

Fall 2003 • Volume XXIV, No. 1

ICPSR BULLETIN

In this issue:

JoAnne McFarland O'Rourke

Disclosure Analysis at ICPSR

- ICPSR Leadership Over the Years:
Photo Collage
- Privacy in the Information Age:
A Symposium in Honor of the 40th
Anniversary of ICPSR
- Complex Modelling: 33rd Spring
Seminar to be Held at the Zentralarchiv
- Research Scientist Position Available for the
National Archive of Criminal Justice Data
- Ronald Rindfuss Appointed to ICPSR Council

***40 Years of Service to the
Social Sciences***

• Data Resources for the Future •

Inter-university Consortium for Political and Social Research



Inter-university Consortium for Political and Social Research

Institute for Social Research • The University of Michigan

426 Thompson Street • Ann Arbor, MI 48106-2321

ICPSR ADMINISTRATION AND STAFF

Myron Gutmann, Director
gutmann@icpsr.umich.edu (734) 615-8400

Rita Bantom, Human Resources Manager
rybantom@icpsr.umich.edu (734) 647-5000

Stacey Kubitz, Business Manager
skubitz@icpsr.umich.edu (734) 615-2315

Bree Gunter, Administrative Assistant
bgunter@icpsr.umich.edu (734) 615-8400

ARCHIVAL DEVELOPMENT

Erik Austin, Director
erik@icpsr.umich.edu (734) 615-7667

Peter Granda, Assistant Archival Director
peterg@icpsr.umich.edu (734) 615-2977

Christopher S. Dunn, Assistant Archival Director
cdunn@icpsr.umich.edu (734) 647-5000

JoAnne McFarland O'Rourke, Director
Substance Abuse and Mental Health Archive
jmcfar@icpsr.umich.edu (734) 615-9528

James McNally, Director, Program on Aging
jmcnally@icpsr.umich.edu (734) 615-9520

Duane Alwin, NACDA Resident Scientist
dalwin@pop.psu.edu Pennsylvania State University

EDUCATIONAL RESOURCES

Henry Heitowit, Director
hank@icpsr.umich.edu (734) 763-7400

Dieter Burrell, Assistant Director
dburrell@icpsr.umich.edu (734) 763-7400

COMPUTER AND NETWORK SERVICES

Bryan Beecher, Director
bryan@icpsr.umich.edu (734) 615-2976

MEMBER RELATIONS

Henry Heitowit, Director
hank@icpsr.umich.edu (734) 763-7400

Bree Gunter, Membership Coordinator
bgunter@icpsr.umich.edu (734) 615-8400

TECHNICAL SERVICES

Janet Vavra, Director
jan@icpsr.umich.edu (734) 647-2200

Mary Morris, Assistant Director
morris@icpsr.umich.edu (734) 647-2200

WEB RESOURCES DEVELOPMENT

Mary Vardigan, Director
maryv@icpsr.umich.edu (734) 615-7908

ICPSR COUNCIL MEMBERS, 2002-2004

Ann Green, Chair
Yale University, ann.green@yale.edu

William Darity Jr.
University of North Carolina-Chapel Hill & Duke University, darity@unc.edu

Ilona Einowski
University of California, Berkeley, ilona@ucdata.berkeley.edu

Elisabeth R. Gerber
University of Michigan, ergerber@umich.edu

John Handy
Morehouse College, jhandy@morehouse.edu

Mark Hayward
Pennsylvania State University, hayward@pop.psu.edu

John Korey
California State Polytechnic University, Pomona, jkorey@csupomona.edu

J. Scott Long
Indiana University, slong@indiana.edu

Ronald Rindfuss
University of North Carolina-Chapel Hill, rindfuss@email.unc.edu

Steven Ruggles
University of Minnesota, ruggles@hist.umn.edu

James Sweet
University of Wisconsin-Madison, sweet@ssc.wisc.edu

Bo Wandschneider
University of Guelph, bo@uoguelph.ca

Margo Anderson, Past Chair
University of Wisconsin-Milwaukee, margo@csd.uwm.edu

ICPSR BULLETIN

Fall 2003 • Volume XXIV, No. 1

The Inter-university Consortium for Political and Social Research (ICPSR), located at the Institute for Social Research at the University of Michigan in Ann Arbor, is the world's largest repository of computer-readable social science data. For 40 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, visit the ICPSR Web site at www.icpsr.umich.edu.

