

Contents

Figures	<i>page</i> xv
Tables	xvii
Acknowledgments	xix
1 The Scientific Study of Politics	1
Overview	1
1.1 Political <i>Science</i> ?	1
1.2 Approaching Politics Scientifically: The Search for Causal Explanations	3
1.3 Thinking about the World in Terms of Variables and Causal Explanations	7
1.4 Models of Politics	14
1.5 Rules of the Road to Scientific Knowledge about Politics	15
1.5.1 Make Your Theories Causal	15
1.5.2 Don't Let Data Alone Drive Your Theories	16
1.5.3 Consider Only Empirical Evidence	17
1.5.4 Avoid Normative Statements	17
1.5.5 Pursue Both Generality and Parsimony	18
1.6 A Quick Look Ahead	18
Concepts Introduced in This Chapter	19
Exercises	20
2 The Art of Theory Building	22
Overview	22
2.1 Good Theories Come from Good Theory-Building Strategies	22
2.2 Identifying Interesting Variation	23
2.2.1 Time-Series Example	24
2.2.2 Cross-Sectional Example	25
2.3 Learning to Use Your Knowledge	26
2.3.1 Moving from a Specific Event to More General Theories	26

viii **Contents**

2.3.2	Know Local, Think Global: Can You Drop the Proper Nouns?	27
2.4	Examine Previous Research	28
2.4.1	What Did the Previous Researchers Miss?	29
2.4.2	Can Their Theory Be Applied Elsewhere?	29
2.4.3	If We Believe Their Findings, Are There Further Implications?	30
2.4.4	How Might This Theory Work at Different Levels of Aggregation (Micro \longleftrightarrow Macro)?	30
2.5	Think Formally about the Causes That Lead to Variation in Your Dependent Variable	31
2.5.1	Utility and Expected Utility	32
2.5.2	The Puzzle of Turnout	34
2.6	Think about the Institutions: The Rules Usually Matter	36
2.6.1	Legislative Rules	36
2.6.2	The Rules Matter!	38
2.7	Extensions	39
2.8	How Do I Know If I Have a “Good” Theory?	40
2.8.1	Is Your Theory Causal?	40
2.8.2	Can You Test Your Theory on Data That You Have Not Yet Observed?	41
2.8.3	How General Is Your Theory?	41
2.8.4	How Parsimonious Is Your Theory?	41
2.8.5	How New Is Your Theory?	41
2.8.6	How Nonobvious Is Your Theory?	42
2.9	Conclusion	42
	Concepts Introduced in This Chapter	43
	Exercises	43
3	Evaluating Causal Relationships	45
	Overview	45
3.1	Causality and Everyday Language	45
3.2	Four Hurdles along the Route to Establishing Causal Relationships	48
3.2.1	Putting It All Together – Adding Up the Answers to Our Four Questions	50
3.2.2	Identifying Causal Claims Is an Essential Thinking Skill	50
3.2.3	What Are the Consequences of Failing to Control for Other Possible Causes?	53
3.3	Why Is Studying Causality So Important? Three Examples from Political Science	54
3.3.1	Life Satisfaction and Democratic Stability	54
3.3.2	School Choice and Student Achievement	55
3.3.3	Electoral Systems and the Number of Political Parties	57

ix

Contents

3.4

Why Is Studying Causality So Important? Three Examples from Everyday Life

61

3.4.1

Alcohol Consumption and Income

61

3.4.2

Treatment Choice and Breast Cancer Survival

62

3.4.3

Explicit Lyrics and Teen Sexual Behavior

63

3.5

Wrapping Up

65

Concepts Introduced in This Chapter

65

Exercises

65

4

Research Design

67

Overview

67

4.1

Comparison as the Key to Establishing Causal Relationships

67

4.2

Experimental Research Designs

68

4.2.1

“Random Assignment” versus “Random Sampling”

74

4.2.2

Are There Drawbacks to Experimental Research Designs?

74

4.3

Observational Studies (in Two Flavors)

77

4.3.1

Datum, Data, Data Set

79

4.3.2

Cross-Sectional Observational Studies

81

4.3.3

Time-Series Observational Studies

82

4.3.4

The Major Difficulty with Observational Studies

83

4.4

Summary

83

Concepts Introduced in This Chapter

84

Exercises

84

5

Measurement

86

Overview

86

5.1

Why Measurement Matters

86

5.2

Social Science Measurement: The Varying Challenges of Quantifying Humanity

88

5.3

Problems in Measuring Concepts of Interest

91

5.3.1

Conceptual Clarity

91

5.3.2

Reliability

92

5.3.3

Measurement Bias and Reliability

93

5.3.4

Validity

94

5.3.5

The Relationship between Validity and Reliability

95

5.4

Controversy 1: Measuring Democracy

96

5.5

Controversy 2: Measuring Political Tolerance

99

5.6

Are There Consequences to Poor Measurement?

