

# Introduction to Econometrics

# The Pearson Series in Economics

- Abel/Bernanke/Croushore**  
*Macroeconomics\**
- Acemoglu/Laibson/List**  
*Economics\**
- Bade/Parkin**  
*Foundations of Economics\**
- Berck/Helfand**  
*The Economics of the Environment*
- Bierman/Fernandez**  
*Game Theory with Economic Applications*
- Blair/Rush**  
*The Economics of Managerial Decisions\**
- Blanchard**  
*Macroeconomics\**
- Boyer**  
*Principles of Transportation Economics*
- Branson**  
*Macroeconomic Theory and Policy*
- Bruce**  
*Public Finance and the American Economy*
- Carlton/Perloff**  
*Modern Industrial Organization*
- Case/Fair/Oster**  
*Principles of Economics\**
- Chapman**  
*Environmental Economics: Theory, Application, and Policy*
- Daniels/VanHoose**  
*International Monetary & Financial Economics*
- Downs**  
*An Economic Theory of Democracy*
- Farnham**  
*Economics for Managers*
- Froyen**  
*Macroeconomics: Theories and Policies*
- Fusfeld**  
*The Age of the Economist*
- Gerber**  
*International Economics\**
- Gordon**  
*Macroeconomics\**
- Greene**  
*Econometric Analysis*
- Gregory/Stuart**  
*Russian and Soviet Economic Performance and Structure*
- Hartwick/Olewiler**  
*The Economics of Natural Resource Use*
- Heilbroner/Milberg**  
*The Making of the Economic Society*
- Heyne/Boettke/Prychitko**  
*The Economic Way of Thinking*
- Hubbard/O'Brien**  
*Economics\**
- InEcon*
- Money, Banking, and the Financial System\**
- Hubbard/O'Brien/Rafferty**  
*Macroeconomics\**
- Hughes/Cain**  
*American Economic History*
- Husted/Melvin**  
*International Economics*
- Jehle/Reny**  
*Advanced Microeconomic Theory*
- Keat/Young/Erffle**  
*Managerial Economics*
- Klein**  
*Mathematical Methods for Economics*
- Krugman/Obstfeld/Melitz**  
*International Economics: Theory & Policy\**
- Laidler**  
*The Demand for Money*
- Lynn**  
*Economic Development: Theory and Practice for a Divided World*
- Miller**  
*Economics Today\**
- Miller/Benjamin**  
*The Economics of Macro Issues*
- Miller/Benjamin/North**  
*The Economics of Public Issues*
- Mishkin**  
*The Economics of Money, Banking, and Financial Markets\**
- The Economics of Money, Banking, and Financial Markets, Business School Edition\**
- Macroeconomics: Policy and Practice\**
- Murray**  
*Econometrics: A Modern Introduction*
- O'Sullivan/Sheffrin/Perez**  
*Economics: Principles, Applications and Tools\**
- Parkin**  
*Economics\**
- Perloff**  
*Microeconomics\**
- Microeconomics: Theory and Applications with Calculus\**
- Perloff/Brander**  
*Managerial Economics and Strategy\**
- Pindyck/Rubinfeld**  
*Microeconomics\**
- Riddell/Shackelford/Stamos/Schneider**  
*Economics: A Tool for Critically Understanding Society*
- Roberts**  
*The Choice: A Fable of Free Trade and Protection*
- Scherer**  
*Industry Structure, Strategy, and Public Policy*
- Schiller**  
*The Economics of Poverty and Discrimination*
- Sherman**  
*Market Regulation*
- Stock/Watson**  
*Introduction to Econometrics*
- Studenmund**  
*Using Econometrics: A Practical Guide*
- Todaro/Smith**  
*Economic Development*
- Walters/Walters/Appel/Callahan/Centanni/Maex/O'Neill**  
*Econversations: Today's Students Discuss Today's Issues*
- Williamson**  
*Macroeconomics*

\*denotes **MyLab Economics** titles. Visit [www.pearson.com/mylab/economics](http://www.pearson.com/mylab/economics) to learn more.

# Introduction to Econometrics

FOURTH EDITION

**James H. Stock**

Harvard University

**Mark W. Watson**

Princeton University



New York, NY

Vice President, Business, Economics, and UK Courseware:  
Donna Battista  
Director of Portfolio Management: Adrienne D'Ambrosio  
Specialist Portfolio Manager: David Alexander  
Editorial Assistant: Nicole Nedwidek  
Vice President, Product Marketing: Roxanne McCarley  
Product Marketing Assistant: Mariana Silvestri  
Manager of Field Marketing, Business Publishing:  
Adam Goldstein  
Executive Field Marketing Manager: Carlie Marvel  
Vice President, Production and Digital Studio,  
Arts and Business: Etain O'Dea  
Director, Production and Digital Studio, Business  
and Economics: Ashley Santora  
Managing Producer, Business: Alison Kalil  
Content Producer: Christine Donovan  
Operations Specialist: Carol Melville

Design Lead: Kathryn Foot  
Manager, Learning Tools: Brian Surette  
Senior Learning Tools Strategist: Emily Biberger  
Managing Producer, Digital Studio and GLP: James Bateman  
Managing Producer, Digital Studio: Diane Lombardo  
Digital Studio Producer: Melissa Honig  
Digital Studio Producer: Alana Coles  
Digital Content Team Lead: Noel Lotz  
Digital Content Project Lead: Noel Lotz  
Project Manager: Rose Kernan, Cengage Publisher Services  
Interior Design: Cengage Publisher Services  
Cover Design: Studio Montage  
Cover Art: Courtesy of authors  
Printer/Binder: LSC Communications, Inc./Kendallville  
Cover Printer: Phoenix Color/Terre Haute

About the cover: The cover shows a time series plot of 72 indicators of real economic activity in the United States beginning in 1959. The plot shows the growth of these variables since 1959 and their (roughly) synchronized downturns associated with recessions. These series are a subset of the 131-variable dataset used in Chapter 17 to construct dynamic factor model forecasts of future growth in real GDP.

Copyright © 2019, 2015, 2011 by Pearson Education, Inc. or its affiliates. All Rights Reserved. Manufactured in the United States of America. This publication is protected by copyright, and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise. For information regarding permissions, request forms, and the appropriate contacts within the Pearson Education Global Rights and Permissions department, please visit [www.pearsoned.com/permissions/](http://www.pearsoned.com/permissions/).

Acknowledgments of third-party content appear on the appropriate page within the text.

PEARSON, ALWAYS LEARNING, and MYLAB are exclusive trademarks owned by Pearson Education, Inc. or its affiliates in the U.S. and/or other countries.

Unless otherwise indicated herein, any third-party trademarks, logos, or icons that may appear in this work are the property of their respective owners, and any references to third-party trademarks, logos, icons, or other trade dress are for demonstrative or descriptive purposes only. Such references are not intended to imply any sponsorship, endorsement, authorization, or promotion of Pearson's products by the owners of such marks, or any relationship between the owner and Pearson Education, Inc., or its affiliates, authors, licensees, or distributors.