The *ICPSR Bulletin* is published to inform interested scholars, including Official Representatives at the member campuses and ICPSR Council members, about topics and events relevant to ICPSR and its data collections. For subscription information, please contact the Editor.

Subscription rate: \$15 per year

Editor: Ruth Shamraj

ICPSR BULLETIN: FALL 2003

DISCLOSURE ANALYSIS AT ICPSR

JoAnne McFarland O'Rourke

Inter-university Consortium for Political and Social Research,
University of Michigan

Disclosure analysis involves the careful examination of a data file for indirect identifiers that could pose the risk of re-identification of a respondent. By examining variables containing detailed personal characteristics such as education, income, race, ethnicity, and military service or organizational characteristics such as capacity, services offered, and programs for special populations, it quickly becomes possible to begin to narrow identity. The availability of large and inexpensive disk storage capacity, the increased sophistication of databases that may be used to link data, and the easy availability of data via the Internet, combined with federal mandates to release data and user demand for data, create increasing awareness and concern regarding disclosure risk.

Disclosure analysis typically focuses on attributes that can be known by an outsider (e.g., age) rather than attitudes or beliefs (e.g., feelings about international military involvement). A critical factor in disclosure assessment is the availability of geographic data. The more specific the geography, the more attention must be paid to disclosure risk. Two types of disclosure are distinguished. Type I disclosure occurs when an intruder has knowledge that a

given person (or organization) is included in a survey and the intruder attempts to find this record. Type II disclosure occurs when an intruder does not know the identity ahead of time and uses externally available resources (linking databases) to attempt to find survey respondents.

Some of the disclosure techniques in common use include:

- Releasing only samples from large data collections
- “Coarsening” the data (e.g., collapsing categories, top- and bottom-coding, converting continuous variables to categories)
- Suppressing data that illuminate unique cases (e.g., blanking variables with sensitive information or removing unique cases altogether)
- Perturbing values of the data (e.g., adding random noise or distortion, microaggregating, swapping cases, suppressing, and re-imputing data)

These techniques are used to “blur” the data such that an intruder cannot be certain she or he has actually found a given respondent, and in a case where an intruder claims



JoAnne McFarland O'Rourke, MSW, is Director of the Substance Abuse and Mental Health Data Archive at ICPSR at the University of Michigan. Her research experience includes studies on HIV/AIDS, adoption assistance for special needs children, child abuse and neglect, substance abuse among women, and domestic violence education, among other topics.

to have found a respondent, provide deniability for the respondent.

Goals of Disclosure Analysis

Goals of disclosure analysis include determining specific disclosure issues and then addressing those issues using disclosure protection methods to reduce risk. We use the term *reduce risk* to acknowledge that risk is never completely eliminated. Analysts are often interested in subgroups of survey populations (e.g., pregnant women, racial minorities, persons having committed crimes, and persons with health conditions such as HIV) and comparisons of subsets within the survey. Yet these are often the very characteristics that create disclosure risk. Disclosure techniques must take into account the key uses of the data and balance the trade-off between analytic utility and data protection.

Different methods of reducing disclosure risk are used based on the type of data and the analytic concerns. For example, cell suppression disproportionately affects sub-groups in the data, such as racial minorities, because these records tend to stand out (or be unique) in the data. Categorizing continuous variables such as income makes it impossible to calculate means. For some data collections, this would not be a serious impediment, but for others it would. Longitudinal files pose different risks than cross-sectional data, simply because more data are collected on the same individual (or organization) across time. Changes in location, marital status, career, etc., make a record more unique, and therefore easier to identify.

Rather than applying simplistic techniques that would render the public-use data unsatisfactory for key analyses or make the data available on a restricted basis only, the application of disclosure techniques, in most cases, allows dissemination of a public-use version of the

data. Disclosure analysis often involves trial and error in determining the most appropriate methods for reducing disclosure risk and must include an analysis of the impact on the data: What is lost and gained by the methods proposed? Which analytic capabilities are diminished? Which are preserved? How will the information lost affect data interpretation? Can the lost information be released in some other way?

