female Body Image Poll, March 1996 — ABC News (ICPSR 6818)
ABC News United States in Bosnia Poll, November 1995 — ABC News (ICPSR 6686)


Errata

In May 1998, a user reported a value label problem in the SAS and SPSS data definition statements of the Latino National Political Survey, 1989–1990 (ICPSR 6841) collection. The error exists for variable RCOUNTRY, which is found on page 55 of the codebook and refers to Question 39: “In what country were you born?” The value labels for codes 2 and 3 in the codebook as well as in the SAS and SPSS data definition statement files are incorrect. The correct labels are as follows:

code 2 is ‘Puerto Rico’
code 3 is ‘Mexico’

Additionally, the Latino National Political Survey, 1989–1990 collection is found on the Minority Research Data CD-ROM (CD0016, ICPSR 6941). All users of the data on CD0016 should take note of the above problems. ICPSR staff updated the Website version of ICPSR 6841 on May 22, 1998; the diskette version became available June 9, 1998.

Users of the World Values Survey, 1981–1984 and 1990–1993 (ICPSR 6160) should be aware that there are several special characteristics of the data file and its accompanying documentation. This item is listed in detail in the README file associated with the collection on the ICPSR Website. While ICPSR expects to receive an update to the collection in the future, users of the current file should make special note of the README file’s contents.
CD-ROMs NOW
AVAILABLE

A series of CD-ROMs containing National Corrections Reporting Program (NCRP) data has been developed by ICPSR with funding from the Bureau of Justice Statistics. These CDs can be ordered under ICPSR 2017 and cover the following years:
CD0017: 1983–1986 NCRP
CD0019: 1989–1990 NCRP
CD0020: 1991 NCRP
CD0021: 1992 NCRP
CD0022: 1993 NCRP
CD0023: 1994 NCRP
CD0025: 1995 NCRP

Also newly available is a CD-ROM providing access to data from the Record of American Democracy, 1984–1990 (see the May 1998 bulletin for more information on this product). The ROAD data CD-ROM (ICPSR 2162) includes 337 data files providing election returns, socioeconomic summaries, and demographic measures of the American public at low levels of aggregation.

ICPSR is also distributing the CD-ROM Twenty Years of Pollibarometers, 1977–1995 (ICPSR 2316), which includes East and West Pollibarometers from the Zentralarchiv in Koln, Germany.

And finally, a CD featuring 50 years of the American National Election Studies (1948–1997) was in production at press time. This CD provides all of the ANES time series and pilot studies, along with a new Java front-end to facilitate exploration of the data. Also included is the NES Guide to Public Opinion and Electoral Behavior, a collection of tables and graphs offering easy access to information on voting behavior. All data collection instruments for the election studies are available in PDF form on the CD.

Additions to holdings, continued


ABC News Whitewater Poll, December 1996 — ABC News (ICPSR 2168)


Availability and Use of Intermediate Sanctions by Judges and Corrections Professionals in the United States, 1994 — Helen G. Corrathers (ICPSR 6788)

Brazilian Survey on Nutrition and Health, 1989 — Instituto Nacional de Alimentacao e Nutricao (INAN), Ministerio da Saude. Instituto de Pesquisa Economica Aplicada (IPEA) (ICPSR 2294)

CBS News Monthly Poll #2, July 1996 — CBS News (ICPSR 2303)

CBS News Monthly Poll #1, July 1997 — CBS News (ICPSR 2304)

CBS News Monthly Poll #1, August 1997 — CBS News (ICPSR 2305)

CBS News Monthly Poll #2, August 1997 — CBS News (ICPSR 2306)

CBS News Monthly Poll #2, September 1997 — CBS News (ICPSR 2308)

CBS News Monthly Poll #1, November 1997 — CBS News (ICPSR 2309)


Census of State and Federal Adult Correctional Facilities, 1995 — United States Department of Justice, Bureau of Justice Statistics (ICPSR 6953)


Drug Offender Treatment in Local Corrections in California and New York, 1991–1993 — Sandra Tunis, James Austin, Mark Morris, Patricia Hardyman, and Melissa Bolyard (ICPSR 6628)


Eurobarometer 43.0: Cross-Border Purchases, Smoking Habits, and Cancer Risks, March–April 1995 — Karlheinz Reif and Eric Marlier (ICPSR 6662)


