

CONTENTS

Tables, Figures, Exhibits, and Boxes	xi
Preface	xxiii
The Author	xxvii
Introduction	xxix
1 CROSS-TABULATIONS	1
What This Chapter Is About	1
Introduction to the Book via a Concrete Example	2
Cross-Tabulations	8
What This Chapter Has Shown	19
2 MORE ON TABLES	21
What This Chapter Is About	21
The Logic of Elaboration	22
Suppressor Variables	25
Additive and Interaction Effects	26
Direct Standardization	28
A Final Note on Statistical Controls Versus Experiments	43
What This Chapter Has Shown	45
3 STILL MORE ON TABLES	47
What This Chapter Is About	47
Reorganizing Tables to Extract New Information	48
When to Percentage a Table “Backwards”	50
Cross-Tabulations in Which the Dependent Variable	
Is Represented by a Mean	52
Index of Dissimilarity	58
Writing About Cross-Tabulations	61
What This Chapter Has Shown	63

4	ON THE MANIPULATION OF DATA BY COMPUTER	65
	What This Chapter Is About	65
	Introduction	66
	How Data Files Are Organized	67
	Transforming Data	72
	What This Chapter Has Shown	80
	Appendix 4.A Doing Analysis Using Stata	80
	Tips on Doing Analysis Using Stata	80
	Some Particularly Useful Stata 10.0 Commands	84
5	INTRODUCTION TO CORRELATION AND REGRESSION (ORDINARY LEAST SQUARES)	87
	What This Chapter Is About	87
	Introduction	88
	Quantifying the Size of a Relationship: Regression Analysis	89
	Assessing the Strength of a Relationship: Correlation Analysis	91
	The Relationship Between Correlation and Regression Coefficients	94
	Factors Affecting the Size of Correlation (and Regression) Coefficients	94
	Correlation Ratios	99
	What This Chapter Has Shown	102
6	INTRODUCTION TO MULTIPLE CORRELATION AND REGRESSION (ORDINARY LEAST SQUARES)	103
	What This Chapter Is About	103
	Introduction	104
	A Worked Example: The Determinants of Literacy in China	113
	Dummy Variables	120
	A Strategy for Comparisons Across Groups	124
	A Bayesian Alternative for Comparing Models	133
	Independent Validation	135
	What This Chapter Has Shown	136
7	MULTIPLE REGRESSION TRICKS: TECHNIQUES FOR HANDLING SPECIAL ANALYTIC PROBLEMS	139
	What This Chapter Is About	139
	Nonlinear Transformations	140

Testing the Equality of Coefficients	147
Trend Analysis: Testing the Assumption of Linearity	149
Linear Splines	152
Expressing Coefficients as Deviations from the Grand Mean (Multiple Classification Analysis)	164
Other Ways of Representing Dummy Variables	166
Decomposing the Difference Between Two Means	172
What This Chapter Has Shown	179
8 MULTIPLE IMPUTATION OF MISSING DATA	181
What This Chapter Is About	181
Introduction	182
A Worked Example: The Effect of Cultural Capital on Educational Attainment in Russia	187
What This Chapter Has Shown	194
9 SAMPLE DESIGN AND SURVEY ESTIMATION	195
What This Chapter Is About	195
Survey Samples	196
Conclusion	223
What This Chapter Has Shown	224
10 REGRESSION DIAGNOSTICS	225
What This Chapter Is About	225
Introduction	226
A Worked Example: Societal Differences in Status Attainment	229
Robust Regression	237
Bootstrapping and Standard Errors	238
What This Chapter Has Shown	240
11 SCALE CONSTRUCTION	241
What This Chapter Is About	241
Introduction	242
Validity	242
Reliability	243

