

SUB Hamburg



B/123639

THE FUNDAMENTALS OF

# **Political Science Research**

*Second Edition*

**Paul M. Kellstedt**

Texas A&M University

**Guy D. Whitten**

Texas A&M University



**CAMBRIDGE**  
UNIVERSITY PRESS



# Contents

List of Figures	<i>page</i> xv
List of Tables	xvii
Preface to the Second Edition	xix
Acknowledgments to the Second Edition	xxi
Acknowledgments to the First Edition	xxiii



<b>The Scientific Study of Politics . . . . .</b>	<b>1</b>
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	15
1.5 Rules of the Road to Scientific Knowledge about Politics	16
1.5.1 Make Your Theories Causal	16
1.5.2 Don't Let Data Alone Drive Your Theories	17
1.5.3 Consider Only Empirical Evidence	18
1.5.4 Avoid Normative Statements	18
1.5.5 Pursue Both Generality and Parsimony	19
1.6 A Quick Look Ahead	19
Concepts Introduced in This Chapter	20
Exercises	21



<b>The Art of Theory Building . . . . .</b>	<b>24</b>
Overview	24
2.1 Good Theories Come from Good Theory-Building Strategies	24
2.2 Promising Theories Offer Answers to Interesting Research Questions	25
2.3 Identifying Interesting Variation	26

2.3.1	Time-Series Example	27
2.3.2	Cross-Sectional Example	28
2.4	Learning to Use Your Knowledge	29
2.4.1	Moving from a Specific Event to More General Theories	30
2.4.2	Know Local, Think Global: Can You Drop the Proper Nouns?	31
2.5	Examine Previous Research	32
2.5.1	What Did the Previous Researchers Miss?	32
2.5.2	Can Their Theory Be Applied Elsewhere?	33
2.5.3	If We Believe Their Findings, Are There Further Implications?	33
2.5.4	How Might This Theory Work at Different Levels of Aggregation (Micro $\iff$ Macro)?	34
2.6	Think Formally about the Causes That Lead to Variation in Your Dependent Variable	34
2.6.1	Utility and Expected Utility	36
2.6.2	The Puzzle of Turnout	38
2.7	Think about the Institutions: The Rules Usually Matter	40
2.7.1	Legislative Rules	40
2.7.2	The Rules Matter!	42
2.8	Extensions	43
2.9	How Do I Know If I Have a “Good” Theory?	44
2.9.1	Does Your Theory Offer an Answer to an Interesting Research Question?	44
2.9.2	Is Your Theory Causal?	44
2.9.3	Can You Test Your Theory on Data That You Have Not Yet Observed?	45
2.9.4	How General Is Your Theory?	45
2.9.5	How Parsimonious Is Your Theory?	45
2.9.6	How New Is Your Theory?	46
2.9.7	How Nonobvious Is Your Theory?	46
2.10	Conclusion	46
	Concepts Introduced in This Chapter	47
	Exercises	47
<b>3</b>	<b>Evaluating Causal Relationships . . . . .</b>	<b>51</b>
	Overview	51
3.1	Causality and Everyday Language	51
3.2	Four Hurdles Along the Route to Establishing Causal Relationships	54
3.2.1	Putting It All Together – Adding Up the Answers to Our Four Questions	56
3.2.2	Identifying Causal Claims Is an Essential Thinking Skill	57
3.2.3	What Are the Consequences of Failing to Control for Other Possible Causes?	61

3.3	Why Is Studying Causality So Important? Three Examples from Political Science	62
3.3.1	Life Satisfaction and Democratic Stability	62
3.3.2	Race and Political Participation in the United States	63
3.3.3	Evaluating Whether Head Start Is Effective	64
3.4	Wrapping Up	66
	Concepts Introduced in This Chapter	66
	Exercises	66

**4**

<b>Research Design</b>	69	
Overview	69	
4.1	Comparison as the Key to Establishing Causal Relationships	69
4.2	Experimental Research Designs	70
4.2.1	“Random Assignment” versus “Random Sampling”	76
4.2.2	Varieties of Experiments and Near-Experiments	77
4.2.3	Are There Drawbacks to Experimental Research Designs?	78
4.3	Observational Studies (in Two Flavors)	82
4.3.1	Datum, Data, Data Set	84
4.3.2	Cross-Sectional Observational Studies	85
4.3.3	Time-Series Observational Studies	86
4.3.4	The Major Difficulty with Observational Studies	87
4.4	Summary	88
	Concepts Introduced in This Chapter	89
	Exercises	90

