

Contents

Preface vii
Acknowledgments x
About the Authors xi

1 Introduction to Econometrics	1
What is Econometrics?	2
Step 1: What Do You Want to Do?	3
Step 2: Formulate Your Research Design and Specify the Econometric Model	5
Step 3: Apply Statistical Theory	7
An Illustration with the Population Mean	9
Putting It Together: Poverty and Test Scores	10
From Statistics to Econometrics	12
2 Simple Regression	19
The Least-Squares Criterion	20
Estimating a Simple Regression Model of Academic Performance	25
The R^2	26
Beyond Simple Regression	31
3 Multiple Regression	33
The Multiple Regression Model	34
When Does the MR Estimator Collapse to the SR Estimator?	37
Back to API and FLE (and PE, Too)	38
From Two to Many Explanatory Variables: The General Multiple Regression Model	43
Perfect Multicollinearity	45
Interpreting MR Coefficients	46
Using Matrix Algebra to Estimate a Multiple Regression Model	47
4 Generalizing from a Sample	51
Three Steps to Generalizing from a Sample	51
Step 1: Define Your Population and Research Goal	52

Step 2: Make Assumptions about Your Population (and How Your Sample Represents It)	54
Step 3: Compute Statistics to Measure OLS Accuracy	57
Data Types	58
5 Properties of Our Estimators	65
An Experiment in Random Sampling	65
BLUE Estimators	69
The Sample Mean	69
The OLS Regression	71
Multiple Regression	76
Back to Schools and Free Lunch	77
Consistent Estimators	79
Properties of the Regression Line (or Plan, or Hyperplane)	80
6 Hypothesis Testing and Confidence Intervals	83
Hypothesis Testing	83
Hypothesis Testing with Small Samples	86
Confidence Intervals	87
What Determines Confidence Interval Width?	88
How Large a Sample is Large Enough?	89
Hypothesis Testing and Confidence Intervals in Multiple Regressions	91
Presenting Your Results	97
7 Predicting in a Nonlinear World	99
Getting the Model Right	100
How Could We Get the Model Wrong?	102
What If the True Model is Nonlinear?	104
Testing Functional Forms	117
8 Best of BLUE I: Cross-Section Data and Heteroskedasticity (Assumption CR2)	121
The Problem of Heteroskedasticity	122
Testing for Heteroskedasticity	127

Fixing the Problem	130
Back to Climate Change	132
9 Best of Blue II: Correlated Errors (Assumption CR3)	137
Time-Series Data: Autocorrelation	139
Ignoring Autocorrelation	140
How to Test for Autocorrelation	142
Fixing Autocorrelation	146
An Ex Post Error-Correction Method: Newey-West	147
Ex Ante: Using What We Know to Improve on OLS	149
Clustered and Spatial Data	156
10 Sample Selection Bias (Assumption CR1)	159
What If We Have a Nonrepresentative Sample?	160
How Does It Happen?	161
How Sample Selection Bias Affects Regression Models	162
What to Do About It	166
Self-Selection	166
Experiments	172
11 Identifying Causation	177
Why Care About Causation?	178
A New Classical Regression Assumption	179
The Endogeneity Problem	179
Measurement Error	181
X and Y Jointly Determined	184
Omitted Variables	187
12 Instrumental Variables: A Solution to the Endogeneity Problem	193
What Are Instrumental Variables?	193
The Quest for Good Instruments	199
The Gold Standard: Randomized Treatments	208
A Brave New World	211
Appendix: Critical Values for Commonly Used Tests in Econometrics	215
Notes	221
Index	223