

CONTENTS

List of Figures	xii
List of Tables	xvii
Useful Commands for Stata	xxii
Useful Commands for R	xxiv
Preface for Students: How This Book Can Help You Learn Econometrics	xxvi
Preface for Instructors: How to Help Your Students Learn Econometrics	xxx
Acknowledgments	xxxvi

1 The Quest for Causality 1

1.1 The Core Model	2
1.2 Two Major Challenges: Randomness and Endogeneity	7
CASE STUDY: Flu Shots	13
CASE STUDY: Country Music and Suicide	15
1.3 Randomized Experiments as the Gold Standard	18
Conclusion 22 • Key Terms 23	

2 Stats in the Wild: Good Data Practices 24

2.1 Know Our Data	26
2.2 Replication	28
CASE STUDY: Violent Crime in the United States	31
2.3 Statistical Software	32
Conclusion 33 • Further Reading 34 • Key Terms 34 • Computing Corner 34 • Exercises 39	

I THE OLS FRAMEWORK 43

3 Bivariate OLS: The Foundation of Econometric Analysis 45

3.1 Bivariate Regression Model	47
3.2 Random Variation in Coefficient Estimates	53

3.3	Endogeneity and Bias	57
3.4	Precision of Estimates	61
3.5	Probability Limits and Consistency	65
3.6	Solvable Problems: Heteroscedasticity and Correlated Errors	67
3.7	Goodness of Fit	70
	CASE STUDY: Height and Wages	74
3.8	Outliers	77
	Conclusion 80 • Further Reading 81 • Key Terms 82 • Computing Corner 82 • Exercises 86	

4 Hypothesis Testing and Interval Estimation: Answering Research Questions 91

4.1	Hypothesis Testing	92
4.2	t Tests	98
4.3	p Values	106
4.4	Power	109
4.5	Straight Talk about Hypothesis Testing	115
4.6	Confidence Intervals	117
	Conclusion 120 • Further Reading 120 • Key Terms 121 • Computing Corner 121 • Exercises 123	

5 Multivariate OLS: Where the Action Is 127

5.1	Using Multivariate OLS to Fight Endogeneity	129
5.2	Omitted Variable Bias	137
	CASE STUDY: Does Education Support Economic Growth?	140
5.3	Measurement Error	143
5.4	Precision and Goodness of Fit	146
	CASE STUDY: Institutions and Human Rights	152
5.5	Standardized Coefficients	155
5.6	Hypothesis Testing about Multiple Coefficients	158
	CASE STUDY: Comparing Effects of Height Measures	164
	Conclusion 166 • Further Reading 168 • Key Terms 168 • Computing Corner 168 • Exercises 172	

6 Dummy Variables: Smarter than You Think 179

6.1	Using Bivariate OLS to Assess Difference of Means	180
	CASE STUDY: Sex Differences in Heights	187

6.2	Dummy Independent Variables in Multivariate OLS	190
6.3	Transforming Categorical Variables to Multiple Dummy Variables	193
	CASE STUDY: When Do Countries Tax Wealth?	197
6.4	Interaction Variables	202
	CASE STUDY: Energy Efficiency	207
	Conclusion 211 • Further Reading 212 • Key Terms 212	
	• Computing Corner 212 • Exercises 214	

7 Specifying Models 220

7.1	Quadratic and Polynomial Models	221
	CASE STUDY: Global Warming	227
7.2	Logged Variables	230
7.3	Post-Treatment Variables	236
7.4	Model Specification	243
	Conclusion 245 • Further Reading 245 • Key Terms 246	
	• Computing Corner 246 • Exercises 247	

II THE CONTEMPORARY ECONOMETRIC TOOLKIT 253

8 Using Fixed Effects Models to Fight Endogeneity in Panel Data and Difference-in-Difference Models 255

8.1	The Problem with Pooling	256
8.2	Fixed Effects Models	261
8.3	Working with Fixed Effects Models	267
8.4	Two-Way Fixed Effects Model	271
	CASE STUDY: Trade and Alliances	274
8.5	Difference-in-Difference	276
	Conclusion 282 • Further Reading 283 • Key Terms 284	
	• Computing Corner 284 • Exercises 288	