Scale Construction	246
Errors-in-Variables Regression	258
What This Chapter Has Shown	261
12 LOG-LINEAR ANALYSIS	263
What This Chapter Is About	263
Introduction	264
Choosing a Preferred Model	265
Parsimonious Models	277
A Bibliographic Note	294
What This Chapter Has Shown	295
Appendix 12.A Derivation of the Effect Parameters	295
Appendix 12.B Introduction to Maximum Likelihood Estimation	297
Mean of a Normal Distribution	298
Log-Linear Parameters	299
13 BINOMIAL LOGISTIC REGRESSION	301
What This Chapter Is About	301
Introduction	302
Relation to Log-Linear Analysis	303
A Worked Logistic Regression Example:	
Predicting Prevalence of Armed Threats	304
A Second Worked Example: Schooling Progression Ratios in Japan	314
A Third Worked Example (Discrete-Time Hazard-Rate Models):	
Age at First Marriage	318
A Fourth Worked Example (Case-Control Models):	
Who Was Appointed to a <i>Nomenklatura</i> Position in Russia?	327
What This Chapter Has Shown	329
Appendix 13.A Some Algebra for Logs and Exponents	330
Appendix 13.B Introduction to Probit Analysis	330
14 MULTINOMIAL AND ORDINAL LOGISTIC REGRESSION AND TOBIT REGRESSION	335
What This Chapter Is About	335
Multinomial Logit Analysis	336

Ordinal Logistic Regression	342
Tobit Regression (and Allied Procedures) for Censored Dependent Variables	353
Other Models for the Analysis of Limited Dependent Variables	360
What This Chapter Has Shown	361
15 IMPROVING CAUSAL INFERENCE: FIXED EFFECTS AND RANDOM EFFECTS MODELING	363
What This Chapter Is About	363
Introduction	364
Fixed Effects Models for Continuous Variables	365
Random Effects Models for Continuous Variables	371
A Worked Example: The Determinants of Income in China	372
Fixed Effects Models for Binary Outcomes	375
A Bibliographic Note	380
What This Chapter Has Shown	380
16 FINAL THOUGHTS AND FUTURE DIRECTIONS: RESEARCH DESIGN AND INTERPRETATION ISSUES	381
What this Chapter is About	381
Research Design Issues	382
The Importance of Probability Sampling	397
A Final Note: Good Professional Practice	400
What This Chapter Has Shown	405
Appendix A: Data Descriptions and Download Locations for the Data Used in This Book	407
Appendix B: Survey Estimation with the General Social Survey	411
References	417
Index	431

TABLES, FIGURES, EXHIBITS, AND BOXES

TABLES

1.1.	Joint Frequency Distribution of Militancy by Religiosity Among Urban Negroes in the U.S., 1964.	8
1.2.	Percent Militant by Religiosity Among Urban Negroes in the U.S., 1964.	10
1.3.	Percentage Distribution of Religiosity by Educational Attainment, Urban Negroes in the U.S., 1964.	13
1.4.	Percent Militant by Educational Attainment, Urban Negroes in the U.S., 1964.	13
1.5.	Percent Militant by Religiosity and Educational Attainment, Urban Negroes in the U.S., 1964.	15
1.6.	Percent Militant by Religiosity and Educational Attainment, Urban Negroes in the U.S., 1964 (Three-Dimensional Format).	18
2.1.	Percentage Who Believe Legal Abortions Should Be Possible Under Specified Circumstances, by Religion and Education, U.S. 1965 (N = 1,368; Cell Frequencies in Parentheses).	27
2.2.	Percentage Accepting Abortion by Religion and Education (Hypothetical Data).	28
2.3.	Percent Militant by Religiosity, and Percent Militant by Religiosity Adjusting (Standardizing) for Religiosity Differences in Educational Attainment, Urban Negroes in the U.S., 1964 (N = 993).	30
2.4.	Percentage Distribution of Beliefs Regarding the Scientific View of Evolution (U.S. Adults, 1993, 1994, and 2000).	32
2.5.	Percentage Accepting the Scientific View of Evolution by Religious Denomination (N = 3,663).	33
2.6.	Percentage Accepting the Scientific View of Evolution by Level of Education.	34
2.7.	Percentage Accepting the Scientific View of Evolution by Age.	34
2.8.	Percentage Distribution of Educational Attainment by Religion.	35
2.9.	Percentage Distribution of Age by Religion.	35
2.10.	Joint Probability Distribution of Education and Age.	36