Disclosure Analyses at ICPSR

Disclosure analysis is perhaps more easily understood with examples of specific applications. ICPSR has conducted two full disclosure analyses within the Substance Abuse and Mental Health Data Archive (SAMHDA), as detailed below.

Treatment Episode Data Set

The first formal disclosure analysis conducted at ICPSR was for the 1998 Treatment Episode Data Set (TEDS),¹ which is based on an administrative data system that collects approximately 1.6 million records annually on admissions for substance abuse treatment. TEDS presented a disclosure risk because it contained records that were unique based on combinations of client characteristics that could

“A critical factor in disclosure assessment is the availability of geographic data. The more specific the geography, the more attention must be paid to disclosure risk.”

be known, such as race, ethnicity, and age. For analytic purposes and to satisfy requests of a key

user constituency for the public-use data, it was important to leave Primary Metropolitan Statistical Area (PMSA) and detailed race and ethnicity codes on the file.

Our goals in conducting the disclosure analysis were (1) to protect the confidentiality of treatment clients, (2) to leave as much detail as possible on the files, (3) to make the files as analytically useful as possible, (4) to provide a level of detail that was adequate for key constituencies, and (5) to create a disclosure analysis plan that could serve as a model for future disclosure analyses. TEDS is an

annual file and each year of the file is updated as many as five times. Therefore, we needed fully automated procedures that could be employed repeatedly.

Our methodology reduced the risk of identification of an individual by deleting one variable type that presented a significant disclosure risk and yet had very limited analytic utility. Risk of identification was further reduced by recoding variables such as age and education. These measures helped eliminate the “uniqueness” of any given record. For the unique records that remained, we swapped a sample of records that matched on certain variables.

Data swapping involves identifying the set of variables that when taken together could potentially identify an individual, and then substituting “matching” records. The set of identifying variables is referred to as the “uniques key.” For TEDS, these variables included age, sex, methadone planned as part of treatment,² race, ethnicity, pregnancy, and

veteran status. Unique records within the file were then identified and a random sample of

these records was drawn with probability p . A record was considered unique if it was the only record with a unique set of characteristics on the above set of variables within a PMSA.

Matching records within the file were identified for each unique record in the sample. Records were matched based on race, ethnicity, sex, age, pregnancy, primary substance of abuse, and methadone planned as part of treatment; this is referred to as the “swapping key.” The “swapping attribute” is the variable over which swapping occurs, and this is typically a geographic variable. Matches were first sought within Census division, then Census region, and then across the entire file. The unique records were then swapped with a matched record selected at random. Matches were found for more than 90 percent of the unique records on the first run of the procedures. For the remaining unmatched records, the swapping key was loosened by allowing matching on fewer variables or broader

categories, until matches were found for all records. In summary, our procedures were as follows:

- 1) Perform variable deletions and recodes
- 2) Identify unique records
- 3) Run a coin toss simulation to determine the subset of records to be swapped
- 4) Randomize all records
- 5) Identify pairs of swappable records based on the swapping key match and geographic hierarchy
- 6) Loosen the swapping key for remaining unmatched records
- 7) Run the swap program and finalize swaps
- 8) Manually review a sample of swapped records

These procedures introduced a significant factor of uncertainty into the public-use file in terms of identifying

any individual record while retaining data integrity and utility. Data swapping had several benefits

“Disclosure techniques must take into account the key uses of the data and balance the trade-off between analytic utility and data protection.”

over other disclosure protection options for TEDS: (1) The overall impact on the data was very small, and less than 1 percent of the records were impacted by the disclosure procedures; (2) Data for special populations (e.g., minorities, pregnant women) were no more impacted than other data; and (3) The procedures allowed greater detail to remain on the public-use file (e.g., PMSA and the original race and ethnicity codes).^{3,4}

The initial TEDS public-use file took about two months to produce. We have streamlined our automated procedures such that the files now take just a few days to produce, and most of this is the run time required due to the large size of the files; little time is required by the data processor.