#### **Library of Congress Cataloging-in-Publication Data**

Names: Stock, James H., author. | Watson, Mark W., author.  
Title: Introduction to econometrics / James H. Stock, Harvard University,  
Mark W. Watson, Princeton University.  
Description: Fourth edition. | New York, NY : Pearson, [2019] | Series:  
The Pearson series in economics | Includes bibliographical references and index.  
Identifiers: LCCN 2018035117 | ISBN 9780134461991  
Subjects: LCSH: Econometrics.  
Classification: LCC HB139 .S765 2019 | DDC 330.01/5195—dc23  
LC record available at <https://lcn.loc.gov/2018035117>



ISBN-10: 0-13-446199-1  
ISBN-13: 978-0-13-446199-1

# Brief Contents

## **PART ONE Introduction and Review**

---

- Chapter 1 Economic Questions and Data 1
- Chapter 2 Review of Probability 13
- Chapter 3 Review of Statistics 61

## **PART TWO Fundamentals of Regression Analysis**

---

- Chapter 4 Linear Regression with One Regressor 101
- Chapter 5 Regression with a Single Regressor: Hypothesis Tests and Confidence Intervals 136
- Chapter 6 Linear Regression with Multiple Regressors 169
- Chapter 7 Hypothesis Tests and Confidence Intervals in Multiple Regression 205
- Chapter 8 Nonlinear Regression Functions 235
- Chapter 9 Assessing Studies Based on Multiple Regression 288

## **PART THREE Further Topics in Regression Analysis**

---

- Chapter 10 Regression with Panel Data 319
- Chapter 11 Regression with a Binary Dependent Variable 350
- Chapter 12 Instrumental Variables Regression 385
- Chapter 13 Experiments and Quasi-Experiments 432
- Chapter 14 Prediction with Many Regressors and Big Data 472

## **PART FOUR Regression Analysis of Economic Time Series Data**

---

- Chapter 15 Introduction to Time Series Regression and Forecasting 512
- Chapter 16 Estimation of Dynamic Causal Effects 567
- Chapter 17 Additional Topics in Time Series Regression 607

## **PART FIVE Regression Analysis of Economic Time Series Data**

---

- Chapter 18 The Theory of Linear Regression with One Regressor 645
- Chapter 19 The Theory of Multiple Regression 671



# Contents

Preface xxvii

## PART ONE Introduction and Review

---

### CHAPTER 1 Economic Questions and Data 1

#### 1.1 Economic Questions We Examine 1

Question #1: Does Reducing Class Size Improve Elementary School Education? 1

Question #2: Is There Racial Discrimination in the Market for Home Loans? 2

Question #3: How Much Do Cigarette Taxes Reduce Smoking? 3

Question #4: By How Much Will U.S. GDP Grow Next Year? 4

Quantitative Questions, Quantitative Answers 4

#### 1.2 Causal Effects and Idealized Experiments 5

Estimation of Causal Effects 5

Prediction, Forecasting, and Causality 6

#### 1.3 Data: Sources and Types 6

Experimental versus Observational Data 7

Cross-Sectional Data 7

Time Series Data 8

Panel Data 9

### CHAPTER 2 Review of Probability 13

#### 2.1 Random Variables and Probability Distributions 14

Probabilities, the Sample Space, and Random Variables 14

Probability Distribution of a Discrete Random Variable 14

Probability Distribution of a Continuous Random Variable 16

#### 2.2 Expected Values, Mean, and Variance 18

The Expected Value of a Random Variable 18

The Standard Deviation and Variance 19

Mean and Variance of a Linear Function of a Random Variable 20

Other Measures of the Shape of a Distribution 21

Standardized Random Variables 23

#### 2.3 Two Random Variables 23

Joint and Marginal Distributions 23

Conditional Distributions 24

Independence 28

Covariance and Correlation 28

The Mean and Variance of Sums of Random Variables 29

- 2.4 The Normal, Chi-Squared, Student  $t$ , and  $F$  Distributions 33
  - The Normal Distribution 33
  - The Chi-Squared Distribution 38
  - The Student  $t$  Distribution 38
  - The  $F$  Distribution 38
- 2.5 Random Sampling and the Distribution of the Sample Average 39
  - Random Sampling 39
  - The Sampling Distribution of the Sample Average 40
- 2.6 Large-Sample Approximations to Sampling Distributions 43
  - The Law of Large Numbers and Consistency 43
  - The Central Limit Theorem 44
  - APPENDIX 2.1 Derivation of Results in Key Concept 2.3 58
  - APPENDIX 2.2 The Conditional Mean as the Minimum Mean Squared Error Predictor 59

### CHAPTER 3 Review of Statistics 61

- 3.1 Estimation of the Population Mean 62
  - Estimators and Their Properties 62
  - Properties of  $\bar{Y}$  64
  - The Importance of Random Sampling 65
- 3.2 Hypothesis Tests Concerning the Population Mean 66
  - Null and Alternative Hypotheses 67
  - The  $p$ -Value 67
  - Calculating the  $p$ -Value When  $\sigma_Y$  Is Known 68
  - The Sample Variance, Sample Standard Deviation, and Standard Error 69
  - Calculating the  $p$ -Value When  $\sigma_Y$  Is Unknown 71
  - The  $t$ -Statistic 71
  - Hypothesis Testing with a Prespecified Significance Level 72
  - One-Sided Alternatives 74
- 3.3 Confidence Intervals for the Population Mean 75
- 3.4 Comparing Means from Different Populations 77
  - Hypothesis Tests for the Difference Between Two Means 77
  - Confidence Intervals for the Difference Between Two Population Means 78
- 3.5 Differences-of-Means Estimation of Causal Effects Using Experimental Data 79
  - The Causal Effect as a Difference of Conditional Expectations 79
  - Estimation of the Causal Effect Using Differences of Means 79
- 3.6 Using the  $t$ -Statistic When the Sample Size Is Small 81
  - The  $t$ -Statistic and the Student  $t$  Distribution 83
  - Use of the Student  $t$  Distribution in Practice 84



- 3.7 Scatterplots, the Sample Covariance, and the Sample Correlation 85
  - Scatterplots 85
  - Sample Covariance and Correlation 85
  - APPENDIX 3.1 The U.S. Current Population Survey 99
  - APPENDIX 3.2 Two Proofs That  $\bar{Y}$  Is the Least Squares Estimator of  $\mu_Y$  99
  - APPENDIX 3.3 A Proof That the Sample Variance Is Consistent 100

## PART TWO Fundamentals of Regression Analysis

---

### CHAPTER 4 Linear Regression with One Regressor 101

- 4.1 The Linear Regression Model 102
- 4.2 Estimating the Coefficients of the Linear Regression Model 105
  - The Ordinary Least Squares Estimator 106
  - OLS Estimates of the Relationship Between Test Scores and the Student-Teacher Ratio 107
  - Why Use the OLS Estimator? 109
- 4.3 Measures of Fit and Prediction Accuracy 111
  - The  $R^2$  111
  - The Standard Error of the Regression 112
  - Prediction Using OLS 113
  - Application to the Test Score Data 113
- 4.4 The Least Squares Assumptions for Causal Inference 114
  - Assumption 1: The Conditional Distribution of  $u_i$  Given  $X_i$  Has a Mean of Zero 115
  - Assumption 2:  $(X_i, Y_i), i = 1, \dots, n$ , Are Independently and Identically Distributed 116
  - Assumption 3: Large Outliers Are Unlikely 117
  - Use of the Least Squares Assumptions 118
- 4.5 The Sampling Distribution of the OLS Estimators 119
- 4.6 Conclusion 122
  - APPENDIX 4.1 The California Test Score Data Set 130
  - APPENDIX 4.2 Derivation of the OLS Estimators 130
  - APPENDIX 4.3 Sampling Distribution of the OLS Estimator 131
  - APPENDIX 4.4 The Least Squares Assumptions for Prediction 134

### CHAPTER 5 Regression with a Single Regressor: Hypothesis Tests and Confidence Intervals 136

- 5.1 Testing Hypotheses About One of the Regression Coefficients 136
  - Two-Sided Hypotheses Concerning  $\beta_1$  137
  - One-Sided Hypotheses Concerning  $\beta_1$  140
  - Testing Hypotheses About the Intercept  $\beta_0$  142
- 5.2 Confidence Intervals for a Regression Coefficient 142

- 5.3 Regression When  $X$  Is a Binary Variable 144
  - Interpretation of the Regression Coefficients 144
- 5.4 Heteroskedasticity and Homoskedasticity 146
  - What Are Heteroskedasticity and Homoskedasticity? 146
  - Mathematical Implications of Homoskedasticity 148
  - What Does This Mean in Practice? 150
- 5.5 The Theoretical Foundations of Ordinary Least Squares 152
  - Linear Conditionally Unbiased Estimators and the Gauss–Markov Theorem 152
  - Regression Estimators Other Than OLS 153
- 5.6 Using the  $t$ -Statistic in Regression When the Sample Size Is Small 154
  - The  $t$ -Statistic and the Student  $t$  Distribution 154
  - Use of the Student  $t$  Distribution in Practice 155
- 5.7 Conclusion 155
  - [APPENDIX 5.1](#) Formulas for OLS Standard Errors 164
  - [APPENDIX 5.2](#) The Gauss–Markov Conditions and a Proof of the Gauss–Markov Theorem 165