2.11.	Percentage Accepting the Scientific View of Evolution by Religion, Age, and Sex (Percentage Bases in Parentheses)	37
2.12.	Observed Proportion Accepting the Scientific View of Evolution, and Proportion Standardized for Education and Age.	39
2.13.	Percentage Distribution of Occupational Groups by Race, South African Males Age 20–69, Early 1990s (Percentages Shown Without Controls and also Directly Standardized for Racial Differences in Educational Attainment ^a ; N = 4,004).	41
2.14.	Mean Number of Chinese Characters Known (Out of 10), for Urban and Rural Residents Age 20–69, China 1996 (Means Shown Without Controls and Also Directly Standardized for Urban-Rural Differences in Distribution of Education; N = 6,081).	42
3.1.	Frequency Distribution of Acceptance of Abortion by Religion and Education, U.S. Adults, 1965 (N = 1,368).	48
3.2.	Social Origins of Nobel Prize Winners (1901–1972) and Other U.S. Elites (and, for Comparison, the Occupations of Employed Males 1900–1920).	51
3.3.	Mean Annual Income in 1979 Among Those Working Full Time in 1980, by Education and Gender, U.S. Adults (Category Frequencies Shown in Parentheses).	52
3.4.	Means and Standard Deviations of Income in 1979 by Education and Gender, U.S. Adults, 1980.	57
3.5.	Median Annual Income in 1979 Among Those Working Full Time in 1980, by Education and Gender, U.S. Adults (Category Frequencies Shown in Parentheses).	58
3.6.	Percentage Distribution Over Major Occupation Groups by Race and Sex, U.S. Labor Force, 1979 (N = 96,945).	60
5.1.	Mean Number of Positive Responses to an Acceptance of Abortion Scale (Range: 0–7), by Religion, U.S. Adults, 2006.	101
6.1.	Means, Standard Deviations, and Correlations Among Variables Affecting Knowledge of Chinese Characters, Employed Chinese Adults Age 20–69, 1996 (N = 4,802)	115
6.2.	Determinants of the Number of Chinese Characters Correctly Identified on a Ten-Item Test, Employed Chinese Adults Age 20–69, 1996 (Standard Errors in Parentheses).	116
6.3.	Coefficients of Models of Acceptance of Abortion, U.S. Adults, 1974 (Standard Errors Shown in Parentheses); N = 1,481.	127
6.4.	Goodness-of-Fit Statistics for Alternative Models of the Relationship Among Religion, Education, and Acceptance of Abortion, U.S. Adults, 1973 (N = 1,499).	136
7.1.	Demonstration That Inclusion of a Linear Term Does Not Affect Predicted Values.	153

7.2.	Coefficients for a Linear Spline Model of Trends in Years of School Completed by Year of Birth, U.S. Adults Age 25 and Older, and Comparisons with Other Models (Pooled Data for 1972–2004, $N = 39,324$).	157
7.3.	Goodness-of-Fit Statistics for Models of Knowledge of Chinese Characters by Year of Birth, Controlling for Years of Schooling, with Various Specifications of the Effect of the Cultural Revolution (Those Affected by the Cultural Revolution Are Defined as People Turning Age 11 During the Period 1966 through 1977), Chinese Adults Age 20 to 69 in 1996 ($N = 6,086$).	160
7.4.	Coefficients for Models 4, 5, and 7 Predicting Knowledge of Chinese Characters by Year of Birth, Controlling for Years of Schooling (p Values in Parentheses).	161
7.5.	Coefficients of Models of Tolerance of Atheists, U.S. Adults, 2000 to 2004 ($N = 4,299$).	165
7.6.	Design Matrices for Alternative Ways of Coding Categorical Variables (See Text for Details).	168
7.7.	Coefficients for a Model of the Determinants of Vocabulary Knowledge, U.S. Adults, 1994 ($N = 1,757$; $R^2 = .2445$; Wald Test That Categorical Variables All Equal Zero: $F_{(3,1752)} = 12.48$; $p < .0000$).	169
7.8.	Means, Standard Deviations, and Correlations for Variables Included in a Model of Educational Attainment for U.S. Adults, 1990 to 2004, by Race (Blacks Above the Diagonal, Non-Blacks Below).	176
7.9.	Coefficients of a Model of Educational Attainment, for Blacks and Non-Blacks, U.S. Adults, 1990 to 2004.	177
7.10.	Decomposition of the Difference in the Mean Years of School Completed by Non-Blacks and Blacks, U.S. Adults, 1990 to 2004.	178
8.1.	Descriptive Statistics for the Variables Used in the Analysis, Russian Adults Age Twenty-Two to Sixty-Nine in 1993 ($N = 4,685$).	189
8.2.	Comparison of Coefficients for a Model of Educational Attainment Estimated from a Casewise-Deleted Data Set [C] ($N = 2,661$) and from a Multiply Imputed Data Set [M] ($N = 4,685$), Russian Adults Age Twenty-Two to Sixty-Nine in 1993.	192
9.1.	Portion of a Table of Random Numbers.	196
9.2.	The Population Size, Cumulative Population Size, and Percentage of the Total Population Residing in Each of the Ten Largest Cities in California, 1990.	201
9.3.	Design Effects for Selected Statistics, Samples of 3,000 with Clustering (50 Counties as Primary Sampling Units, 2 Villages or	