9 Instrumental Variables: Using Exogenous Variation to Fight Endogeneity 295

9.1	2SLS Example	296
9.2	Two-Stage Least Squares (2SLS)	298
	CASE STUDY: Emergency Care for Newborns	305
9.3	Multiple Instruments	309

9.4	Quasi and Weak Instruments	310
9.5	Precision of 2SLS	313
9.6	Simultaneous Equation Models	315
	CASE STUDY: Supply and Demand Curves for the Chicken Market	319
	Conclusion 323 • Further Reading 324 • Key Terms 325	
	• Computing Corner 325 • Exercises 327	

10 Experiments: Dealing with Real-World Challenges 333

10.1	Randomization and Balance	335
	CASE STUDY: Development Aid and Balancing	338
10.2	Compliance and Intention-to-Treat Models	340
10.3	Using 2SLS to Deal with Non-compliance	346
	CASE STUDY: Minneapolis Domestic Violence Experiment	350
10.4	Attrition	354
	CASE STUDY: Health Insurance and Attrition	357
10.5	Natural Experiments	360
	CASE STUDY: Crime and Terror Alerts	362
	Conclusion 363 • Further Reading 365 • Key Terms 365	
	• Computing Corner 365 • Exercises 366	

11 Regression Discontinuity: Looking for Jumps in Data 373

11.1	Basic RD Model	375
11.2	More Flexible RD Models	380
11.3	Windows and Bins	386
	CASE STUDY: Universal Prekindergarten	389
11.4	Limitations and Diagnostics	391
	CASE STUDY: Alcohol and Grades	395
	Conclusion 397 • Further Reading 398 • Key Terms 398	
	• Computing Corner 399 • Exercises 400	

III LIMITED DEPENDENT VARIABLES 407

12 Dummy Dependent Variables 409

12.1	Linear Probability Model	410
------	--------------------------	-----

12.2 Using Latent Variables to Explain Observed Variables	414
12.3 Probit and Logit Models	418
12.4 Estimation	423
12.5 Interpreting Probit and Logit Coefficients	426
CASE STUDY: Econometrics in the Grocery Store	431
12.6 Hypothesis Testing about Multiple Coefficients	436
CASE STUDY: Civil Wars	440
Conclusion 443 • Further Reading 443 • Key Terms 444	
• Computing Corner 444 • Exercises 449	

IV ADVANCED MATERIAL 457

13 Time Series: Dealing with Stickiness over Time 459

13.1 Modeling Autocorrelation	460
13.2 Detecting Autocorrelation	463
13.3 Fixing Autocorrelation	467
CASE STUDY: Using an AR(1) Model to Study Global Temperature Changes	471
13.4 Dynamic Models	473
13.5 Stationarity	476
CASE STUDY: Dynamic Model of Global Temperature	482
Conclusion 486 • Further Reading 487 • Key Terms 488	
• Computing Corner 488 • Exercises 490	

14 Advanced OLS 493

14.1 How to Derive the OLS Estimator and Prove Unbiasedness	493
14.2 How to Derive the Equation for the Variance of $\hat{\beta}_1$	499
14.3 Calculating Power	501
14.4 How to Derive the Omitted Variable Bias Conditions	502
14.5 Anticipating the Sign of Omitted Variable Bias	505
14.6 Omitted Variable Bias with Multiple Variables	507
14.7 Omitted Variable Bias due to Measurement Error	508

14.8 Collider Bias with Post-Treatment Variables	510
Conclusion 513 • Further Reading 514 • Key Term 514	
• Computing Corner 514 • Exercises 515	

15 Advanced Panel Data 518

15.1 Panel Data Models with Serially Correlated Errors	518
15.2 Temporal Dependence with a Lagged Dependent Variable	520
15.3 Random Effects Models	524
Conclusion 526 • Further Reading 526 • Key Term 527	
• Computing Corner 527 • Exercises 530	

16 Conclusion: How to Be an Econometric Realist 533

Further Reading 537

APPENDICES

Math and Probability Background 538

A	Summation	538
B	Expectation	538
C	Variance	539
D	Covariance	540
E	Correlation	541
F	Probability Density Functions	541
G	Normal Distributions	543
H	Other Useful Distributions	549
I	Sampling	551

Further Reading 554 • Key Terms 554 • Computing Corner 554

Citations and Additional Notes 556

Guide to Review Questions 567

Bibliography 577

Photo Credits 586

Glossary 587

Index 596