## CHAPTER 6 Linear Regression with Multiple Regressors 169

- 6.1 Omitted Variable Bias 169
  - Definition of Omitted Variable Bias 170
  - A Formula for Omitted Variable Bias 172
  - Addressing Omitted Variable Bias by Dividing the Data into Groups 173
- 6.2 The Multiple Regression Model 175
  - The Population Regression Line 175
  - The Population Multiple Regression Model 176
- 6.3 The OLS Estimator in Multiple Regression 177
  - The OLS Estimator 178
  - Application to Test Scores and the Student–Teacher Ratio 179
- 6.4 Measures of Fit in Multiple Regression 180
  - The Standard Error of the Regression ( $SER$ ) 180
  - The  $R^2$  181
  - The Adjusted  $R^2$  181
  - Application to Test Scores 182
- 6.5 The Least Squares Assumptions for Causal Inference in Multiple Regression 183
  - Assumption 1: The Conditional Distribution of  $u_i$ , Given  $X_{1i}, X_{2i}, \dots, X_{ki}$ , Has a Mean of 0 183
  - Assumption 2:  $(X_{1i}, X_{2i}, \dots, X_{ki}, Y_i), i = 1, \dots, n$ , Are i.i.d. 183
  - Assumption 3: Large Outliers Are Unlikely 183
  - Assumption 4: No Perfect Multicollinearity 184

- 6.6 The Distribution of the OLS Estimators in Multiple Regression 185
- 6.7 Multicollinearity 186
  - Examples of Perfect Multicollinearity 186
  - Imperfect Multicollinearity 188
- 6.8 Control Variables and Conditional Mean Independence 189
  - Control Variables and Conditional Mean Independence 190
- 6.9 Conclusion 192
  - [APPENDIX 6.1](#) Derivation of Equation (6.1) 200
  - [APPENDIX 6.2](#) Distribution of the OLS Estimators When There Are Two Regressors and Homoskedastic Errors 201
  - [APPENDIX 6.3](#) The Frisch–Waugh Theorem 201
  - [APPENDIX 6.4](#) The Least Squares Assumptions for Prediction with Multiple Regressors 202
  - [APPENDIX 6.5](#) Distribution of OLS Estimators in Multiple Regression with Control Variables 203

## **CHAPTER 7 Hypothesis Tests and Confidence Intervals in Multiple Regression 205**

- 7.1 Hypothesis Tests and Confidence Intervals for a Single Coefficient 205
  - Standard Errors for the OLS Estimators 205
  - Hypothesis Tests for a Single Coefficient 206
  - Confidence Intervals for a Single Coefficient 207
  - Application to Test Scores and the Student–Teacher Ratio 207
- 7.2 Tests of Joint Hypotheses 209
  - Testing Hypotheses on Two or More Coefficients 210
  - The  $F$ -Statistic 211
  - Application to Test Scores and the Student–Teacher Ratio 213
  - The Homoskedasticity-Only  $F$ -Statistic 214
- 7.3 Testing Single Restrictions Involving Multiple Coefficients 216
- 7.4 Confidence Sets for Multiple Coefficients 217
- 7.5 Model Specification for Multiple Regression 218
  - Model Specification and Choosing Control Variables 219
  - Interpreting the  $R^2$  and the Adjusted  $R^2$  in Practice 220
- 7.6 Analysis of the Test Score Data Set 220
- 7.7 Conclusion 226
  - [APPENDIX 7.1](#) The Bonferroni Test of a Joint Hypothesis 232

**CHAPTER 8 Nonlinear Regression Functions 235**

- 8.1 A General Strategy for Modeling Nonlinear Regression Functions 237
  - Test Scores and District Income 237
  - The Effect on Y of a Change in X in Nonlinear Specifications 240
  - A General Approach to Modeling Nonlinearities Using Multiple Regression 243
- 8.2 Nonlinear Functions of a Single Independent Variable 244
  - Polynomials 244
  - Logarithms 246
  - Polynomial and Logarithmic Models of Test Scores and District Income 254
- 8.3 Interactions Between Independent Variables 255
  - Interactions Between Two Binary Variables 256
  - Interactions Between a Continuous and a Binary Variable 258
  - Interactions Between Two Continuous Variables 263
- 8.4 Nonlinear Effects on Test Scores of the Student–Teacher Ratio 268
  - Discussion of Regression Results 268
  - Summary of Findings 272
- 8.5 Conclusion 273
  - [APPENDIX 8.1 Regression Functions That Are Nonlinear in the Parameters](#) 283
  - [APPENDIX 8.2 Slopes and Elasticities for Nonlinear Regression Functions](#) 286

**CHAPTER 9 Assessing Studies Based on Multiple Regression 288**

- 9.1 Internal and External Validity 288
  - Threats to Internal Validity 289
  - Threats to External Validity 290
- 9.2 Threats to Internal Validity of Multiple Regression Analysis 291
  - Omitted Variable Bias 292
  - Misspecification of the Functional Form of the Regression Function 294
  - Measurement Error and Errors-in-Variables Bias 294
  - Missing Data and Sample Selection 297
  - Simultaneous Causality 299
  - Sources of Inconsistency of OLS Standard Errors 301
- 9.3 Internal and External Validity When the Regression Is Used for Prediction 302
- 9.4 Example: Test Scores and Class Size 303
  - External Validity 304
  - Internal Validity 310
  - Discussion and Implications 311
- 9.5 Conclusion 312
  - [APPENDIX 9.1 The Massachusetts Elementary School Testing Data](#) 318

**PART THREE Further Topics in Regression Analysis****CHAPTER 10 Regression with Panel Data 319**

- 10.1 Panel Data 320
  - Example: Traffic Deaths and Alcohol Taxes 320
- 10.2 Panel Data with Two Time Periods: “Before and After” Comparisons 323
- 10.3 Fixed Effects Regression 325
  - The Fixed Effects Regression Model 325
  - Estimation and Inference 327
  - Application to Traffic Deaths 328
- 10.4 Regression with Time Fixed Effects 329
  - Time Effects Only 329
  - Both Entity and Time Fixed Effects 330
- 10.5 The Fixed Effects Regression Assumptions and Standard Errors for Fixed Effects Regression 332
  - The Fixed Effects Regression Assumptions 332
  - Standard Errors for Fixed Effects Regression 334
- 10.6 Drunk Driving Laws and Traffic Deaths 335
- 10.7 Conclusion 339
  - APPENDIX 10.1** The State Traffic Fatality Data Set 345
  - APPENDIX 10.2** Standard Errors for Fixed Effects Regression 346

**CHAPTER 11 Regression with a Binary Dependent Variable 350**

- 11.1 Binary Dependent Variables and the Linear Probability Model 351
  - Binary Dependent Variables 351
  - The Linear Probability Model 353
- 11.2 Probit and Logit Regression 355
  - Probit Regression 355
  - Logit Regression 359
  - Comparing the Linear Probability, Probit, and Logit Models 361
- 11.3 Estimation and Inference in the Logit and Probit Models 362
  - Nonlinear Least Squares Estimation 362
  - Maximum Likelihood Estimation 363
  - Measures of Fit 364
- 11.4 Application to the Boston HMDA Data 365
- 11.5 Conclusion 371
  - APPENDIX 11.1** The Boston HMDA Data Set 379
  - APPENDIX 11.2** Maximum Likelihood Estimation 379
  - APPENDIX 11.3** Other Limited Dependent Variable Models 382