	Neighborhoods per County, and 30 Adults Age 20 to 69 per Village or Neighborhood), With and Without Stratification, by Level of Education.	210
9.4.	Determinants of the Number of Chinese Characters Correctly Identified on a 10-Item Test, Employed Chinese Adults Age 20–69, 1996 (N = 4,802).	216
9.5.	Coefficients for Models of the Determinants of Income, U.S. Adult Women, 1994, Under Various Design Assumptions (N = 1,015).	221
9.6.	Coefficients of a Model of Educational Attainment, U.S. Adults, 1990 to 2004 (N = 15,932).	223
10.1.	Coefficients for Models of the Determinants of the Strength of the Occupation-Education Connection in Eighteen Nations.	236
11.1.	Values of Cronbach’s Alpha for Multiple-Item Scales with Various Combinations of the Number of Items and the Average Correlation Among Items.	246
11.2.	Factor Loadings for Abortion Acceptance Items Before Rotation.	253
11.3.	Abortion Factor Loadings After Varimax Rotation.	254
11.4.	Means, Standard Deviations, and Correlations Among Variables Included in Models of the Acceptance of Legal Abortion, U.S. Adults, 1984 (N = 1,459).	256
11.5.	Coefficients of Two Models Predicting Acceptance of Abortion, U.S. Adults, 1984.	256
11.6.	Mean Score on the ISEI by Level of Education, Chinese Males Age Twenty to Sixty-Nine, 1996.	259
11.7.	Coefficients of a Model of the Determinants of Political Conservatism Estimated by Conventional OLS and Errors-in-Variables Regression, U.S. Adults, 1984 (N = 1,294).	260
12.1.	Frequency Distribution of Program by Sex in a Graduate Course.	265
12.2.	Frequency Distribution of Level of Stratification by Level of Political Integration and Level of Technology, in Ninety-Two Societies.	268
12.3.	Models of the Relationship Between Technology, Political Integration, and Level of Stratification in Ninety-Two Societies.	269
12.4.	Percentage Distribution of Expected Level of Stratification by Level of Political Integration and Level of Technology, in Ninety-Two Societies (Expected Frequencies from Model 7 Are Percentaged).	272
12.5.	Frequency Distribution of Whether “A Communist Should Be Allowed to Speak in Your Community” by Schooling, Region, and Age, U.S. Adults, 1977 (N = 1,478).	273

12.6.	Goodness-of-Fit Statistics for Log-Linear Models of the Associations Among Whether a Communist Should Be Allowed to Speak in Your Community, Age, Region, and Education, U.S. Adults, 1977.	275
12.7.	Expected Percentage (from Model 8) Agreeing That “A Communist Should Be Allowed to Speak in Your Community” by Education, Age, and Region, U.S. Adults, 1977.	276
12.8.	Frequency Distribution of Voting by Race, Education, and Voluntary Association Membership.	278
12.9.	Frequency Distribution of Occupation by Father’s Occupation, Chinese Adults, 1996.	280
12.10.	Interaction Parameters for the Saturated Model Applied to Table 12.9.	282
12.11.	Goodness-of-Fit Statistics for Alternative Models of Intergenerational Occupational Mobility in China (Six-by-Six Table).	284
12.12.	Frequency Distribution of Educational Attainment by Size of Place of Residence at Age Fourteen, Chinese Adults Not Enrolled in School, 1996.	289
13.1.	Percentage Ever Threatened by a Gun, by Selected Variables, U.S. Adults, 1973 to 1994 (N = 19,260).	306
13.2.	Goodness-of-Fit Statistics for Various Models Predicting the Prevalence of Armed Threat to U.S. Adults, 1973 to 1994.	308
13.3.	Effect Parameters for Models 2 and 4 of Table 13.2.	310
13.4.	Goodness-of-Fit Statistics for Various Models of the Process of Educational Transition in Japan (Preferred Model Shown in Boldface).	315
13.5.	Effect Parameters for Model 3 of Table 13.4.	316
13.6.	Odds Ratios for a Model Predicting the Likelihood of Marriage from Age at Risk, Sex, Race, and Mother’s Education, with Interactions Between Age at Risk and the Other Variables.	323
13.7.	Coefficients for a Model of Determinants of <i>Nomenklatura</i> Membership, Russia, 1988.	328
13.B.1.	Effect Parameters for a Probit Analysis of Gun Threat (Corresponding to Models 2 and 4 of Table 13.3).	331
14.1.	Effect Parameters for a Model of the Determinants of English and Russian Language Competence in the Czech Republic, 1993 (N = 3,945). (Standard Errors in Parentheses; <i>p</i> Values in Italic.)	339
14.2.	Effect Parameters for an Ordered Logit Model of Political Party Identification, U.S. Adults, 1998 (N = 2,443).	345
14.3.	Predicted Probability Distributions of Party Identification for Black and non-Black Males Living in Large Central Cities of Non-Southern SMSAs and Earning \$40,000 to \$50,000 per Year.	349