Alcohol and Drug Services Study (ADSS)

The second disclosure analysis was conducted on the Alcohol and Drug Services Study (ADSS), which included interviews with facility directors for a national probability sample of substance abuse treatment facilities, additional interview data at the facility level, client record abstracts for a subsample of facilities, and client follow-up interviews for clients whose records were abstracted. The key issue with ADSS was the availability of external files that matched certain facility-level client count and revenue variables. While ADSS was a national sample and thus it was not important to include geographic codes, the external files included state, PMSA, and county FIPS. Therefore, the files could be linked on the common client count and revenue variables, allowing a data intruder to be able to identify the geographic location of ADSS facilities and increasing the risk of the identification of respondents.

We discussed several options to resolve the problem of the linkage between the files, including categorizing values, releasing unlinkable facility and client files, and including only selected facility-level variables in the public-use file. While categorizing values is a simple approach, it distorts means and ratios, which are likely to be of interest to researchers (e.g., for revenue and client counts). Releasing unlinkable files or providing a public-use file with selected variables would severely limit the utility of the files and create a situation in which many researchers would likely request the original (restricted-use) files. Therefore, we sought a solution that would allow the release of all of the data, maintain linkages between the files, and yet protect facilities and clients. Our approach was to use microaggregation for the problematic (matching) variables, thus breaking the link to the external files.

In its basic form, microaggregation involves ordering microdata along a single variable and then aggregating records in groups of three or more. Within each grouping, the reported value on all variables included in the procedure (e.g., those considered problematic) is replaced by the average value of the group for each variable. This form of microaggregation works best for variables that are highly correlated.

A refinement of the basic procedure, and the one employed for ADSS, is to repeat the grouping procedure on each problem variable. In this approach, the data are ordered on the first problem variable, grouped by three (or

more), and the values for each record on problem variable one are replaced by the average value for the record's grouping. Next, the data are ordered

“Longitudinal files pose different risks than cross-sectional data, simply because more data are collected on the same individual (or organization) across time. Changes in location, marital status, career, etc., make a record more unique, and therefore easier to identify.”

by the second problem variable, different groupings result, and averages for these new groupings replace the original values for problem variable two. The procedure is repeated for each problem variable.

By design, microaggregation leaves all attribute mean values unchanged, while only slightly reducing variance. Since mean values are the most basic summary statistic, microaggregation has a fundamental advantage over other categorizing and recoding techniques. A more common approach is to recode continuous variables into broad categories, preventing measures of central tendency (e.g., revenue categorizations of (1) less than \$500,000 or (2) \$500,000 or more), thus seriously impacting uses of the data. However, in microaggregation, the data themselves determine the ranges, and the ability to use measures of central tendency is preserved.

Though a confidentiality risk might exist for only a subset of records in a file (as with ADSS), microaggregation is performed across all records. When the groupings are performed over a large number of records, the values are likely to be closer to each other, thereby creating less

distortion than if the groupings are done over a subset, in which the values may be widely disparate.

There were just two variables within ADSS that provided a direct match to the external files. These two variables were carried forward to two additional variables and were also included in tables that needed to be recalculated based on the results of the microaggregation (so that columns within the tables added correctly). Additional measures taken with ADSS included deleting direct identifiers such as facility name, deleting administrative variables such as date and time of the interview, recoding variables such as Diagnostic and Statistical Manual (DSM) codes, and when possible, converting time frames to standard units.

In order to assess the impact on the data, for the microaggregated and recalculated variables, the cells that changed more than 5 percent in either direction were calculated as a percentage of valid cells (including zero) and as a percentage of total cells. Because a large number of valid values in ADSS were zero, we also calculated the cells that changed more than 5 percent as a percentage of non-missing and non-zero cells.

The results show that less than 1 percent of the non-missing and

non-zero microaggregated variables changed more than 5 percent, while 3.6 percent of the recalculated variables changed more than 5 percent. Of all valid cells (including zero) for microaggregated variables, less than 1 percent changed more than 5 percent while fewer than 2 percent of the recalculated variables did so.