**5**

<b>Getting to Know Your Data: Evaluating Measurement and Variations</b>	92	
Overview	92	
5.1	Getting to Know Your Data	93
5.2	Social Science Measurement: The Varying Challenges of Quantifying Humanity	95
5.3	Problems in Measuring Concepts of Interest	98
5.3.1	Conceptual Clarity	99
5.3.2	Reliability	99
5.3.3	Measurement Bias and Reliability	100
5.3.4	Validity	101
5.3.5	The Relationship between Validity and Reliability	102
5.4	Controversy 1: Measuring Democracy	103
5.5	Controversy 2: Measuring Political Tolerance	107
5.6	Are There Consequences to Poor Measurement?	109
5.7	Getting to Know Your Data Statistically	109
5.8	What Is the Variable’s Measurement Metric?	111
5.8.1	Categorical Variables	111
5.8.2	Ordinal Variables	112
5.8.3	Continuous Variables	113

5.8.4	Variable Types and Statistical Analyses	114
5.9	Describing Categorical Variables	114
5.10	Describing Continuous Variables	116
5.10.1	Rank Statistics	117
5.10.2	Moments	120
5.11	Limitations of Descriptive Statistics and Graphs	122
5.12	Conclusions	124
	Concepts Introduced in This Chapter	124
	Exercises	126

## **6 Probability and Statistical Inference . . . . . 129**

	Overview	129
6.1	Populations and Samples	129
6.2	Some Basics of Probability Theory	131
6.3	Learning about the Population from a Sample: The Central Limit Theorem	133
6.3.1	The Normal Distribution	134
6.4	Example: Presidential Approval Ratings	139
6.4.1	What Kind of Sample Was That?	140
6.4.2	A Note on the Effects of Sample Size	141
6.5	A Look Ahead: Examining Relationships between Variables	142
	Concepts Introduced in This Chapter	143
	Exercises	143

## **7 Bivariate Hypothesis Testing . . . . . 145**

	Overview	145
7.1	Bivariate Hypothesis Tests and Establishing Causal Relationships	145
7.2	Choosing the Right Bivariate Hypothesis Test	146
7.3	All Roads Lead to $p$	147
7.3.1	The Logic of $p$ -Values	147
7.3.2	The Limitations of $p$ -Values	148
7.3.3	From $p$ -Values to Statistical Significance	149
7.3.4	The Null Hypothesis and $p$ -Values	150
7.4	Three Bivariate Hypothesis Tests	150
7.4.1	Example 1: Tabular Analysis	150
7.4.2	Example 2: Difference of Means	156
7.4.3	Example 3: Correlation Coefficient	161
7.5	Wrapping Up	167
	Concepts Introduced in This Chapter	167
	Exercises	168

## **8 Bivariate Regression Models . . . . . 171**

	Overview	171
8.1	Two-Variable Regression	171
8.2	Fitting a Line: Population $\Leftrightarrow$ Sample	172

8.3	Which Line Fits Best? Estimating the Regression Line	174
8.4	Measuring Our Uncertainty about the OLS Regression Line	178
8.4.1	Goodness-of-Fit: Root Mean-Squared Error	179
8.4.2	Goodness-of-Fit: R-Squared Statistic	179
8.4.3	Is That a “Good” Goodness-of-Fit?	181
8.4.4	Uncertainty about Individual Components of the Sample Regression Model	182
8.4.5	Confidence Intervals about Parameter Estimates	183
8.4.6	Two-Tailed Hypothesis Tests	185
8.4.7	The Relationship between Confidence Intervals and Two-Tailed Hypothesis Tests	187
8.4.8	One-Tailed Hypothesis Tests	188
8.5	Assumptions, More Assumptions, and Minimal Mathematical Requirements	189
8.5.1	Assumptions about the Population Stochastic Component	190
8.5.2	Assumptions about Our Model Specification	192
8.5.3	Minimal Mathematical Requirements	193
8.5.4	How Can We Make All of These Assumptions?	194
	Concepts Introduced in This Chapter	194
	Exercises	195