14.4.	Effect Parameters for a Generalized Ordered Logit Model of Political Party Identification, U.S. Adults, 1998.	350
14.5.	Effect Parameters for an Ordinary Least-Squares Regression Model of Political Party Identification, U.S. Adults, 1998.	354
14.6.	Codes for Frequency of Sex in the Past Year, U.S. Adults, 2000.	356
14.7.	Alternative Estimates of a Model of Frequency of Sex, U.S. Adults, 2000 (N = 2,258). (Standard Errors in Parentheses; All Coefficients Are Significant at .001 or Beyond.)	357
15.1.	Socioeconomic Characteristics of Chinese Adults by Size of Place of Residence, 1996.	373
15.2.	Comparison of OLS and FE Estimates for a Model of the Determinants of Family Income, Chinese RMB, 1996 (N = 5,342).	374
15.3.	Comparison of OLS and FE Estimates for a Model of the Effect of Migration and Remittances on South African Black Children's School Enrollment, 2002 to 2003. (N(FE) = 2,408 Children; N(full RE) = 12,043 Children.)	379

FIGURES

2.1.	The Observed Association Between X and Y Is Entirely Spurious and Goes to Zero When Z Is Controlled.	23
2.2.	The Observed Association Between X and Y Is Partly Spurious: the Effect of X on Y Is Reduced When Z Is Controlled (Z Affects X and Both Z and X Affect Y).	23
2.3.	The Observed Association Between X and Y Is Entirely Explained by the Intervening Variable Z and Goes to Zero When Z Is Controlled.	24
2.4.	The Observed Association Between X and Y Is Partly Explained by the Intervening Variable Z: the Effect of X on Y Is Reduced When Z Is Controlled (X Affects Z, and Both X and Z Affect Y).	24
2.5.	Both X and Z Affect Y, but There Is no Assumption Regarding the Causal Ordering of X and Z.	25
2.6.	The Size of the Zero-Order Association Between X and Y (and Between Z and Y) Is Suppressed When the Effects of X on Z and Y have Opposite Sign, and the Effects of X and Z on Y have Opposite Sign.	26
4.1.	An IBM punch card.	71
5.1.	Scatter Plot of Years of Schooling by Father's Years of Schooling (Hypothetical Data, N = 10).	88
5.2.	Least-Squares Regression Line of the Relation Between Years of Schooling and Father's Years of Schooling.	89