**CHAPTER 12 Instrumental Variables Regression 385**

- 12.1 The IV Estimator with a Single Regressor and a Single Instrument 386
  - The IV Model and Assumptions 386
  - The Two Stage Least Squares Estimator 387
  - Why Does IV Regression Work? 387
  - The Sampling Distribution of the TSLS Estimator 392
  - Application to the Demand for Cigarettes 393
- 12.2 The General IV Regression Model 395
  - TSLS in the General IV Model 397
  - Instrument Relevance and Exogeneity in the General IV Model 398
  - The IV Regression Assumptions and Sampling Distribution of the TSLS Estimator 399
  - Inference Using the TSLS Estimator 400
  - Application to the Demand for Cigarettes 401
- 12.3 Checking Instrument Validity 402
  - Assumption 1: Instrument Relevance 402
  - Assumption 2: Instrument Exogeneity 404
- 12.4 Application to the Demand for Cigarettes 408
- 12.5 Where Do Valid Instruments Come From? 412
  - Three Examples 413
- 12.6 Conclusion 417
  - APPENDIX 12.1** The Cigarette Consumption Panel Data Set 424
  - APPENDIX 12.2** Derivation of the Formula for the TSLS Estimator in Equation (12.4) 424
  - APPENDIX 12.3** Large-Sample Distribution of the TSLS Estimator 425
  - APPENDIX 12.4** Large-Sample Distribution of the TSLS Estimator When the Instrument Is Not Valid 426
  - APPENDIX 12.5** Instrumental Variables Analysis with Weak Instruments 427
  - APPENDIX 12.6** TSLS with Control Variables 429

**CHAPTER 13 Experiments and Quasi-Experiments 432**

- 13.1 Potential Outcomes, Causal Effects, and Idealized Experiments 433
  - Potential Outcomes and the Average Causal Effect 433
  - Econometric Methods for Analyzing Experimental Data 434
- 13.2 Threats to Validity of Experiments 436
  - Threats to Internal Validity 436
  - Threats to External Validity 439
- 13.3 Experimental Estimates of the Effect of Class Size Reductions 440
  - Experimental Design 440
  - Analysis of the STAR Data 441
  - Comparison of the Observational and Experimental Estimates of Class Size Effects 446

- 13.4 Quasi-Experiments 448
  - Examples 448
  - The Differences-in-Differences Estimator 450
  - Instrumental Variables Estimators 452
  - Regression Discontinuity Estimators 453
- 13.5 Potential Problems with Quasi-Experiments 454
  - Threats to Internal Validity 454
  - Threats to External Validity 456
- 13.6 Experimental and Quasi-Experimental Estimates in Heterogeneous Populations 456
  - OLS with Heterogeneous Causal Effects 457
  - IV Regression with Heterogeneous Causal Effects 458
- 13.7 Conclusion 461
  - [APPENDIX 13.1](#) The Project STAR Data Set 468
  - [APPENDIX 13.2](#) IV Estimation When the Causal Effect Varies Across Individuals 469
  - [APPENDIX 13.3](#) The Potential Outcomes Framework for Analyzing Data from Experiments 470

## **CHAPTER 14 Prediction with Many Regressors and Big Data 472**

- 14.1 What Is “Big Data”? 473
- 14.2 The Many-Predictor Problem and OLS 474
  - The Mean Squared Prediction Error 476
  - The First Least Squares Assumption for Prediction 477
  - The Predictive Regression Model with Standardized Regressors 477
  - The MSPE of OLS and the Principle of Shrinkage 479
  - Estimation of the MSPE 480
- 14.3 Ridge Regression 482
  - Shrinkage via Penalization and Ridge Regression 482
  - Estimation of the Ridge Shrinkage Parameter by Cross Validation 483
  - Application to School Test Scores 484
- 14.4 The Lasso 485
  - Shrinkage Using the Lasso 486
  - Application to School Test Scores 489
- 14.5 Principal Components 490
  - Principals Components with Two Variables 490
  - Principal Components with  $k$  Variables 492
  - Application to School Test Scores 494
- 14.6 Predicting School Test Scores with Many Predictors 495

- 14.7 Conclusion 500
  - APPENDIX 14.1 The California School Test Score Data Set 509
  - APPENDIX 14.2 Derivation of Equation (14.4) for  $k = 1$  509
  - APPENDIX 14.3 The Ridge Regression Estimator When  $k = 1$  509
  - APPENDIX 14.4 The Lasso Estimator When  $k = 1$  510
  - APPENDIX 14.5 Computing Out-of-Sample Predictions in the Standardized Regression Model 510

## **PART FOUR Regression Analysis of Economic Time Series Data**

---

### **CHAPTER 15 Introduction to Time Series Regression and Forecasting 512**

- 15.1 Introduction to Time Series Data and Serial Correlation 513
  - Real GDP in the United States 513
  - Lags, First Differences, Logarithms, and Growth Rates 513
  - Autocorrelation 516
  - Other Examples of Economic Time Series 518
- 15.2 Stationarity and the Mean Squared Forecast Error 519
  - Stationarity 519
  - Forecasts and Forecast Errors 520
  - The Mean Squared Forecast Error 521
- 15.3 Autoregressions 523
  - The First-Order Autoregressive Model 523
  - The  $p^{\text{th}}$ -Order Autoregressive Model 525
- 15.4 Time Series Regression with Additional Predictors and the Autoregressive Distributed Lag Model 526
  - Forecasting GDP Growth Using the Term Spread 527
  - The Autoregressive Distributed Lag Model 528
  - The Least Squares Assumptions for Forecasting with Multiple Predictors 529
- 15.5 Estimation of the MSFE and Forecast Intervals 531
  - Estimation of the MSFE 531
  - Forecast Uncertainty and Forecast Intervals 534
- 15.6 Estimating the Lag Length Using Information Criteria 536
  - Determining the Order of an Autoregression 536
  - Lag Length Selection in Time Series Regression with Multiple Predictors 539
- 15.7 Nonstationarity I: Trends 540
  - What Is a Trend? 540
  - Problems Caused by Stochastic Trends 542
  - Detecting Stochastic Trends: Testing for a Unit AR Root 544
  - Avoiding the Problems Caused by Stochastic Trends 546



15.8	Nonstationarity II: Breaks	547
	What Is a Break?	547
	Testing for Breaks	547
	Detecting Breaks Using Pseudo Out-of-Sample Forecasts	552
	Avoiding the Problems Caused by Breaks	553
15.9	Conclusion	554
	APPENDIX 15.1 Time Series Data Used in Chapter 15	562
	APPENDIX 15.2 Stationarity in the AR(1) Model	563
	APPENDIX 15.3 Lag Operator Notation	564
	APPENDIX 15.4 ARMA Models	565
	APPENDIX 15.5 Consistency of the BIC Lag Length Estimator	565
<b>CHAPTER 16</b>	<b>Estimation of Dynamic Causal Effects</b>	<b>567</b>
16.1	An Initial Taste of the Orange Juice Data	568
16.2	Dynamic Causal Effects	570
	Causal Effects and Time Series Data	570
	Two Types of Exogeneity	573
16.3	Estimation of Dynamic Causal Effects with Exogenous Regressors	575
	The Distributed Lag Model Assumptions	575
	Autocorrelated $u_t$ , Standard Errors, and Inference	576
	Dynamic Multipliers and Cumulative Dynamic Multipliers	576
16.4	Heteroskedasticity- and Autocorrelation-Consistent Standard Errors	578
	Distribution of the OLS Estimator with Autocorrelated Errors	578
	HAC Standard Errors	579
16.5	Estimation of Dynamic Causal Effects with Strictly Exogenous Regressors	582
	The Distributed Lag Model with AR(1) Errors	583
	OLS Estimation of the ADL Model	585
	GLS Estimation	586
16.6	Orange Juice Prices and Cold Weather	588
16.7	Is Exogeneity Plausible? Some Examples	595
	U.S. Income and Australian Exports	595
	Oil Prices and Inflation	595
	Monetary Policy and Inflation	596
	The Growth Rate of GDP and the Term Spread	596
16.8	Conclusion	597
	APPENDIX 16.1 The Orange Juice Data Set	604
	APPENDIX 16.2 The ADL Model and Generalized Least Squares in Lag Operator Notation	605

