

## **Real Econometrics**

The Right Tools to Answer Important Questions

Second Edition

**Michael Bailey**

### **Table of Contents**

List of Figures
List of Tables
Useful Commands for Stata
Useful Commands for R
Preface for Students: How This Book Can Help You Learn Econometrics
Preface for Instructors: How to Help Your Students Learn Econometrics
Acknowledgments
1 The Quest for Causality
The Core Model
Two Challenges: Randomness and Endogeneity
CASE STUDY: Flu Shots
CASE STUDY: Country Music and Suicides
Randomized Experiments as the Gold Standard
2 Stats in the Wild: Good Data Practices
2.1 Know Our Data
2.2 Replication
CASE STUDY: Violent Crime in the United States
2.3 Statistical Software
I The OLS FRAMEWORK
3 Bivariate OLS: The Foundation of Econometric Analysis
3.1 Bivariate Regression Model
3.2 Random Variation in Coefficient Estimates
3.3 Exogeneity and Unbiasedness
3.4 Precision of Estimates
3.5 Probability Limits and Consistency
3.6 Solvable Problems: Heteroscedasticity and Correlated Errors
3.7 Goodness of Fit
CASE STUDY: Height and Wages
3.8 Outliers
4 Hypothesis Testing and Interval Estimation: Answering Research Questions
4.1 Hypothesis Testing
4.2 t Tests
4.3 p Values
4.4 Power
4.5 Straight Talk about Hypothesis Testing
4.6 Confidence Intervals
5 Multivariate OLS: Where the Action Is
5.1 Using Multivariate OLS to Fight Endogeneity
5.2 Omitted Variable Bias
CASE STUDY: Does Education Support Economic Growth?
5.3 Measurement Error
5.4 Precision and Goodness of Fit
CASE STUDY: Institutions and Human Rights
5.5 Model Specification
6 Dummy Variables: Smarter Than You Think
6.1 Using Bivariate OLS to Assess Difference of Means
CASE STUDY: Sex Differences in Heights
6.2 Dummy Independent Variables in Multivariate OLS
6.3 Transforming Categorical Variables to Multiple Dummy Variables
CASE STUDY: When Do Countries Tax Wealth?
6.4 Interaction Variables
CASE STUDY: Energy Efficiency
7 Specifying Models
7.1 Quadratic and Polynomial Models
CASE STUDY: Global Warming
7.2 Logged Variables
7.3 Standardized Coefficients
7.4 Hypothesis Testing about Multiple Coefficients

## CASE STUDY: Comparing Effects of Height Measures

### II THE CONTEMPORARY ECONOMETRIC TOOLKIT

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

##### 8.1 The Problem with Pooling

##### 8.2 Fixed Effects Models

##### 8.3 Working with Fixed Effects Models

##### 8.4 Two-Way Fixed Effects Model

#### CASE STUDY: Trade and Alliances

##### 8.5 Difference-in-Difference

#### 9 Instrumental Variables: Using Exogenous Variation to Fight Endogeneity

##### 9.1 2SLS Example

##### 9.2 Two-Stage Least Squares (2SLS)

#### CASE STUDY: Emergency Care for Newborns

##### 9.3 Multiple Instruments

##### 9.4 Quasi and Weak Instruments

##### 9.5 Precision of 2SLS

##### 9.6 Simultaneous Equation Models

#### CASE STUDY: Supply and Demand Curves for the Chicken Market

#### 10 Experiments: Dealing with Real-World Challenges

##### 10.1 Randomization and Balance

#### CASE STUDY: Development Aid and Balancing

##### 10.2 Compliance and Intention-to-Treat Models

##### 10.3 Using 2SLS to Deal with Non-compliance

#### CASE STUDY: Minneapolis Domestic Violence Experiment

##### 10.4 Attrition

#### CASE STUDY: Health Insurance and Attrition

##### 10.5 Natural Experiments

#### CASE STUDY: Crime and Terror Alerts 354

#### 11 Regression Discontinuity: Looking for Jumps in Data

##### 11.1 Basic RD Model

##### 11.2 More Flexible RD Models

##### 11.3 Windows and Bins

#### CASE STUDY: Universal Prekindergarten

##### 11.4 Limitations and Diagnostics

#### CASE STUDY: Alcohol and Grades

### III LIMITED DEPENDENT VARIABLES

#### 12 Dummy Dependent Variables

##### 12.1 Linear Probability Model

##### 12.2 Using Latent Variables to Explain Observed Variables

##### 12.3 Probit and Logit Models

##### 12.4 Estimation

##### 12.5 Interpreting Probit and Logit Coefficients

#### CASE STUDY: Econometrics in the Grocery Store

##### 12.6 Hypothesis Testing about Multiple Coefficients

#### CASE STUDY: Civil Wars

### IV ADVANCED MATERIAL

#### 13 Time Series: Dealing with Stickiness over Time

##### 13.1 Modeling Autocorrelation

##### 13.2 Detecting Autocorrelation

##### 13.3 Fixing Autocorrelation

#### CASE STUDY: Using an AR(1) Model to Study Global Temperature Changes

##### 13.4 Dynamic Models

##### 13.5 Stationarity

#### CASE STUDY: Dynamic Model of Global Temperature

#### 14 Advanced OLS

##### 14.1 How to Derive the OLS Estimator and Prove Unbiasedness

##### 14.2 How to Derive the Equation for the Variance of $O\hat{\beta}_1$

##### 14.3 How to Derive the Omitted Variable Bias Conditions

##### 14.4 Anticipating the Sign of Omitted Variable Bias

##### 14.5 Omitted Variable Bias with Multiple Variables

##### 14.6 Omitted Variable Bias due to Measurement Error

#### 15 Advanced Panel Data

##### 15.1 Panel Data Models with Serially Correlated Errors

##### 15.2 Temporal Dependence with a Lagged Dependent Variable

##### 15.3 Random Effects Models

## 16 Conclusion: How to Be an Econometric Realist

### APPENDICES

Math and Probability Background

A. Summation

B. Expectation

C. Variance

D. Covariance

E. Correlation

F. Probability Density Functions

G. Normal Distributions

H. Other Useful Distributions

I. Sampling

Citations and Additional Notes

Guide to Review Questions

Bibliography

Photo Credits

Glossary

Index