

# TABLE OF CONTENTS

	Page
<b>I. Introduction .....</b>	<b>1</b>
A. Subject and Purposes of This Report .....	1
B. Some Definitions .....	2
1. Confidentiality and Disclosure .....	2
2. Tables and Microdata .....	3
3. Restricted Data and Restricted Access .....	3
C. Report of the Panel on Confidentiality and Data Access .....	4
D. Organization of the Report .....	4
E. Underlying Themes of the Report .....	5
<b>II. Statistical Disclosure Limitation: A Primer .....</b>	<b>6</b>
A. Background .....	6
B. Definitions .....	7
1. Tables of Magnitude Data Versus Tables of Frequency Data .....	7
2. Table Dimensionality .....	8
3. What is Disclosure? .....	8
C. Tables of Counts or Frequencies .....	10
1. Sampling as a Statistical Disclosure Limitation Method .....	10
2. Special Rules .....	10
3. The Threshold Rule .....	12
a. Suppression .....	12
b. Random Rounding .....	14
c. Controlled Rounding .....	15
d. Confidentiality Edit .....	15
D. Tables of Magnitude .....	19
E. Microdata .....	20
1. Sampling, Removing Identifiers and Limiting Geographic Detail .....	21
2. High Visibility Variables .....	21
a. Top-coding, Bottom-Coding, Recoding into Intervals .....	21
b. Adding Random Noise .....	23
c. Swapping or Rank Swapping .....	23
d. Blank and Impute for Randomly Selected Records .....	24
e. Blurring .....	24
F. Summary .....	24

## TABLE OF CONTENTS (Continued)

	Page
<b>III. Current Federal Statistical Agency Practices.....</b>	<b>25</b>
A. Agency Summaries.....	25
1. Department of Agriculture.....	25
a. Economic Research Service (ERS).....	25
b. National Agricultural Statistics Service (NASS).....	26
2. Department of Commerce.....	27
a. Bureau of Economic Analysis (BEA).....	27
b. Bureau of the Census (BOC).....	29
3. Department of Education:	
National Center for Education Statistics (NCES).....	31
4. Department of Energy:	
Energy Information Administration (EIA).....	32
5. Department of Health and Human Services.....	33
a. National Center for Health Statistics (NCHS).....	33
b. Social Security Administration (SSA).....	34
6. Department of Justice: Bureau of Justice Statistics (BJS).....	35
7. Department of Labor: Bureau of Labor Statistics (BLS).....	35
8. Department of the Treasury: Internal Revenue Service, Statistics of Income Division (IRS, SOI).....	36
9. Environmental Protection Agency (EPA).....	37
B. Summary.....	38
1. Magnitude and Frequency Data.....	38
2. Microdata.....	39
<b>IV. Methods for Tabular Data.....</b>	<b>42</b>
A. Tables of Frequency Data.....	42
1. Controlled Rounding.....	43
2. The Confidentiality Edit.....	44
B. Tables of Magnitude Data.....	44
1. Definition of Sensitive Cells.....	45
a. The p-Percent Rule.....	46
b. The pq Rule.....	47
c. The (n,k) Rule.....	48
d. The Relationship Between (n,k) and p-Percent or pq Rules.....	49
2. Complementary Suppression.....	50
a. Audits of Proposed Complementary Suppression.....	51
b. Automatic Selection of Cells for Complementary Suppression.....	52
3. Information in Parameter Values.....	54
C. Technical Notes:	
Relationships Between Common Linear Sensitivity Measures.....	54

## TABLE OF CONTENTS (Continued)

	Page
<b>V. Methods for Public-Use Microdata Files.....</b>	<b>61</b>
A. Disclosure Risk of Microdata.....	62
1. Disclosure Risk and Intruders.....	62
2. Factors Contributing to Risk.....	62
3. Factors that Naturally Decrease Risk.....	63
B. Mathematical Methods of Addressing the problem.....	64
1. Proposed Measures of Risk.....	65
2. Methods of Reducing Risk by Reducing the Amount of Information Released.....	66
3. Methods of Reducing Risk by Disturbing Microdata.....	66
4. Methods of Analyzing Disturbed Microdata to Determine Usefulness.....	68
C. Necessary Procedures for Releasing Microdata Files.....	68
1. Removal of Identifiers.....	68
2. Limiting Geographic Detail.....	69
3. Top-coding of Continuous High Visibility Variables.....	69
4. Precautions for Certain Types of Microdata.....	70
a. Establishment Microdata.....	70
b. Longitudinal Microdata.....	70
c. Microdata Containing Administrative Data.....	70
d. Consideration of Potentially Matchable Files and Population Uniques.....	71
D. Stringent Methods of Limiting Disclosure Risk.....	71
1. Do Not Release the Microdata.....	71
2. Recode Data to Eliminate Uniques.....	71
3. Disturb Data to Prevent Matching to External Files.....	71
E. Conclusion.....	
<b>VI. Recommended Practices.....</b>	<b>73</b>
A. Introduction.....	73
B. Recommendations.....	74
1. General Recommendations for Tables and Microdata.....	74
2. Tables of Frequency Count Data.....	76
3. Tables of Magnitude Data.....	76
4. Microdata files.....	78

## TABLE OF CONTENTS (Continued)

	Page
<b>VII. Research Agenda .....</b>	<b>79</b>
A. Microdata .....	79
1. Defining Disclosure .....	79
2. Effects of Disclosure Limitation on Data Quality and Usefulness ....	80
a. Disturbing Data.....	80
b. More Information about Recoded Values.....	80
3. Reidentification Issues .....	80
4. Economic Microdata .....	81
5. Longitudinal Microdata.....	81
6. Contextual Variable Data.....	81
7. Implementation Issues for Microdata .....	81
B. Tabular Data.....	82
1. Effects of Disclosure Limitation on Data Quality and Usefulness ....	82
a. Frequency Count Data .....	82
b. Magnitude Data.....	82
2. Near-Optimal Cell Suppression in Two-Dimensional Tables .....	83
3. Evaluating CONFID .....	83
4. Faster Software .....	83
5. Reducing Over-suppression .....	84
C. Data Products Other Than Microdata and Tabular Data.....	84
1. Database Systems.....	85
2. Disclosure Risk in Analytic Reports.....	87

## TABLE OF CONTENTS (Continued)

Page

### Appendices

<b>A. Technical Notes: Extending Primary Suppression Rules to Other Common Situations.....</b>	<b>89</b>
1. Background .....	89
2. Extension of Disclosure Limitation Practices .....	89
a. Sample Survey Data.....	89
b. Tables Containing Imputed Data .....	90
c. Tables that Report Negative Values.....	90
d. Tables Where Differences Between Positive Values are Reported ...	90
e. Tables Reporting Net Changes (that is, Difference Between Values Reported at Different Times) .....	91
f. Tables Reporting Weighted Averages .....	91
g. Output from Statistical Models .....	91
3. Simplifying Procedures .....	91
a. Key Item Suppression .....	91
b. Preliminary and Final Data .....	91
c. Time Series Data .....	92
<b>B. Government References.....</b>	<b>93</b>
<b>C. Annotated Bibliography.....</b>	<b>94</b>