**CHAPTER 17 Additional Topics in Time Series Regression 607**

- 17.1 Vector Autoregressions 607
  - The VAR Model 608
  - A VAR Model of the Growth Rate of GDP and the Term Spread 611
- 17.2 Multi-period Forecasts 612
  - Iterated Multi-period Forecasts 612
  - Direct Multi-period Forecasts 614
  - Which Method Should You Use? 616
- 17.3 Orders of Integration and the Nonnormality of Unit Root Test Statistics 616
  - Other Models of Trends and Orders of Integration 617
  - Why Do Unit Root Tests Have Nonnormal Distributions? 619
- 17.4 Cointegration 621
  - Cointegration and Error Correction 621
  - How Can You Tell Whether Two Variables Are Cointegrated? 622
  - Estimation of Cointegrating Coefficients 623
  - Extension to Multiple Cointegrated Variables 624
- 17.5 Volatility Clustering and Autoregressive Conditional Heteroskedasticity 625
  - Volatility Clustering 625
  - Realized Volatility 626
  - Autoregressive Conditional Heteroskedasticity 627
  - Application to Stock Price Volatility 628
- 17.6 Forecasting with Many Predictors Using Dynamic Factor Models and Principal Components 629
  - The Dynamic Factor Model 630
  - The DFM: Estimation and Forecasting 631
  - Application to U.S. Macroeconomic Data 634
- 17.7 Conclusion 640
  - APPENDIX 17.1** The Quarterly U.S. Macro Data Set 644

**PART FIVE Regression Analysis of Economic Time Series Data**

---

**CHAPTER 18 The Theory of Linear Regression with One Regressor 645**

- 18.1 The Extended Least Squares Assumptions and the OLS Estimator 646
  - The Extended Least Squares Assumptions 646
  - The OLS Estimator 647
- 18.2 Fundamentals of Asymptotic Distribution Theory 648
  - Convergence in Probability and the Law of Large Numbers 648
  - The Central Limit Theorem and Convergence in Distribution 650

	Slutsky's Theorem and the Continuous Mapping Theorem	651
	Application to the $t$ -Statistic Based on the Sample Mean	652
18.3	Asymptotic Distribution of the OLS Estimator and $t$ -Statistic	653
	Consistency and Asymptotic Normality of the OLS Estimators	653
	Consistency of Heteroskedasticity-Robust Standard Errors	653
	Asymptotic Normality of the Heteroskedasticity-Robust $t$ -Statistic	654
18.4	Exact Sampling Distributions When the Errors Are Normally Distributed	655
	Distribution of $\hat{\beta}_1$ with Normal Errors	655
	Distribution of the Homoskedasticity-Only $t$ -Statistic	656
18.5	Weighted Least Squares	657
	WLS with Known Heteroskedasticity	658
	WLS with Heteroskedasticity of Known Functional Form	659
	Heteroskedasticity-Robust Standard Errors or WLS?	661
	<b>APPENDIX 18.1</b> The Normal and Related Distributions and Moments of Continuous Random Variables	667
	<b>APPENDIX 18.2</b> Two Inequalities	669
<b>CHAPTER 19</b>	<b>The Theory of Multiple Regression</b>	<b>671</b>
19.1	The Linear Multiple Regression Model and OLS Estimator in Matrix Form	672
	The Multiple Regression Model in Matrix Notation	672
	The Extended Least Squares Assumptions	673
	The OLS Estimator	674
19.2	Asymptotic Distribution of the OLS Estimator and $t$ -Statistic	675
	The Multivariate Central Limit Theorem	676
	Asymptotic Normality of $\hat{\beta}$	676
	Heteroskedasticity-Robust Standard Errors	677
	Confidence Intervals for Predicted Effects	678
	Asymptotic Distribution of the $t$ -Statistic	678
19.3	Tests of Joint Hypotheses	679
	Joint Hypotheses in Matrix Notation	679
	Asymptotic Distribution of the $F$ -Statistic	679
	Confidence Sets for Multiple Coefficients	680
19.4	Distribution of Regression Statistics with Normal Errors	680
	Matrix Representations of OLS Regression Statistics	681
	Distribution of $\hat{\beta}$ with Independent Normal Errors	682
	Distribution of $s_{\hat{\beta}_1}^2$	682
	Homoskedasticity-Only Standard Errors	682
	Distribution of the $t$ -Statistic	683
	Distribution of the $F$ -Statistic	683

19.5	Efficiency of the OLS Estimator with Homoskedastic Errors	684
	The Gauss–Markov Conditions for Multiple Regression	684
	Linear Conditionally Unbiased Estimators	684
	The Gauss–Markov Theorem for Multiple Regression	685
19.6	Generalized Least Squares	686
	The GLS Assumptions	687
	GLS When $\Omega$ Is Known	688
	GLS When $\Omega$ Contains Unknown Parameters	689
	The Conditional Mean Zero Assumption and GLS	689
19.7	Instrumental Variables and Generalized Method of Moments Estimation	691
	The IV Estimator in Matrix Form	691
	Asymptotic Distribution of the TSLS Estimator	692
	Properties of TSLS When the Errors Are Homoskedastic	693
	Generalized Method of Moments Estimation in Linear Models	696
	APPENDIX 19.1 Summary of Matrix Algebra	706
	APPENDIX 19.2 Multivariate Distributions	710
	APPENDIX 19.3 Derivation of the Asymptotic Distribution of $\hat{\beta}$	711
	APPENDIX 19.4 Derivations of Exact Distributions of OLS Test Statistics with Normal Errors	712
	APPENDIX 19.5 Proof of the Gauss–Markov Theorem for Multiple Regression	713
	APPENDIX 19.6 Proof of Selected Results for IV and GMM Estimation	714
	APPENDIX 19.7 Regression with Many Predictors: MSPE, Ridge Regression, and Principal Components Analysis	716
	 <b>Appendix</b>	<b>721</b>
	<b>References</b>	<b>729</b>
	<b>Glossary</b>	<b>733</b>
	<b>Index</b>	<b>743</b>

# Key Concepts

## PART ONE Introduction and Review

---

- 1.1 Cross-Sectional, Time Series, and Panel Data 11
- 2.1 Expected Value and the Mean 18
- 2.2 Variance and Standard Deviation 19
- 2.3 Means, Variances, and Covariances of Sums of Random Variables 32
- 2.4 Computing Probabilities and Involving Normal Random Variables 34
- 2.5 Simple Random Sampling and i.i.d. Random Variables 40
- 2.6 Convergence in Probability, Consistency, and the Law of Large Numbers 44
- 2.7 The Central Limit Theorem 47
- 3.1 Estimators and Estimates 62
- 3.2 Bias, Consistency, and Efficiency 63
- 3.3 Efficiency of  $\bar{Y}$ :  $\bar{Y}$  Is BLUE 65
- 3.4 The Standard Error of  $\bar{Y}$  71
- 3.5 The Terminology of Hypothesis Testing 73
- 3.6 Testing the Hypothesis  $E(Y) = \mu_{Y,0}$  Against the Alternative  $E(Y) \neq \mu_{Y,0}$  74
- 3.7 Confidence Intervals for the Population Mean 76

## PART TWO Fundamentals of Regression Analysis

---

- 4.1 Terminology for the Linear Regression Model with a Single Regressor 104
- 4.2 The OLS Estimator, Predicted Values, and Residuals 108
- 4.3 The Least Squares Assumptions for Causal Inference 118
- 4.4 Large-Sample Distributions of  $\hat{\beta}_0$  and  $\hat{\beta}_1$  120
- 5.1 General Form of the  $t$ -Statistic 137
- 5.2 Testing the Hypothesis  $\beta_1 = \beta_{1,0}$  Against the Alternative  $\beta_1 \neq \beta_{1,0}$  139
- 5.3 Confidence Interval for  $\beta_1$  143
- 5.4 Heteroskedasticity and Homoskedasticity 148
- 5.5 The Gauss–Markov Theorem for  $\hat{\beta}_1$  153
- 6.1 Omitted Variable Bias in Regression with a Single Regressor 171
- 6.2 The Multiple Regression Model 177
- 6.3 The OLS Estimators, Predicted Values, and Residuals in the Multiple Regression Model 179
- 6.4 The Least Squares Assumptions for Causal Inference in the Multiple Regression Model 185
- 6.5 Large-Sample Distribution of  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$  186
- 6.6 The Least Squares Assumptions for Causal Inference in the Multiple Regression Model with Control Variables 191
- 7.1 Testing the Hypothesis  $\beta_j = \beta_{j,0}$  Against the Alternative  $\beta_j \neq \beta_{j,0}$  207

