

Overview of 2002 CIPSEA: Methods to Protect Confidential Tabular Data

Amrut Champaneri, Ph.D.

BTS Confidentiality Seminar Series, March 2003



U.S. Department of Transportation
Bureau of Transportation Statistics



Outline

- What is 2002 CIPSEA?
- How does CIPSEA affect Federal agencies?
- Methods to protect confidential tabular data
- Today's seminar

Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA)

- New law affects all executive branch Federal agencies
- Public Law 107-347: Title V of E-Gov't Act of 2002
- Law available at -

http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107_cong_public_laws&docid=f:publ347.107.pdf

CIPSEA Subtitle A, Confidential Information Protection

- An agency may collect information under a pledge of confidentiality for statistical purposes
- This information may not be disclosed in identifiable form for any non-statistical purpose without the informed consent of a respondent
- This information is exempt from release under the Freedom of Information Act (FOIA)

CIPSEA Benefits for Federal Agencies

- Most agencies did not have specific laws ensuring confidentiality of information
- Agencies can now better protect data collected for only statistical purposes
- Higher level of confidentiality may encourage respondents to participate in data collections
- Agencies can avoid disputes about withholding information under FOIA requests

Information Distinction Under CIPSEA

- Statistical purposes - using information to describe or make estimates about whole or subgroups of the economy, society, or environment
- Non-statistical purposes - using information for administrative, regulatory, law enforcement, judicial, or other purposes that may affect the rights, privileges, or benefits of a respondent

Information Collected Under CIPSEA

- Statistical purposes
 - Must protect data; cannot allow direct or indirect identification of data provider
 - Data cannot be shared for non-statistical purposes
 - Data can be shared for statistical purposes by written agreements; data user bound to provide same level of protection
 - Class E felony for disclosing confidential data (5 years prison and/or \$250,000 fine)

Information Collected Under CIPSEA

- Non-statistical purposes
 - A Federal agency must clearly explain to data provider before any data is collected, that it will be used for non-statistical purposes
 - Data is considered non-confidential

Protecting Confidential Data

If information is collected for a statistical purpose –

- Federal agency must have controls to ensure protection of confidential information
- Must act to protect the information
 - Internal agency procedures
 - Information in tables, charts, graphs, and text
 - Microdata; i.e., information about individual data providers

Methods to Protect Confidential Tabular Data

- Aggregation – collapse columns or rows
- Perturbation – add “noise” to tabular data
 - Add random noise to cells in table
 - Round cell values
 - Controlled adjustment – target sensitive cells
 - Markov – unbiased cycle of cell modification
 - Add noise to microdata before creating tables

Methods to Protect Confidential Tabular Data

- Complementary cell suppression –
 - Blank out sensitive cell data
 - Blank out other cells that could be used to recover sensitive cells

Practical Considerations in Selecting Statistical Disclosure Methodology for Tabular Data

Rich Allen

Deputy Administrator, NASS