Considering likely uses of the data, we further examined the impact on the data by comparing pre- and post-microaggregation ratios and means and by running regression models on the pre- and post-microaggregated data to determine if significance results were comparable between the files. The results were favorable, showing only minor differences in the means and ratios and indicating that the regression coefficients and the interpretation of the sig-

nificance of the associated effects were robust against the changes to the data resulting from the microaggregation.⁵

Publicly Available Documentation of Procedures

To the extent possible, it is important that disclosure protection procedures are made publicly available so that analysts are aware of the changes to the data. For both TEDS and ADSS, we identified the details that we deemed important to relay to data users that could be safely released to the public. Based on these decisions, we documented our procedures and the results of our disclosure limitation techniques for inclusion in the public-use codebooks.⁶

New Disclosure Project at ICPSR

So what is the real risk of disclosure? How successfully can indirect identifiers be used to re-identify a respondent? No one really knows. Current practices rely largely upon rules of thumb for protecting data or upon the

“To the extent possible, it is important that disclosure protection procedures are made publicly available so that analysts are aware of the changes to the data.”

federal statistical agencies’ provision of cleansed datasets.⁷ Some data producers and distributors

may do too little to reduce disclosure risk thus exposing respondents to risk, while others may do too much, thereby unnecessarily reducing analytic utility. Further, most data producers and distributors do not have the resources to conduct formal disclosure analyses. Rarely does an organization such as ICPSR receive a directive to conduct formal disclosure analyses as we did with TEDS and ADSS.⁸

Recognizing the significant and increasing issues regarding disclosure risk, several ICPSR and Survey Research Center (SRC) researchers began meeting last year to discuss submitting a proposal to the National Institutes of Health (NIH) that would tackle some of the remaining questions regarding disclosure risk and related topics. The result was a proposal for a program project submitted in February 2003 to

the National Institute on Child Health and Human Development (NICHD). We are delighted to report that we received a very good score on our proposal and are awaiting formal notification of award. The program includes four interrelated projects:

Project 1, headed by **Eleanor Singer** of the SRC is “Informed Consent and Perceptions of Risk and Harm in Survey Participation.” Singer and SRC co-investigators

Fred Conrad,

Mick Couper,

and **Bob Groves**

will look at the

level of risk of

disclosure that

the public is

willing to ac-

cept; whether disclosure of some kinds of information is considered more harmful than others; whether some data intruders are perceived as more harmful than others; whether people perceive the relationship of expected risk of harm *versus* magnitude of harm and risk of disclosure in a manner consistent with the mathematical probability of such occurrences; and how researchers can accurately inform participants of the risks without unnecessarily deterring them from participation. The project will use Web-based experiments, including tests of informed consent statements. This will mark the first time that such an investigation will be informed by statistical analyses of the actual risks of disclosure present in the research. The group hopes to provide guidance for researchers who are struggling to balance accurate communication of the risk of harm with credible assurances of confidentiality.

T.E. Raghunathan (SRC) will lead Project 2, entitled “Estimation of Disclosure Risk and Statistical Methods for Disclosure Limitation.” Co-investigators on this project will be **Ben Hansen**, **Rod Little**, and **Richard Valliant**, also from the SRC. Their objectives include (1) assessment of the risk of disclosure using test-bed national probability surveys covering diverse topics, including an assessment of Type I and II disclosure; (2) development and evaluation of new methods to prevent disclosure; and (3) development of strategies for replacing variables in

public-use datasets that might increase the risk of disclosure with other variables that can be used without knowing their actual value.

Project 3 is headed by **JoAnne McFarland O’Rourke** at ICPSR, with **Myron Gutmann** as co-investigator. The project, “Best Practices and Tools for the Social Sciences,” will develop best practices for disclosure limitation by reviewing the literature on disclosure and then surveying

the principal investigators and others involved in disclosure decisions for a sample of studies funded by NIH and NSF. The

“Some data producers and distributors may do too little to reduce disclosure risk thus exposing respondents to risk, while others may do too much, thereby unnecessarily reducing analytic utility.”

survey is intended to ascertain the level of knowledge among researchers and archivists regarding disclosure risk and the practices used to limit risk. Results from these two components of the project, along with relevant findings from Projects 1 and 2, will be used to define best practices. ICPSR data will be used to test best practices for different types of data. Tools that incorporate best practices will then be designed. We anticipate disseminating these tools, much as we currently distribute the *ICPSR Guide to Social Science Data Preparation and Archiving*. **Corey Colyer** of SAMHDA will assist with Project 3.