Eurobarometer 44.3: Health Care Issues and Public Security, February–April 1996 — Karlheinz Reif and Eric Marlier (ICPSR 6732)

Eurobarometer 45.1: European Union Rights, Sun Exposure, Work Safety, and Privacy Issues, April–May 1996 — Karlheinz Reif and Eric Marlier (ICPSR 6749)

Evaluation of the Weed and Seed Initiative in the United States, 1994 — Jan Roehi (ICPSR 6789)


NEW AT ICPSR


The 1997 Pilot Study is part of the National Election Studies (NES) effort to develop new instrumentation. Previous pilot studies were conducted in 1979, 1983, 1985, 1987, 1989, 1991, 1993, and 1995. As in earlier pilot studies (except for 1979), the 1997 study respondents were a subset of the previous year's traditional time-series respondents.

This study is a one-wave re-interview of a randomly-selected subset of respondents with telephones from the fresh cross-section portion of the American National Election Study, 1996: Post-Election Survey (ICPSR 6896). The 1997 Pilot Study was conducted between September 5 and October 1, 1997. Specific topic areas in the study include: (1) a battery designed to improve NES instrumentation on nonelectoral political participation and mobilization; (2) testing of NES instrumentation on group closeness, group difference, and group conflict as a basis of current mass politics, and group threat as a basis of group-based politics; (3) evaluations of the President, Congress, and the Supreme Court using a new battery of items; and (4) the role of religion in citizens' political thinking.

The use of Computer-Assisted Telephone Interviewing (CATI) enabled a number of experimental treatments within the survey instrumentation, including random assignment, early-late placement, and presentation order. In addition, rosters of items such as the thermometer were randomized in administration to minimize order effects.
NEW AT ICPSR

Additions to holdings, continued


Integrated Postsecondary Education Data System (IPEDS): Fall Staff, 1995 — United States Department of Education, National Center for Education Statistics (ICPSR 2514)


Linked Birth/Infant Death Period Data, 1995: [United States, Puerto Rico, Virgin Islands, and Guam] — United States Department of Health and Human Services, National Center for Health Statistics (ICPSR 2285)

Multiple Cause of Death, 1995 — United States Department of Health and Human Services, National Center for Health Statistics (ICPSR 2392)


Natality Detail File, 1993: [United States] — United States Department of Health and Human Services, National Center for Health Statistics (ICPSR 6847)

National Black Election Study, 1996 — Katherine Tate (ICPSR 2029)


National Corrections Reporting Program, 1995: [United States] — United States Department of Justice, Bureau of Justice Statistics (ICPSR 2194)


National Household Education Survey, 1996 — United States Department of Education, National Center for Education Statistics (ICPSR 2149)

National Household Survey on Drug Abuse, 1996 — United States Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies (ICPSR 2391)


National Survey of Jails: Jurisdiction-Level Data, 1996 — United States Department of Justice, Bureau of Justice Statistics (ICPSR 6856)

National Survey of Jails: Jurisdiction-Level Data, 1997 — United States Department of Justice, Bureau of Justice Statistics (ICPSR 2313)


Uniform Crime Reporting Program Data [United States]: County-Level Detailed Arrest and Offense Data, 1996 — United States Department of Justice, Federal Bureau of Investigation (ICPSR 2389)

United States Courts of Appeals Database Phase 1, 1925–1988 — Donald R. Songer (ICPSR 2086)

United States Historical Data on Bank Market Structure, 1896–1955 — Mark D. Hood (ICPSR 2393)

Washington, DC, Metropolitan Area Drug Study (DC*MADS), 1991: Household and Nonhousehold Populations — United States Department of Health and Human Services, National Institute on Drug Abuse (ICPSR 2153)

Washington, DC, Metropolitan Area Drug Study (DC*MADS), 1992: Drug Use Among DC Women Delivering Live Births in DC Hospitals — United States Department of Health and Human Services, National Institute on Drug Abuse (ICPSR 2347)


Washington Post Virginia Governor Poll, October 1997 — The Washington Post (ICPSR 2502)

World Health Organization Collaborative Study on Social and Health Aspects of Aging in Bahrain, Egypt, Jordan, and Tunisia, 1991 — Gary Andrews (ICPSR 6250)

Zambia Democratic Governance Project Political Attitudes Survey, 1993 and 1996 — Michael Bratton (ICPSR 2232)

NEW AT ICPSR


This supplement to the National Crime Victimization Survey was designed to collect data on crime victimization in schools in the United States. Student respondents were asked a series of questions to determine their school attendance in the last six months. Other questions concerning schools were posed, including type of school, distance from home, and general attendance and monitoring policies.