<b>9</b>	<b>Multiple Regression: The Basics</b> . . . . .	197
	Overview	197
9.1	Modeling Multivariate Reality	197
9.2	The Population Regression Function	198
9.3	From Two-Variable to Multiple Regression	198
9.4	Interpreting Multiple Regression	202
9.5	Which Effect Is “Biggest”?	206
9.6	Statistical and Substantive Significance	207
9.7	What Happens When We Fail to Control for Z?	209
9.7.1	An Additional Minimal Mathematical Requirement in Multiple Regression	213
9.8	An Example from the Literature: Competing Theories of How Politics Affects International Trade	214
9.9	Implications	216
	Concepts Introduced in This Chapter	217
	Exercises	218

<b>10</b>	<b>Multiple Regression Model Specification</b> . . . . .	220
	Overview	220
10.1	Extensions of OLS	220
10.2	Being Smart with Dummy Independent Variables in OLS	220
10.2.1	Using Dummy Variables to Test Hypotheses about a Categorical Independent Variable with Only Two Values	221

10.2.2	Using Dummy Variables to Test Hypotheses about a Categorical Independent Variable with More Than Two Values	225
10.2.3	Using Dummy Variables to Test Hypotheses about Multiple Independent Variables	228
10.3	Testing Interactive Hypotheses with Dummy Variables	229
10.4	Outliers and Influential Cases in OLS	232
10.4.1	Identifying Influential Cases	233
10.4.2	Dealing with Influential Cases	236
10.5	Multicollinearity	238
10.5.1	How Does Multicollinearity Happen?	238
10.5.2	Detecting Multicollinearity	239
10.5.3	Multicollinearity: A Simulated Example	240
10.5.4	Multicollinearity: A Real-World Example	242
10.5.5	Multicollinearity: What Should I Do?	244
10.6	Wrapping Up	244
	Concepts Introduced in This Chapter	245
	Exercises	245

## 11

<b>Limited Dependent Variables and Time-Series Data . . . .</b>	<b>247</b>	
Overview	247	
11.1	Extensions of OLS	247
11.2	Dummy Dependent Variables	248
11.2.1	The Linear Probability Model	248
11.2.2	Binomial Logit and Binomial Probit	251
11.2.3	Goodness-of-Fit with Dummy Dependent Variables	254
11.3	Being Careful with Time Series	256
11.3.1	Time-Series Notation	256
11.3.2	Memory and Lags in Time-Series Analysis	257
11.3.3	Trends and the Spurious Regression Problem	259
11.3.4	The Differenced Dependent Variable	262
11.3.5	The Lagged Dependent Variable	263
11.4	Example: The Economy and Presidential Popularity	265
11.5	Wrapping Up	269
	Concepts Introduced in This Chapter	269
	Exercises	270

## 12

<b>Putting It All Together to Produce Effective Research . . . .</b>	<b>273</b>	
Overview	273	
12.1	Two Routes toward a New Scientific Project	273
12.1.1	Project Type 1: A New Y (and Some X)	274
12.1.2	Project Type 2: An Existing Y and a New X	275
12.1.3	Variants on the Two Project Types	276
12.2	Using the Literature without Getting Buried in It	278
12.2.1	Identifying the Important Work on a Subject – Using Citation Counts	278

12.2.2	Oh No! Someone Else Has Already Done What I Was Planning to Do. What Do I Do Now?	279
12.2.3	Dissecting the Research by Other Scholars	279
12.2.4	Read Effectively to Write Effectively	280
12.3	Writing Effectively about Your Research	281
12.3.1	Write Early, Write Often because Writing Is Thinking	281
12.3.2	Document Your Code – Writing and Thinking while You Compute	282
12.3.3	Divide and Conquer – a Section-by-Section Strategy for Building Your Project	282
12.3.4	Proofread, Proofread, and Then Proofread Again	285
12.4	Making Effective Use of Tables and Figures	286
12.4.1	Constructing Regression Tables	287
12.4.2	Writing about Regression Tables	290
12.4.3	Other Types of Tables and Figures	292
	Exercises	293
	<b>Appendix A. Critical Values of Chi-Square . . . . .</b>	<b>295</b>
	<b>Appendix B. Critical Values of <math>t</math> . . . . .</b>	<b>296</b>
	<b>Appendix C. The A Link Function for Binomial Logit Models</b>	<b>297</b>
	<b>Appendix D. The <math>\Phi</math> Link Function for Binomial Probit Models</b>	<b>299</b>
	Bibliography	301
	Index	307