- 7.2 Confidence Intervals for a Single Coefficient in Multiple Regression 208
- 7.3  $R^2$  and  $\bar{R}^2$ : What They Tell You—and What They Don't 221
- 8.1 The Expected Change in  $Y$  from a Change in  $X_1$  in the Nonlinear Regression Model [Equation (8.3)] 241
- 8.2 Logarithms in Regression: Three Cases 253
- 8.3 A Method for Interpreting Coefficients in Regressions with Binary Variables 251
- 8.4 Interactions Between Binary and Continuous Variables 260
- 8.5 Interactions in Multiple Regression 264
- 9.1 Internal and External Validity 289
- 9.2 Omitted Variable Bias: Should I Include More Variables in My Regression? 293
- 9.3 Functional Form Misspecification 294
- 9.4 Errors-in-Variables Bias 296
- 9.5 Sample Selection Bias 298
- 9.6 Simultaneous Causality Bias 301
- 9.7 Threats to the Internal Validity of a Multiple Regression Study 302

### PART THREE Further Topics in Regression Analysis

---

- 10.1 Notation for Panel Data 320
- 10.2 The Fixed Effects Regression Model 327
- 10.3 The Fixed Effects Regression Assumptions 333
- 11.1 The Linear Probability Model 354
- 11.2 The Probit Model, Predicted Probabilities, and Estimated Effects 358
- 11.3 Logit Regression 360
- 12.1 The General Instrumental Variables Regression Model and Terminology 396
- 12.2 Two Stage Least Squares 398
- 12.3 The Two Conditions for Valid Instruments 399
- 12.4 The IV Regression Assumptions 400
- 12.5 A Rule of Thumb for Checking for Weak Instruments 404
- 12.6 The Overidentifying Restrictions Test (The  $J$ -Statistic) 407
- 14.1  $m$ -Fold Cross Validation 481
- 14.2 The Principal Components of  $X$  493

### PART FOUR Regression Analysis of Economic Time Series Data

---

- 15.1 Lags, First Differences, Logarithms, and Growth Rates 515
- 15.2 Autocorrelation (Serial Correlation) and Autocovariance 517
- 15.3 Stationarity 520
- 15.4 Autoregressions 526
- 15.5 The Autoregressive Distributed Lag Model 529
- 15.6 The Least Squares Assumptions for Forecasting with Time Series Data 530
- 15.7 Pseudo Out-of-Sample Forecasts 533
- 15.8 The QLR Test for Coefficient Stability 550
- 16.1 The Distributed Lag Model and Exogeneity 574

- 16.2 The Distributed Lag Model Assumptions 576
- 16.3 HAC Standard Errors 582
- 17.1 Vector Autoregressions 608
- 17.2 Iterated Multi-period Forecasts 614
- 17.3 Direct Multi-period Forecasts 616
- 17.4 Orders of Integration, Differencing, and Stationarity 618
- 17.5 Cointegration 622

## **PART FIVE**    **Regression Analysis of Economic Time Series Data**

---

- 18.1 The Extended Least Squares Assumptions for Regression with a Single Regressor 647
- 19.1 The Extended Least Squares Assumptions in the Multiple Regression Model 673
- 19.2 The Multivariate Central Limit Theorem 676
- 19.3 Gauss–Markov Theorem for Multiple Regression 685
- 19.4 The GLS Assumptions 687





# General Interest Boxes

The Distribution of Earnings in the United States in 2015	30
A Bad Day on Wall Street	35
Financial Diversification and Portfolios	42
Landon Wins!	66
The Gender Gap of Earnings of College Graduates in the United States	80
A Novel Way to Boost Retirement Savings	82
The “Beta” of a Stock	110
The Economic Value of a Year of Education: Homoskedasticity or Heteroskedasticity?	151
The Mozart Effect: Omitted Variable Bias?	172
The Return to Education and the Gender Gap	262
The Demand for Economics Journals	265
Do Stock Mutual Funds Outperform the Market?	299
James Heckman and Daniel McFadden, Nobel Laureates	372
Who Invented Instrumental Variables Regression?	388
The First IV Regression	405
The Externalities of Smoking	409
The Hawthorne Effect	438
Text as Data	501
Can You Beat the Market?	522
The River of Blood	535
Orange Trees on the March	593
NEWS FLASH: Commodity Traders Send Shivers Through Disney World	594
Nobel Laureates in Time Series Econometrics	638



# Preface

Econometrics can be a fun course for both teacher and student. The real world of economics, business, and government is a complicated and messy place, full of competing ideas and questions that demand answers. Is it more effective to tackle drunk driving by passing tough laws or by increasing the tax on alcohol? Can you make money in the stock market by buying when prices are historically low, relative to earnings, or should you just sit tight, as the random walk theory of stock prices suggests? Can we improve elementary education by reducing class sizes, or should we simply have our children listen to Mozart for 10 minutes a day? Econometrics helps us sort out sound ideas from crazy ones and find quantitative answers to important quantitative questions. Econometrics opens a window on our complicated world that lets us see the relationships on which people, businesses, and governments base their decisions.

*Introduction to Econometrics* is designed for a first course in undergraduate econometrics. It is our experience that to make econometrics relevant in an introductory course, interesting applications must motivate the theory and the theory must match the applications. This simple principle represents a significant departure from the older generation of econometrics books, in which theoretical models and assumptions do not match the applications. It is no wonder that some students question the relevance of econometrics after they spend much of their time learning assumptions that they subsequently realize are unrealistic so that they must then learn “solutions” to “problems” that arise when the applications do not match the assumptions. We believe that it is far better to motivate the need for tools with a concrete application and then to provide a few simple assumptions that match the application. Because the methods are immediately relevant to the applications, this approach can make econometrics come alive.

To improve student results, we recommend pairing the text content with MyLab Economics, which is the teaching and learning platform that empowers you to reach every student. By combining trusted author content with digital tools and a flexible platform, MyLab personalizes the learning experience and will help your students learn and retain key course concepts while developing skills that future employers are seeking in their candidates. MyLab Economics helps you teach your course, your way. Learn more at [www.pearson.com/mylab/economics](http://www.pearson.com/mylab/economics).

## New To This Edition

- New chapter on “Big Data” and machine learning
- Forecasting in time series data with large data sets

- Dynamic factor models
- Parallel treatment of prediction and causal inference using regression
- Now covers realized volatility as well as autoregressive conditional heteroskedasticity
- Updated discussion of weak instruments

Very large data sets are increasingly being used in economics and related fields. Applications include predicting consumer choices, measuring the quality of hospitals or schools, analyzing nonstandard data such as text data, and macroeconomic forecasting with many variables. The three main additions in this edition incorporate the fundamentals of this growing and exciting area of application.

First, we have a new chapter (Chapter 14) that focuses on big data and machine learning methods. Within economics, many of the applications to date have focused on the so called many-predictor problem, where the number of predictors is large relative to the sample size—perhaps even exceeding the sample size. With many predictors, ordinary least squares (OLS) provides poor predictions, and other methods, such as the LASSO, can have much lower out-of-sample prediction errors. This chapter goes over the concepts of out-of-sample prediction, why OLS performs poorly, and how shrinkage can improve upon OLS. The chapter introduces shrinkage methods and prediction using principal components, shows how to choose tuning parameters by cross-validation, and explains how these methods can be used to analyze nonstandard data such as text data. As usual, this chapter has a running empirical example, in this case, prediction of school-level test scores given school-level characteristics, for California elementary schools.

Second, in Chapter 17 (newly renumbered), we extend the many-predictor focus of Chapter 14 to time series data. Specifically, we show how the dynamic factor model can handle a very large number of time series, and show how to implement the dynamic factor model using principal components analysis. We illustrate the dynamic factor model and its use for forecasting with a 131-variable dataset of U.S. quarterly macroeconomic time series.

Third, we now lay out these two uses of regression—causal inference and prediction—up front, when regression is first introduced in Chapter 4. Regression is a statistical tool that can be used to make causal inferences or to make predictions; the two applications place different demands on how the data are collected. When the data are from a randomized controlled experiment, OLS estimates the causal effect. In observational data, if we are interested in estimating the causal effect, then the econometrician needs to use control variables and/or instruments to produce as-if randomization of the variable of interest. In contrast, for prediction, one is not interested in the causal effect so one does not need as-if random variation; however, the estimation (“training”) data set must be drawn from the same population as the observations for which one wishes to make the prediction.