The data present information on the response of the school to student violation of rules, accessibility of drugs, and violence in school, including types of violence and student reaction. Other variables cover general violent crimes, personal larceny crimes, and household crimes and offer information on date, time, and place of crime.

Demographic characteristics of household members such as age, sex, race, education, employment, median family income, and marital status are provided.
NEW AT ICPSR


The National Household Education Survey (NHES) series reports information on the condition of education in the United States by collecting data at the household level rather than using a traditional school-based data collection system. The surveys attempt to address many current issues in education, such as preprimary education, school safety and discipline, adult education, and activities related to citizenship.

This survey included two topical survey components: Parent/Family Involvement in Education (PFI) and Adult and Youth Civic Involvement (CI). The PFI component focused on four areas: types and frequency of family involvement in children's schools, communication with teachers or other school personnel, children's homework and behavior, and learning activities with children outside of school. Other information collected for this component pertained to student experiences at school, children's personal and demographic characteristics, household characteristics, and children's health and disability status.

The CI component of the survey gathered information on civic participation, sources of information about government issues, and knowledge and attitudes about government. This component also addressed opportunities for youth to develop personal responsibility and skills that would facilitate their taking an active role in civic life.

A screener component of the survey collected demographic and educational information on all members in every household contacted. Items on use of public libraries by the household were also administered.

REVISIONS/UPDATES

A National Survey of Access to Medical Care, 1975-1976 — Ronald Andersen and Lu Ann Aday (ICPSR 7730)


Assessment of a Program of Public Information on Health Care Reform, 1992-1993: [Wichita, Kansas, and Des Moines, Iowa] — Sally Daniels and Andrew Kully (ICPSR 6066)

Business Leaders' Views on American Health Care, 1990 — Joel C. Cantor (ICPSR 6032)

Community Policing in Baltimore, 1986-1987 — Antony M. Pate and Sampson O. Annan (ICPSR 9401)


Disorder and Community Decline in Forty Neighborhoods of the United States, 1977-1983 — Wesley G. Skogan (ICPSR 8944)


Effects of Preferred Provider Organizations on Health Care Use and Costs: Pooled Cross-Sectional Time Series, First Quarter 1988 Through First Quarter 1990 — Dean G. Smith (ICPSR 6373)


Euro-Barometer 34.0: Perceptions of the European Community, and Employment Patterns and Child Rearing, October-November 1990 — Karlheinz Reif and Anna Melich (ICPSR 9576)

Euro-Barometer 36.0: Regional Identity and Perceptions of the Third World, Fall 1991 — Karlheinz Reif and Anna Melich (ICPSR 9771)

Euro-Barometer 41.1: Post-European Election, June-July 1994 — Karlheinz Reif and Eric Marlier (ICPSR 6535)

Eurobarometer 43.1: Social, Regional Development and Consumer and Environmental Issues, May-June 1995 — Karlheinz Reif and Eric Marlier (ICPSR 6840)
Eurobarometer 44.0: Cancer, Education Issues, and the Single European Currency, October–November 1995 — Karlheinz Reif and Eric Marlier (ICPSR 6721)

Expanded United States Supreme Court Judicial Database, 1946–1968 Terms — Harold J. Spaeth (ICPSR 6557)

Expenditure and Employment Data for the Criminal Justice System [United States]: Extract File, 1984 — United States Department of Justice. Bureau of Justice Statistics (ICPSR 9162)

Expenditure and Employment Data for the Criminal Justice System [United States]: Extract File, 1985 — United States Department of Justice. Bureau of Justice Statistics (ICPSR 9161)

Extending Health Insurance to the Working Poor: An Assessment of Health Status and Health Care Utilization Effects Among New York City Home Health Attendants, February 1990–June 1991 — Beth Weitzman (ICPSR 9774)