James McNally (ICPSR) leads the fourth and final project, “Resources for the Secure Dissemination of Human Subject Data,” with **Myron Gutmann** as co-investigator. This project will draw upon the strengths of the ICPSR and SRC as recognized leaders in the training of professional social science researchers toward its goal of disseminating tools and information to enhance the understanding of disclosure analysis. It is expected that a series of programs, training seminars, and information distribution systems will be initiated based on the outcomes of this project.

More Information

The ICPSR Web site contains information that we have compiled regarding our confidentiality and disclosure protection procedures. The Web address is: <http://www.icpsr.umich.edu/irb>. Information on the NICHD disclosure project, as well as other relevant information regarding confidentiality, disclosure analysis, and disclosure limitation procedures at ICPSR, will be posted to these pages.

Footnotes

¹ We distinguish between the routine confidentiality reviews that are conducted on every dataset ICPSR receives and the thorough and detailed work involved in a formal disclosure analysis. While a confidentiality review, which includes examination of direct identifiers such as Social Security number and date of birth, and in some cases includes techniques used in disclosure analysis such as top- and bottom-coding, a formal disclosure analysis is significantly more comprehensive in that it includes in-depth analysis of indirect identifiers and consideration of externally available files that may be linkable.

² All methadone clinics are federally licensed. Some sparsely populated states have just one or two licensed facilities, and the names of the facilities are easy to find. Therefore, knowing that methadone was planned as part of treatment could indicate the approximate geographic location of the client.

³ O'Rourke, J.M., S.F. Roehrig, W.C. Birdsall, B.G. Reed, S.G. Heeringa, M.A. Overcashier, and A. Stanulis. *Disclosure Analysis Plan and Results for The 1998 Treatment Episode Data Set*. Confidential Report submitted to the Office of Applied Studies, Substance Abuse and Mental Health Services Administration, May 2001.

⁴ For further discussion of data swapping see (1) Steel, P., and L. Zayatz, "Disclosure Limitation for the 2000 Census of Housing and Population," in *Statistical Data Protection: Proceedings of the Conference*, Lisbon, 25-27

March 1998, Eurostat, 1999; and (2) Reiss, S., "Practical Data Swapping: The First Steps," *ACM Transactions on Database Systems*, 9 (March 1984).

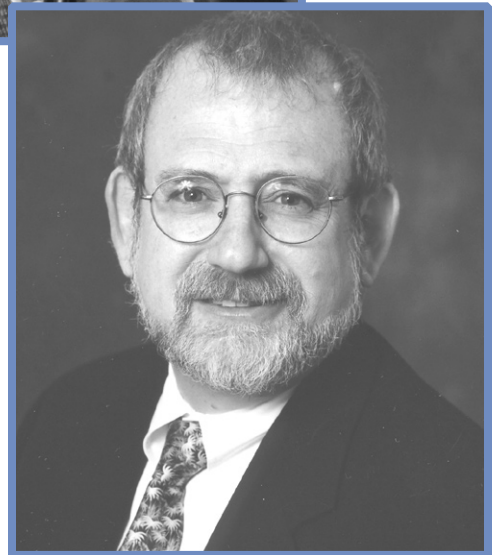
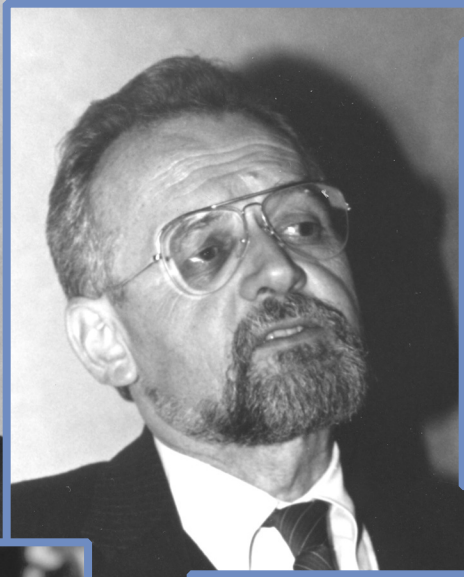
⁵ O'Rourke, J.M., S.G. Heeringa, S.F. Roehrig, M.A. Overcashier, W.C. Birdsall, and B.G. Reed. *Disclosure Analysis and Data Protection Plan for the Alcohol and Drug Services Study*. Confidential Report submitted to the Office of Applied Studies, Substance Abuse and Mental Health Services Administration, December 2001.