This edition has several smaller changes. For example, we now introduce realized volatility as a complement to the GARCH model when analyzing time series data with volatility clustering. In addition, we now extend the discussion (in a new general interest box) of the historical origins of instrumental variables regression in Chapter 12. This treatment now includes a first-ever reproduction of the original derivation of the IV estimator, which was in a letter from Philip Wright to his son Sewall in the spring of 1926, and a discussion of the first IV regression, an estimate of the elasticity of supply of flaxseed.

## Solving Teaching and Learning Challenges

*Introduction to Econometrics* differs from other texts in three main ways. First, we integrate real-world questions and data into the development of the theory, and we take seriously the substantive findings of the resulting empirical analysis. Second, our choice of topics reflects modern theory and practice. Third, we provide theory and assumptions that match the applications. Our aim is to teach students to become sophisticated consumers of econometrics and to do so at a level of mathematics appropriate for an introductory course.

### Real-World Questions and Data

We organize each methodological topic around an important real-world question that demands a specific numerical answer. For example, we teach single-variable regression, multiple regression, and functional form analysis in the context of estimating the effect of school inputs on school outputs. (Do smaller elementary school class sizes produce higher test scores?) We teach panel data methods in the context of analyzing the effect of drunk driving laws on traffic fatalities. We use possible racial discrimination in the market for home loans as the empirical application for teaching regression with a binary dependent variable (logit and probit). We teach instrumental variable estimation in the context of estimating the demand elasticity for cigarettes. Although these examples involve economic reasoning, all can be understood with only a single introductory course in economics, and many can be understood without any previous economics coursework. Thus the instructor can focus on teaching econometrics, not microeconomics or macroeconomics.

We treat all our empirical applications seriously and in a way that shows students how they can learn from data but at the same time be self-critical and aware of the limitations of empirical analyses. Through each application, we teach students to explore alternative specifications and thereby to assess whether their substantive findings are robust. The questions asked in the empirical applications are important, and we provide serious and, we think, credible answers. We encourage students and instructors to disagree, however, and invite them to reanalyze the

data, which are provided on the text's Companion Website ([www.pearsonhighered.com/stock\\_watson](http://www.pearsonhighered.com/stock_watson)) and in MyLab Economics.

Throughout the text, we have focused on helping students understand, retain, and apply the essential ideas. *Chapter introductions* provide real-world grounding and motivation, as well as brief road maps highlighting the sequence of the discussion. *Key terms* are boldfaced and defined in context throughout each chapter, and *Key Concept boxes* at regular intervals recap the central ideas. *General interest boxes* provide interesting excursions into related topics and highlight real-world studies that use the methods or concepts being discussed in the text. A *Summary* concluding each chapter serves as a helpful framework for reviewing the main points of coverage.

Available for student practice or instructor assignment in MyLab Economics are *Review the Concepts questions*, *Exercises*, and *Empirical Exercises* from the text. These questions and exercises are auto-graded, giving students practical hands-on experience with solving problems using the data sets used in the text.

- 100 percent of Review the Concepts questions are available in MyLab.
- Select Exercises and Empirical Exercises are available in MyLab. Many of the Empirical Exercises are algorithmic and based on the data sets used in the text. These exercises require students to use Excel or an econometrics software package to analyze the data and derive results.
- New to the 4<sup>th</sup> edition are concept exercises that focus on core concepts and economic interpretations. Many are algorithmic and include the Help Me Solve This learning aid.

## Contemporary Choice of Topics

The topics we cover reflect the best of contemporary applied econometrics. One can only do so much in an introductory course, so we focus on procedures and tests that are commonly (or increasingly) used in practice. For example:

- ***Instrumental variables regression.*** We present instrumental variables regression as a general method for handling correlation between the error term and a regressor, which can arise for many reasons, including omitted variables and simultaneous causality. The two assumptions for a valid instrument—exogeneity and relevance—are given equal billing. We follow that presentation with an extended discussion of where instruments come from and with tests of overidentifying restrictions and diagnostics for weak instruments, and we explain what to do if these diagnostics suggest problems.
- ***Program evaluation.*** Many modern econometric studies analyze either randomized controlled experiments or quasi-experiments, also known as natural experiments. We address these topics, often collectively referred to as program

evaluation, in Chapter 13. We present this research strategy as an alternative approach to the problems of omitted variables, simultaneous causality, and selection, and we assess both the strengths and the weaknesses of studies using experimental or quasi-experimental data.

- **Prediction with “big data.”** Chapter 14 takes up the opportunities and challenges posed by large cross-sectional data sets. An increasingly common application in econometrics is making predictions when the number of predictors is very large. This chapter focuses on methods designed to use many predictors in a way that produces accurate and precise out-of-sample predictions. The chapter covers some of the building blocks of machine learning, and the methods can substantially improve upon OLS when the number of predictors is large. In addition, these methods extend to nonstandard data, such as text data.
- **Forecasting.** The chapter on forecasting (Chapter 15) considers univariate (autoregressive) and multivariate forecasts using time series regression, not large simultaneous equation structural models. We focus on simple and reliable tools, such as autoregressions and model selection via an information criterion, that work well in practice. This chapter also features a practically oriented treatment of structural breaks (at known and unknown dates) and pseudo out-of-sample forecasting, all in the context of developing stable and reliable time series forecasting models.
- **Time series regression.** The chapter on causal inference using time series data (Chapter 16) pays careful attention to when different estimation methods, including generalized least squares, will or will not lead to valid causal inferences and when it is advisable to estimate dynamic regressions using OLS with heteroskedasticity- and autocorrelation-consistent standard errors.

## Theory That Matches Applications

Although econometric tools are best motivated by empirical applications, students need to learn enough econometric theory to understand the strengths and limitations of those tools. We provide a modern treatment in which the fit between theory and applications is as tight as possible, while keeping the mathematics at a level that requires only algebra.

Modern empirical applications share some common characteristics: The data sets typically have many observations (hundreds or more); regressors are not fixed over repeated samples but rather are collected by random sampling (or some other mechanism that makes them random); the data are not normally distributed; and there is no *a priori* reason to think that the errors are homoskedastic (although often there are reasons to think that they are heteroskedastic).

These observations lead to important differences between the theoretical development in this text and other texts:

- **Large-sample approach.** Because data sets are large, from the outset we use large-sample normal approximations to sampling distributions for hypothesis testing and confidence intervals. In our experience, it takes less time to teach the rudiments of large-sample approximations than to teach the Student  $t$  and exact  $F$  distributions, degrees-of-freedom corrections, and so forth. This large-sample approach also saves students the frustration of discovering that, because of nonnormal errors, the exact distribution theory they just mastered is irrelevant. Once taught in the context of the sample mean, the large-sample approach to hypothesis testing and confidence intervals carries directly through multiple regression analysis, logit and probit, instrumental variables estimation, and time series methods.
- **Random sampling.** Because regressors are rarely fixed in econometric applications, from the outset we treat data on all variables (dependent and independent) as the result of random sampling. This assumption matches our initial applications to cross-sectional data, it extends readily to panel and time series data, and because of our large-sample approach, it poses no additional conceptual or mathematical difficulties.
- **Heteroskedasticity.** Applied econometricians routinely use heteroskedasticity-robust standard errors to eliminate worries about whether heteroskedasticity is present or not. In this book, we move beyond treating heteroskedasticity as an exception or a “problem” to be “solved”; instead, we allow for heteroskedasticity from the outset and simply use heteroskedasticity-robust standard errors. We present homoskedasticity as a special case that provides a theoretical motivation for OLS.

### Skilled Producers, Sophisticated Consumers

We hope that students using this book will become sophisticated consumers of empirical analysis. To do so, they must learn not only how to use the tools of regression analysis but also how to assess the validity of empirical analyses presented to them.

Our approach to teaching how to assess an empirical study is threefold. First, immediately after introducing the main tools of regression analysis, we devote Chapter 9 to the threats to internal and external validity of an empirical study. This chapter discusses data problems and issues of generalizing findings to other settings. It also examines the main threats to regression analysis, including omitted variables, functional form misspecification, errors-in-variables, selection, and simultaneity—and ways to recognize these threats in practice.