101

5.7

Conclusions

101

Concepts Introduced in This Chapter

102

Exercises

102

6

Descriptive Statistics and Graphs

104

Overview

104

6.1

Know Your Data

104

x	Contents	
	6.2 What Is the Variable’s Measurement Metric?	105
	6.2.1 Categorical Variables	106
	6.2.2 Ordinal Variables	106
	6.2.3 Continuous Variables	107
	6.2.4 Variable Types and Statistical Analyses	108
	6.3 Describing Categorical Variables	109
	6.4 Describing Continuous Variables	110
	6.4.1 Rank Statistics	111
	6.4.2 Moments	114
	6.5 Limitations	118
	Concepts Introduced in This Chapter	118
	Exercises	118
7	Statistical Inference	120
	Overview	120
	7.1 Populations and Samples	120
	7.2 Learning about the Population from a Sample: The Central Limit Theorem	122
	7.2.1 The Normal Distribution	122
	7.3 Example: Presidential Approval Ratings	128
	7.3.1 What Kind of Sample Was That?	129
	7.3.2 A Note on the Effects of Sample Size	130
	7.4 A Look Ahead: Examining Relationships between Variables	131
	Concepts Introduced in This Chapter	132
	Exercises	132
8	Bivariate Hypothesis Testing	134
	Overview	134
	8.1 Bivariate Hypothesis Tests and Establishing Causal Relationships	134
	8.2 Choosing the Right Bivariate Hypothesis Test	135
	8.3 All Roads Lead to <i>p</i>	136
	8.3.1 The Logic of <i>p</i> -Values	136
	8.3.2 The Limitations of <i>p</i> -Values	137
	8.3.3 From <i>p</i> -Values to Statistical Significance	138
	8.3.4 The Null Hypothesis and <i>p</i> -Values	138
	8.4 Three Bivariate Hypothesis Tests	139
	8.4.1 Example 1: Tabular Analysis	139
	8.4.2 Example 2: Difference of Means	145
	8.4.3 Example 3: Correlation Coefficient	150
	8.5 Wrapping Up	155
	Concepts Introduced in This Chapter	156
	Exercises	157
9	Bivariate Regression Models	159
	Overview	159
	9.1 Two-Variable Regression	159
	9.2 Fitting a Line: Population ⇔ Sample	160

xi Contents

9.3	Which Line Fits Best? Estimating the Regression Line	162
9.4	Measuring Our Uncertainty about the OLS Regression Line	165
9.4.1	Goodness-of-Fit: Root Mean-Squared Error	167
9.4.2	Goodness-of-Fit: R-Squared Statistic	167
9.4.3	Is That a “Good” Goodness-of-Fit?	169
9.4.4	Uncertainty about Individual Components of the Sample Regression Model	169
9.4.5	Confidence Intervals about Parameter Estimates	171
9.4.6	Hypothesis Testing: Overview	172
9.4.7	Two-Tailed Hypothesis Tests	173
9.4.8	The Relationship between Confidence Intervals and Two-Tailed Hypothesis Tests	175
9.4.9	One-Tailed Hypothesis Tests	175
9.5	Assumptions, More Assumptions, and Minimal Mathematical Requirements	177
9.5.1	Assumptions about the Population Stochastic Component	177
9.5.2	Assumptions about Our Model Specification	180
9.5.3	Minimal Mathematical Requirements	181
9.5.4	How Can We Make All of These Assumptions?	181
	Concepts Introduced in This Chapter	182
	Exercises	182
10	Multiple Regression Models I: The Basics	183
	Overview	183
10.1	Modeling Multivariate Reality	183
10.2	The Population Regression Function	184
10.3	From Two-Variable to Multiple Regression	184
10.4	What Happens When We Fail to Control for Z?	188
10.4.1	An Additional Minimal Mathematical Requirement in Multiple Regression	192
10.5	Interpreting Multiple Regression	193
10.6	Which Effect Is “Biggest”?	196
10.7	Statistical and Substantive Significance	198
10.8	Implications	199
	Concepts Introduced in This Chapter	200
	Exercises	200
11	Multiple Regression Models II: Crucial Extensions	202
	Overview	202
11.1	Extensions of OLS	202
11.2	Being Smart with Dummy Independent Variables in OLS	203
11.2.1	Using Dummy Variables to Test Hypotheses about a Categorical Independent Variable with Only Two Values	203
11.2.2	Using Dummy Variables to Test Hypotheses about a Categorical Independent Variable with More Than Two Values	207

xii	Contents	
	11.3 Testing Interactive Hypotheses with Dummy Variables	210
	11.4 Dummy Dependent Variables	212
	11.4.1 The Linear Probability Model	212
	11.4.2 Binomial Logit and Binomial Probit	215
	11.4.3 Goodness-of-Fit with Dummy Dependent Variables	219
	11.5 Outliers and Influential Cases in OLS	220
	11.5.1 Identifying Influential Cases	221
	11.5.2 Dealing with Influential Cases	224
	11.6 Multicollinearity	225
	11.6.1 How Does Multicollinearity Happen?	226
	11.6.2 Detecting Multicollinearity	227
	11.6.3 Multicollinearity: A Simulated Example	228
	11.6.4 Multicollinearity: A Real-World Example	230
	11.6.5 Multicollinearity: What Should I Do?	232
	11.7 Being Careful with Time Series	233
	11.7.1 Time-Series Notation	233
	11.7.2 Memory and Lags in Time-Series Analysis	234
	11.7.3 Trends and the Spurious Regression Problem	236
	11.7.4 The Differenced Dependent Variable	239
	11.7.5 The Lagged Dependent Variable	241
	11.8 Wrapping Up	242
	Concepts Introduced in This Chapter	243
12	Multiple Regression Models III: Applications	244
	Overview	244
	12.1 Why Controlling for Z Matters	244
	12.2 Example 1: The Economy and Presidential Popularity	245
	12.3 Example 2: Politics, Economics, and Public Support for Democracy	248
	12.4 Example 3: Competing Theories of How Politics Affects International Trade	251
	12.5 Conclusions	253
	Concepts Introduced in This Chapter	254
	Exercises	254
	Appendix A. Critical Values of χ^2	255
	Appendix B. Critical Values of t	256
	Appendix C. The Λ Link Function for BNL Models	257
	Appendix D. The Φ Link Function for BNP Models	259
	Bibliography	261
	Index	265