⁶ For TEDS, see the public-use codebook available at <http://www.icpsr.umich.edu:8080/SAMHDA-SERIES/00056.xml>. For ADSS, see the codebook at <http://www.icpsr.umich.edu:8080/SAMHDA-DISPLAY/03088.xml>.

⁷ Currently, the most widely used tool for assessing risk is the Checklist on Disclosure Potential, produced by the Federal Committee for Statistical Methodology (FCSM) in the 1990s. While the Checklist represents a significant advancement in standardizing disclosure practices and lays out the considerations for disclosure risk assessment, it stops short of providing specific instructions and detailed examples for applying disclosure limitation procedures to different data types (based on content, population, etc.). The Checklist is also primarily targeted at statistical agencies. It presumes the availability of disclosure experts or a disclosure review board with whom the data provider may consult. Moreover, the Checklist does not provide a mechanism for *measuring* disclosure risk. For the Checklist, see the following publication: *Interagency Confidentiality and Data Access Group, Federal Committee on Statistical Methodology* (1999). "Checklist on Disclosure Potential of Proposed Data Releases." Washington, D.C.: Statistical Policy Office. Office of Information and Regulatory Affairs. Office of Management and Budget.

⁸ SAMHDA, ICPSR, and the disclosure committee members thank the Office of Applied Studies, Substance Abuse and Mental Health Services Administration, for the opportunity to conduct the TEDS and ADSS disclosure analyses.

ICPSR Leadership Over the Years



ICPSR Directors

From left, top row: Warren Miller, Founder, 1962-1970; Richard Hofferbert, 1970-1975; Jerome Clubb, 1975-1991;
Center row: Richard Rockwell, 1991-2000; Halliman Winsborough, Acting Director, 2000;
Bottom row: Erik Austin, Acting Director, 2000-2001; Myron Gutmann, 2001-present.

ANNOUNCEMENTS

Privacy in the Information Age: A Symposium in Honor of the 40th Anniversary of ICPSR

The 27th Biennial Meeting of ICPSR Official Representatives (ORs) will take place October 9-12, 2003, at the University of Michigan in Ann Arbor. In honor of ICPSR's 40th anniversary, this year's meeting will include a special symposium, "Privacy in the Information Age."

ICPSR was formed in 1962 as a partnership among 21 universities with the aim of promoting wider access to social science data. Forty years later, the organization has grown to over 500 member institutions and the ICPSR Data Archive is world renowned for its collection. ICPSR data have served as the foundation for tens of thousands of seminal research publications.

ICPSR seeks to stimulate a conversation about the tension between providing broad and equitable access to data and protecting individual privacy, a tension that all of us involved in the research enterprise face. The symposium is designed to shed light on how these issues of access and privacy play out in different domains and to generate discussion on ways to negotiate this complicated terrain and measures we might take to achieve appropriate balance.

The symposium will be held Friday, October 10, 2003, from 3:30 to 5:30 p.m. at the Michigan Union, Anderson Room.

Complex Modelling: 33rd Spring Seminar to be Held at the Zentralarchiv

The Spring Seminar, to be held on March 1-19, 2004, at the Zentralarchiv für Empirische Sozialforschung, Universität zu Köln, is a training course for social scientists interested in advanced techniques of data analysis and in the application of these techniques to data. Participants must have a sound basic knowledge of statistics as well as experience in handling of PCs and working with SPSS. Stata will be used in the third week.