5.3.	Least-squares Regression Line of the Relation Between Years of Schooling and Father's Years of Schooling, Showing How the "Error of Prediction" or "Residual" Is Defined.	90
5.4.	Least-Squares Regression Lines for Three Configurations of Data: (a) Perfect Independence, (b) Perfect Correlation, and (c) Perfect Curvilinear Correlation—a Parabola Symmetrical to the X-Axis.	92
5.5.	The Effect of a Single Deviant Case (High Leverage Point).	95
5.6.	Truncating Distributions Reduces Correlations.	97
5.7.	The Effect of Aggregation on Correlations.	99
6.1.	Three-Dimensional Representation of the Relationship Between Number of Siblings, Father's Years of Schooling, and Respondent's Years of Schooling (Hypothetical Data; N = 10).	105
6.2.	Expected Number of Chinese Characters Identified (Out of Ten) by Years of Schooling and Gender, Urban Origin Chinese Adults Age 20 to 69 in 1996 with Nonmanual Occupations and with Years of Father's Schooling and Level of Cultural Capital Set at Their Means (N = 4,802). (Note: the female line does not extend beyond 16 because there are no females in the sample with post-graduate education.)	120
6.3.	Acceptance of Abortion by Education and Religious Denomination, U.S. Adults, 1974 (N = 1,481).	131
7.1.	The Relationship Between 2003 Income and Age, U.S. Adults Age Twenty to Sixty-Four in 2004 (N = 1,573).	141
7.2.	Expected ln(Income) by Years Of School Completed, U.S. Males and Females, 2004, with Hours Worked per Week Fixed at the Mean for Both Sexes Combined (42.7; N = 1,459).	144
7.3.	Expected Income by Years of School Completed, U.S. Males and Females, 2004, with Hours Worked per Week Fixed at the Mean for Both Sexes Combined (42.7).	145
7.4.	Trend in Attitudes Regarding Gender Equality, U.S. Adults Surveyed in 1974 Through 1998 (Linear Trend and Annual Means; N = 21,464).	151
7.5.	Years of School Completed by Year of Birth, U.S. Adults (Pooled Samples from the 1972 Through 2004 GSS; N = 39,324; Scatter Plot Shown for 5 Percent Sample).	154
7.6.	Mean Years of Schooling by Year of Birth, U.S. adults (Same Data as for Figure 7.5).	155
7.7.	Three-Year Moving Average of Years of Schooling by Year of Birth, U.S. Adults (Same Data as for Figure 7.5).	155
7.8.	Trend in Years of School Completed by Year of Birth, U.S. Adults (Same Data as for Figure 7.5). Predicted Values from a Linear Spline with a Knot at 1947.	158

7.9.	Graphs of Three Models of the Effect of the Cultural Revolution on Vocabulary Knowledge, Holding Constant Education (at Twelve Years), Chinese Adults, 1996 (N = 6,086).	163
7.10.	Figure 7.9 Rescaled to Show the Entire Range of the Y-Axis.	163
10.1.	Four Scatter Plots with Identical Lines.	226
10.2.	Scatter Plot of the Relationship Between X and Y and Also the Regression Line from a Model That Incorrectly Assumes a Linear Relationship Between X and Y (Hypothetical Data).	227
10.3.	Years of School Completed by Number of Siblings, U.S. Adults, 1994 (N = 2,992).	228
10.4.	Years of School Completed by Number of Siblings, U.S. Adults, 1994.	228
10.5.	A Plot of Leverage Versus Squared Normalized Residuals for Equation 7 in Treiman and Yip (1989).	232
10.6.	A Plot of Leverage Versus Studentized Residuals for Treiman and Yip's Equation 7, with Circles Proportional to the Size of Cook's D.	233
10.7.	Added-Variable Plots for Treiman and Yip's Equation 7.	233
10.8.	Residual-Versus-Fitted Plot for Treiman and Yip's Equation 7.	234
10.9.	Augmented Component-Plus-Residual Plots for Treiman and Yip's Equation 7.	235
10.10.	Objective Functions for Three M Estimators: (a) OLS Objective Function, (b) Huber Objective Function, and (c) Bi-Square Objective Function.	238
10.11.	Sampling Distributions of Bootstrapped Coefficients (2,000 Repetitions) for the Expanded Model, Estimated by Robust Regression on Seventeen Countries.	240
11.1.	Loadings of the Seven Abortion-Acceptance Items on the First Two Factors, Unrotated and Rotated 30 Degrees Counterclockwise.	255
13.1.	Expected Probability of Marrying for the First Time by Age at Risk, U.S. Adults, 1994 (N = 1,556).	320
13.2.	Expected Probability of Marrying for the First Time by Age at Risk (Range: Fifteen to Thirty-Six), Discrete-Time Model, U.S. Adults, 1994.	322
13.3.	Expected Probability of Marrying for the First Time by Age at Risk (Range: Fifteen to Thirty-Six), Polynomial Model, U.S. Adults, 1994.	322
13.4.	Expected Probability of Marrying for the First Time by Age at Risk, Sex, and Mother's Education (Twelve and Sixteen Years of Schooling), Non-Black U.S. Adults, 1994.	326
13.5.	Expected Probability of Marrying for the First Time by Age at Risk, Sex, and Mother's Education (Twelve and Sixteen Years of Schooling), Black U.S. Adults, 1994.	326