German Social Survey (ALLBUS), 1994 — Zentralarchiv für Empirische Sozialforschung and Zentrum für Umfragen, Methoden und Analysen (ZUMA) (ICPSR 6524)


Interaction Between Neighborhood Change and Criminal Activity, 1950–1976: Los Angeles County — Solomon Kornblum and Leo A. Schurman (ICPSR 9056)


National Survey of Surgeons on Trauma Care Issues, March–July 1993 — Thomas J. Esposito et al. (ICPSR 6265)

National Survey of the Aged [United States], 1957 — Ethel Shanaz (ICPSR 7686)


Practice Patterns of Young Physicians, 1987: [United States] — American Medical Association Education and Research Foundation (ICPSR 9277)

Practice Patterns of Young Physicians, 1991: [United States] — Jack Hadley (ICPSR 6145)

Providing Help to Victims: A Study of Psychological and Material Outcomes in New York City, 1984–1985 — Robert C. Davis (ICPSR 9479)

Public Health on the Agenda of Counties, 1993: [United States] — Vincent L. Marano (ICPSR 6562)


NEW AT ICPSR


This series measures the prevalence and correlates of drug use in the United States. The surveys are designed to provide quarterly, as well as annual, estimates.

Information is provided on the use of illicit drugs, alcohol, and tobacco among members of United States households aged 12 and older. Questions include age at first use as well as lifetime, annual, and past-month usage for the following drug classes: marijuana, cocaine (and crack), hallucinogens, heroin, inhalants, alcohol, tobacco, and nonmedical use of prescription drugs, including psychotherapeutics.

Respondents were also asked about substance abuse treatment history, illegal activities, problems resulting from the use of drugs, personal and family income sources and amounts, need for treatment for drug or alcohol use, criminal record, and needle-sharing.

Questions on mental health and access to care, which were introduced in the 1994-8 questionnaire (see National Household Survey on Drug Abuse, 1994 [ICPSR 6949]), were retained in this administration of the survey. In 1996, the section on risk/availability of drugs was reintroduced, and sections on driving behavior and personal behavior were added. Demographic data include gender, race, age, ethnicity, marital status, educational level, job status, income level, veteran status, and current household composition.
NEW AT ICPSR


These data were gathered to provide information on the elderly in Bahrain, Egypt, Jordan, and Tunisia. Questions were asked regarding demographic characteristics (age, race, sex, marital status, religion, number of children and siblings, education); economic resources (employment status of respondent and of respondent's spouse, main income source, other income sources, whether house was owned); health (current health status, accidents and/or injuries affecting daily activities, number of times respondent saw a health professional in the previous month, number of days in hospital, nursing home, or rehabilitation center in the previous month, medications currently used, usage of any devices to assist in getting around); activities of daily living; living habits (smoking or drinking); social activities (club membership, whether respondent had a confidant); housing (satisfaction with current living conditions, accessibility to quarters, safety); and mental state of the respondent.

Revisions/Updates, continued


Research on Minorities, [1981]: Race and Crime in Atlanta and Washington, DC — Julius Debro (ICPSR 8459)

Small Business Benefits Study (SBBS), 1990: [Denver, Flint, Tampa, and Tucson] — Catherine G. McLaughlin (ICPSR 6002)


Study of Ambulatory Care Sensitive Diagnoses As a Monitor of Primary Access, 1993: [California] — Andrew B. Bindman (ICPSR 6570)


Trends in Hospital and Health Personnel in the United States and Canada, 1968–1991 — Steffie Woolhandler and David U. Himmelstein (ICPSR 6243)


Uniform Crime Reporting Program Data [United States]: County Level Arrest and Offenses Data, 1977–1983 — United States Department of Justice. Federal Bureau of Investigation (ICPSR 8703)

Uniform Crime Reporting Program Data: [United States] — United States Department of Justice. Federal Bureau of Investigation (ICPSR 9028)

Uniform Crime Reports: County Level Arrest and Offense Data, 1976 — United States Department of Justice. Federal Bureau of Investigation (ICPSR 9119)


United States Supreme Court Judicial Database, 1953–1996 Terms — Harold J. Spaeth (ICPSR 9422)

Use of Radiology and Laboratory Tests Among Selected Inpatients in Canadian and United States Hospitals, 1990–1991 — Steven J. Katz (ICPSR 6539)