The Spring Seminar comprises lectures, exercises, and practical work using personal computers. The general topic will be complex modelling with a focus on teaching multivariate analysis techniques. In addition to the lectures, the participants will be provided with information about func-

tions and services of the Zentralarchiv, which is the German data archive for survey data.

The seminar courses offered are: Multilevel Analysis, Mixture Modelling, and Generalised Linear Latent and Mixed Models (GLLAMM). A more detailed programme is available on the ZA Web site: www.gesis.org/ZA. The number of participants is limited to 40 persons per week. Participants will be accepted by order of application date.

Ronald Rindfuss Appointed to ICPSR Council

The ICPSR Council has selected Ronald Rindfuss, Robert Paul Ziff Distinguished Professor of Sociology at the University of North Carolina, Chapel Hill, to complete the term of Council member Franklin Gilliam, who is vacating his seat on the Council for scheduling reasons. Professor Rindfuss will attend the October 2003 ICPSR Council meeting and will serve until a new Council is seated in March 2006.

A social demographer whose work focuses on the timing and sequencing of cohabitation, marriage, childbearing, divorce, education, migration, and employment, Rindfuss is also working on the relationship between population processes and the environment. In collaboration with several Carolina Population Center (CPC) Fellows, Pramote Prasartkul, and others at the Institute for Population and Social Research, Mahidol University, he is examining migration and social change in Thailand. He is also investigating the consequences of childcare availability on fertility in Norway.

Rindfuss, who received his Ph.D. from Princeton University, is author and editor of four books and over 100 research contributions dealing with demography, including *First Births in America: Changing in the Timing of Parenthood and Changing Numbers*, and *Changing Needs: American Indian Demography and Public Health*. He has been a leader in the interdisciplinary field of the human dimensions of global change, promoting the usefulness of remote sensing and the GIS approach to social science problem-solving as attested in his recent co-edited volume, *People and the Environment*.

Rindfuss is a past president of the Population Association of America and former director of Carolina Population Center. He has been a fellow of the American Association for the Advancement of Science since 1992.

Research Scientist Position Available for the National Archive of Criminal Justice Data

ICPSR is searching for a Research Scientist to direct the National Archive of Criminal Justice Data (NACJD) — a data archive that disseminates data pertinent to crime and the criminal justice system to the public. Responsibilities for this multiple-role position include directing an extensive data processing operation, developing strategies to make the data well-known and easy to use, and conducting research. The primary appointment will be as a Research Scientist at ICPSR. A faculty appointment at the University of Michigan in Ann Arbor or Dearborn is also possible.

To qualify, candidates must possess a Ph.D. in a relevant field of study with specialization in research on crime and criminal justice, along with at least 5-8 years of related post-doctoral professional experience, with two years of experience in the management of related research projects. Experience in teaching or training in the criminal justice area; familiarity and experience with archived data; and a background in serving as a principal investigator are preferred.

ICPSR offers a highly competitive compensation and benefits package. Interviews for this position will begin soon. To apply, please submit a letter of application, a CV, three letters of reference, and relevant writing samples to Myron Gutmann, Director, Inter-university Consortium for Political and Social Research, Institute for Social Research, P.O. Box 1248, Ann Arbor, Michigan, 48106-1248. For more information on this position, please visit www.icpsr.umich.edu. The University of Michigan is a nondiscriminatory, affirmative action employer.

ICPSR Membership Staff Change

After many years of distinguished service, ICPSR's Membership Coordinator, Michelle Humphres, has accepted an offer of employment elsewhere. Membership services will now be provided by Bree Gunter, who also serves as administrative assistant to the Director, Myron Gutmann. If you have membership questions, please contact Bree at membership@icpsr.umich.edu.

Has your institution signed up for ICPSR Direct yet? This service enables all faculty, staff, and students at participating institutions to download data directly from the ICPSR Web site. Sign up at

www.icpsr.umich.edu/or/beta-form.html



Data Resources for the Future



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

ADDRESS SERVICE REQUESTED

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