13.B.1.	Probabilities Associated with Values of Probit and Logit Coefficients.	334
14.1.	Three Estimates of the Expected Frequency of Sex per Year, U.S. Married Women, 2000 (N = 552).	358
14.2.	Expected Frequency of Sex Per Year by Gender and Marital Status, U.S. Adults, 2000 (N = 2,258).	359
16.1.	1980 Male Disability by Quarter of Birth (Prevented from Work by a Physical Disability).	386
16.2.	Blau and Duncan's Basic Model of the Process of Stratification.	394

EXHIBITS

4.1.	Illustration of How Data Files Are Organized.	67
4.2.	A Codebook Corresponding to Exhibit 4.1.	68

BOXES

Open-Ended Questions	3
Samuel A. Stouffer	6
Technical Points on Table 1.1	9
Technical Points on Table 1.2	10
Technical Points on Table 1.3	14
Technical Points on Table 1.4	15
Technical Points on Table 1.5	16
Technical Points on Table 1.6	18
Paul Lazarsfeld	22
Hans Zeisel	27
Stata -do- Files and -log- Files	30
Direct Standardization In Earlier Survey Research	31
The Weakness of Matching and a Useful Fix	44
Technical Points on Table 3.3	53
Substantive Points On Table 3.3	54
A Historical Note on Social Science Computer Packages	66
Herman Hollerith	70
The Way Things Were	72
Treating Missing Values as If They Were Not	75

XX Tables, Figures, Exhibits, and Boxes

People Generally Like to Respond to (Well-Designed and Well-Administered) Surveys	77
Why Use the “Least Squares” Criterion to Determine the Best-Fitting Line?	91
Karl Pearson	93
A Useful Computational Formula for r	93
A “Real Data” Example of the Effect of Truncating the Distribution	97
A Useful Computational Formula for η^2	101
Multicollinearity	108
Reminder Regarding the Variance of Dichotomous Variables	110
A Formula for Computing R^2 from Correlations	111
Adjusted R^2	111
Always Present Descriptive Statistics	114
Technical Point on Table 6.2	117
Why You Should Include the Entire Sample in Your Analysis	122
Getting p -values via Stata	125
Using Stata to Compare the Goodness-of-fit of Regression Models	125
R. A. (Ronald Aylmer) Fisher	126
How to Test the Significance of the Difference Between Two Coefficients	129
Alternative Ways to Estimate BIC	134
Why the Relationship Between Income and Age Is Curvilinear	140
A Trick to Reduce Collinearity	145
In Some Years of the GSS, Only a Subset of Respondents Was Asked Certain Questions	150
An Alternative Specification of Spline Functions	156
Why Black versus Non-black Is Better Than White versus Non-white for Social Analysis in the United States	175
A Comment on Credit in Science	175
Why Pairwise Deletion Should Be Avoided	183
Technical Details on the Variables	188
Telephone Surveys	198
Mail Surveys	199
Web Surveys	200
Philip M. Hauser	202
A Superior Sampling Procedure	205

Sources of Nonresponse	207
Leslie Kish	208
How the Chinese Stratified Sample Used in the Design Experiments was Constructed	212
Weighting Data in Stata	213
Limitations of the Stata 10.0 Survey Estimation Procedure	215
An Alternative to Survey Estimation	219
How to Downweight Sample Size in Stata	219
Ways to Assess Reliability	244
Why the SAT and GRE Tests Include Several Hundred Items	245
Transforming Variables so That “High” has a Consistent Meaning	248
Constructing Scales from Incomplete Information	249
In Log-Linear Analysis “Interaction” Simply Means “Association”	264
L^2 Defined	267
Other Software for Estimating Log-Linear Models	294
Maximum Likelihood Estimation	302
Probit Analysis	302
Technical Point on Table 13.1	305
Limitations of Wald Tests	309
Smoothing Distributions	325
Estimating Generalized Order Logit Models With Stata	349
James Tobin	354
Panel Surveys in the Public Domain	369
Otis Dudley Duncan	395
Sewell Wright	396
Ask a Foreigner To Do It	397
George Peter Murdock	398
In the United States, Publicly Funded Studies Must be Made Available to the Research Community	401
An “Available from Author” Archive	404