P

Publication-Related Archive

National Black Election Study Pilot, 1996 — Katherine Tate (ICPSR 1125)

New Approach to the Study of Ticket-Splitting — Barry C. Burden and David C. Kimball (ICPSR 1140)

Of Time and Consensual Norms in the Supreme Court — Gregory A. Caldeira and Christopher J.W. Zorn (ICPSR 1142)

Political Economy and the Dynamics of Party Support in Contemporary Britain — Marianne C. Stewart, Harold D. Clarke, and Paul Whiteley, University of Sheffield (ICPSR 1141)
Announcements

ICPSR Becomes Independent Unit in ISR

ICPSR was founded within the Survey Research Center (SRC) of the Institute for Social Research (ISR), University of Michigan, as a partnership between SRC and 21 universities. When the Center for Political Studies (CPS) was created out of programs in SRC, ICPSR moved to become a part of the new CPS. On July 1, 1998, ICPSR became a free-standing unit of ISR reporting directly to the Director of ISR. Final negotiations are in progress on a new Memorandum of Agreement, under the terms of which ISR will continue to host ICPSR.

Both the ICPSR Council and the ICPSR staff, including the Executive Director, believe that this new reporting relationship has potential benefits for ICPSR. Among these are new possibilities for ICPSR to work across the four centers of ISR. There may be financial benefits as well.

We envision continued cooperation with the many colleagues in CPS who have been so supportive of ICPSR over the years, as well as with personnel and programs of ISR's other three research centers.

Richard Rockwell Takes Sabbatical

Beginning this fall, ICPSR Executive Director Richard Rockwell will be taking a sabbatical to pursue his research interests in global change. In order to facilitate ICPSR's transition to independent status in ISR and its move into new quarters (see other items on this page), he plans to take his sabbatical as a series of short research and teaching trips.

In Richard's absence, Erik W. Austin, Director of Archival Development, will be Acting Executive Director of ICPSR. Erik is an historian with over 30 years of service to ICPSR.

Richard's sabbatical includes several trips to Asia and to Bonn, Germany, for teaching and writing during fall 1998 through spring 1999, followed by a writing period in Ann Arbor during the summer.

ICPSR Moves to New Quarters

As a result of recent additions to the ICPSR staff and projected staffing increases in the near future, ICPSR is expanding into new quarters in the building formerly occupied by Borders Group International at 311 Maynard, Ann Arbor, MI.

To accommodate ICPSR's needs for space, ISR will rent 21,000 sq. ft. in the lower level and first floor of the Borders building. At press time, ICPSR was scheduled to take occupancy sometime after October 1, 1998.

Also, beginning January 1, 1999, the University of Michigan's Population Studies Center, which has become a fourth center of ISR, will move into the second floor of the Borders space.

During ICPSR's transition to the new space, there is a possibility for brief interruption of service to ICPSR users.

However, ICPSR staff is working to minimize any inconvenience caused by the move. Official Representatives will be notified in advance of any service disruption.

Individuals sending correspondence or shipments to ICPSR should continue to use the following addresses:

P.O. Box 1248
Ann Arbor, MI 48106-1248
or
426 Thompson Street
Ann Arbor, MI 48104-2321

Note also that the area code for Ann Arbor has changed from '313' to '734'. New phone numbers for ICPSR were not available at press time.

Conference Attendance Set

ICPSR has a busy conference schedule this fall, with staff representing ICPSR at several professional meetings. In August and September, ICPSR attended meetings of the American Sociological Association and the American Political Science Association. Look for the ICPSR conference display booth at the following upcoming conferences:

- Bureau of Justice Statistics/Justice Research and Statistics Association Sept. 17–18, San Diego, CA
- American Society of Criminology Nov. 11–14, Washington, DC
- American Public Health Association Nov. 15–19, Washington, DC
- Gerontological Society of America Nov. 20–23, Philadelphia, PA

New Look for Bulletin

Starting with this volume, the ICPSR Bulletin carries a new design and a new ICPSR logo. Please let us know what you think, and feel free to suggest ideas for articles or new features.

Also, please note that the quarterly issues of the Bulletin will now reflect the seasons in which they appear, rather than the months.
ICPSR

P.O. Box 1248, Ann Arbor, MI 48106-1248

ADDRESS CORRECTION REQUESTED

Moving? Please send us your new address, along with your old mailing label.

ICPSR COUNCIL MEMBERS, 1998–2000
Margo Anderson, University of Wisconsin, Milwaukee
margo@cdw.uwm.edu
Charles Betsey, Howard University
cbetsey2@aol.com
Kenneth A. Bollen, University of North Carolina, Chapel Hill
bollen@gbibs.unc.edu
Stephen Fieneberg, Carnegie Mellon University
fienberg@stat.cmu.edu
Diane Geraci, State University of New York, Binghamton
dgeraci@library.lib.binghamton.edu
Gary King, Harvard University
king@harvard.edu
Paula McClain, University of Virginia
pmclain@virginia.edu
Edward Nelson, California State University, Fresno
ed_nelson@csufresno.edu
Huey L. Perry, Southern University, Baton Rouge
hperry@subvm.subr.edu
Carole Shamas, Past Chair, University of Southern California
shamas@usc.edu
Elizabeth Stephenson, University of California, Los Angeles
libbie@ucla.edu
Wendy Watkins, Carleton University
wwatkins@ccs.carleton.ca
Hillman A. Winsborough, Chair, University of Wisconsin, Madison
winsborough@ssc.wisc.edu
To reach all Council members:
council@icpsr.umich.edu

ICPSR ASSOCIATE DIRECTORS
Heinz Eulau
Stanford University
Norval D. Glenn
University of Texas, Austin
M. Kent Jennings
University of California, Santa Barbara
University of Michigan
Warren E. Miller
Arizona State University

ICPSR ADMINISTRATION AND STAFF
(note area code change)
Richard C. Rockwell, Executive Director
richard@icpsr.umich.edu (734) 764-2570
Kathleen Thomson, Administrative Manager
kthomson@sr.umich.edu (734) 764-2570
Michelle Humphreys, Administrative Secretary
michelle@icpsr.umich.edu (734) 764-2570

COMPUTING AND NETWORK SERVICES
John E. Gray, Director
jgray@icpsr.umich.edu (734) 763-3482

EDUCATIONAL RESOURCES
Henry Helitowt, Director
hank@icpsr.umich.edu (734) 764-8392

ARCHIVE STAFF
Erik W. Austin, Director, Archival Development
erik@icpsr.umich.edu (734) 936-1752
Janet Vavra, Technical Director
jan@icpsr.umich.edu (734) 763-5010
Peter Granda, Assistant Archival Director
peter@icpsr.umich.edu (734) 936-1752
Christopher S. Dunn, Assistant Archival Director
cdunn@icpsr.umich.edu (734) 763-5011
Mary Morris, Research Associate, User Support
morris@icpsr.umich.edu (734) 763-5010
JoAnne McFarland, Manager, Substance Abuse and Mental Health Project
jmcfar@icpsr.umich.edu (734) 936-1752
Kenneth F. Ferraro, NACDA Resident Scientist, Purdue University
ferraro@purdue.edu

ICPSR BULLETIN
FALL 1998 — VOL. XIX, NO. 1

The Inter-university Consortium for Political and Social Research (ICPSR), located at the Institute for Social Research in Ann Arbor, is the world's largest repository of computer-readable social science data. For over 35 years, the Consortium has served the social science community by acquiring, processing, and distributing data collections on a broad range of topics. Researchers at the Consortium's member institutions may obtain any of these data collections at no charge; researchers at nonmember institutions may also use the data, after paying an access fee. To find out more about ICPSR's holdings or about a specific data collection, access the ICPSR Website at the URL: http://www.icpsr.umich.edu.

The ICPSR Bulletin is published four times during each academic year to inform Official Representatives at the member campuses, ICPSR Council members, and other interested scholars of activities occurring at ICPSR and at other member institutions and to list the data collections most recently released or updated by ICPSR. For subscription information, contact the Editor.

Subscription Price: $15 per year

ICPSR
Institute for Social Research
426 Thompson St., Ann Arbor, MI 48104-2321
P.O. Box 1248, Ann Arbor, MI 48106-1248
Phone: (734) 764-2570
Fax: (734) 764-8041
E-mail: netmail@icpsr.umich.edu

Mary B. Vardigan, Editor
Amy B. Garber, Associate Editor

Printed on recycled paper