Second, we apply these methods for assessing empirical studies to the empirical analysis of the ongoing examples in the book. We do so by considering alternative specifications and by systematically addressing the various threats to validity of the analyses presented in the book.

Third, to become sophisticated consumers, students need firsthand experience as producers. Active learning beats passive learning, and econometrics is an ideal course for active learning. For this reason, the MyLab Economics and text website feature data sets, software, and suggestions for empirical exercises of different scopes.

### Approach to Mathematics and Level of Rigor

Our aim is for students to develop a sophisticated understanding of the tools of modern regression analysis, whether the course is taught at a “high” or a “low” level of mathematics. Parts I through IV of the text (which cover the substantive material) are written for students with only precalculus mathematics. Parts I through IV have fewer equations and more applications than many introductory econometrics books and far fewer equations than books aimed at mathematical sections of undergraduate courses. But more equations do not imply a more sophisticated treatment. In our experience, a more mathematical treatment does not lead to a deeper understanding for most students.

That said, different students learn differently, and for mathematically well-prepared students, learning can be enhanced by a more explicit mathematical treatment. The appendices in Parts I-IV therefore provide key calculations that are too involved to be included in the text. In addition, Part V contains an introduction to econometric theory that is appropriate for students with a stronger mathematical background. When the mathematical chapters in Part V are used in conjunction with the material in Parts I through IV (including appendices), this book is suitable for advanced undergraduate or master’s level econometrics courses.

### Developing Career Skills

For students to succeed in a rapidly changing job market, they should be aware of their career options and how to go about developing a variety of skills. Data analysis is an increasingly marketable skill. This text prepares the students for a range of data analytic applications, including causal inference and prediction. It also introduces the students to the core concepts of prediction using large data sets.

## Table of Contents Overview

There are five parts to *Introduction to Econometrics*. This text assumes that the student has had a course in probability and statistics, although we review that material in Part I. We cover the core material of regression analysis in Part II. Parts III, IV, and V present additional topics that build on the core treatment in Part II.

### Part I

Chapter 1 introduces econometrics and stresses the importance of providing quantitative answers to quantitative questions. It discusses the concept of causality in statistical studies and surveys the different types of data encountered in econometrics. Material from probability and statistics is reviewed in Chapters 2 and 3, respectively; whether these chapters are taught in a given course or are simply provided as a reference depends on the background of the students.

### Part II

Chapter 4 introduces regression with a single regressor and ordinary least squares (OLS) estimation, and Chapter 5 discusses hypothesis tests and confidence intervals in the regression model with a single regressor. In Chapter 6, students learn how they can address omitted variable bias using multiple regression, thereby estimating the effect of one independent variable while holding other independent variables constant. Chapter 7 covers hypothesis tests, including  $F$ -tests, and confidence intervals in multiple regression. In Chapter 8, the linear regression model is extended to models with nonlinear population regression functions, with a focus on regression functions that are linear in the parameters (so that the parameters can be estimated by OLS). In Chapter 9, students step back and learn how to identify the strengths and limitations of regression studies, seeing in the process how to apply the concepts of internal and external validity.

### Part III

Part III presents extensions of regression methods. In Chapter 10, students learn how to use panel data to control for unobserved variables that are constant over time. Chapter 11 covers regression with a binary dependent variable. Chapter 12 shows how instrumental variables regression can be used to address a variety of problems that produce correlation between the error term and the regressor, and examines how one might find and evaluate valid instruments. Chapter 13 introduces students to the analysis of data from experiments and quasi-, or natural, experiments, topics often referred to as “program evaluation.” Chapter 14 turns to econometric issues that arise with large data sets, and focuses on prediction when there are very many predictors.

## Part IV

Part IV takes up regression with time series data. Chapter 15 focuses on forecasting and introduces various modern tools for analyzing time series regressions, such as tests for stability. Chapter 16 discusses the use of time series data to estimate causal relations. Chapter 17 presents some more advanced tools for time series analysis, including models of volatility clustering and dynamic factor models.

## Part V

Part V is an introduction to econometric theory. This part is more than an appendix that fills in mathematical details omitted from the text. Rather, it is a self-contained treatment of the econometric theory of estimation and inference in the linear regression model. Chapter 18 develops the theory of regression analysis for a single regressor; the exposition does not use matrix algebra, although it does demand a higher level of mathematical sophistication than the rest of the text. Chapter 19 presents the multiple regression model, instrumental variables regression, generalized method of moments estimation of the linear model, and principal components analysis, all in matrix form.

## Prerequisites Within the Book

Because different instructors like to emphasize different material, we wrote this book with diverse teaching preferences in mind. To the maximum extent possible, the chapters in Parts III, IV, and V are “stand-alone” in the sense that they do not require first teaching all the preceding chapters. The specific prerequisites for each chapter are described in Table I. Although we have found that the sequence of topics adopted in the text works well in our own courses, the chapters are written in a way that allows instructors to present topics in a different order if they so desire.

## Sample Courses

This book accommodates several different course structures.

TABLE I Guide to Prerequisites for Special-Topic Chapters in Parts III, IV, and V									
Prerequisite parts or chapters									
Chapter	Part I 1–3	Part II 4–7, 9 8		Part III 10.1, 10.2 12.1, 12.2		Part IV 15.1–15.4 15.5–15.8 16			Part V 18
10	X <sup>a</sup>	X <sup>a</sup>	X						
11	X <sup>a</sup>	X <sup>a</sup>	X						
12.1, 12.2	X <sup>a</sup>	X <sup>a</sup>	X						
12.3–12.6	X <sup>a</sup>	X <sup>a</sup>	X	X	X				
13	X <sup>a</sup>	X <sup>a</sup>	X	X	X				
14	X <sup>a</sup>	X <sup>a</sup>	X						
15	X <sup>a</sup>	X <sup>a</sup>	b						
16	X <sup>a</sup>	X <sup>a</sup>	b			X			
17	X <sup>a</sup>	X <sup>a</sup>	b			X	X	X	
18	X	X	X						
19	X	X	X		X				X

This table shows the minimum prerequisites needed to cover the material in a given chapter. For example, estimation of dynamic causal effects with time series data (Chapter 16) first requires Part I (as needed, depending on student preparation, and except as noted in footnote a), Part II (except for Chapter 8; see footnote b), and Sections 15.1 through 15.4.

<sup>a</sup>Chapters 10 through 17 use exclusively large-sample approximations to sampling distributions, so the optional Sections 3.6 (the Student  $t$  distribution for testing means) and 5.6 (the Student  $t$  distribution for testing regression coefficients) can be skipped.

<sup>b</sup>Chapters 15 through 17 (the time series chapters) can be taught without first teaching Chapter 8 (nonlinear regression functions) if the instructor pauses to explain the use of logarithmic transformations to approximate percentage changes.

### Standard Introductory Econometrics

This course introduces econometrics (Chapter 1) and reviews probability and statistics as needed (Chapters 2 and 3). It then moves on to regression with a single regressor, multiple regression, the basics of functional form analysis, and the evaluation of regression studies (all Part II). The course proceeds to cover regression with panel data (Chapter 10), regression with a limited dependent variable (Chapter 11), and instrumental variables regression (Chapter 12), as time permits. The course then

turns to experiments and quasi-experiments in Chapter 13, topics that provide an opportunity to return to the questions of estimating causal effects raised at the beginning of the semester and to recapitulate core regression methods. If there is time, the students can be introduced to big data and machine learning methods at the end (Chapter 14). *Prerequisites: Algebra II and introductory statistics.*

### Introductory Econometrics with Time Series and Forecasting Applications

Like a standard introductory course, this course covers all of Part I (as needed) and Part II. Optionally, the course next provides a brief introduction to panel data (Sections 10.1 and 10.2) and takes up instrumental variables regression (Chapter 12, or just Sections 12.1 and 12.2). The course then proceeds to Chapter 14 (prediction in large cross sectional data sets). It then turns to Part IV, covering forecasting (Chapter 15) and estimation of dynamic causal effects (Chapter 16). If time permits, the course can include some advanced topics in time series analysis such as volatility clustering (Section 17.5) and forecasting with many predictors (Section 17.6). *Prerequisites: Algebra II and introductory statistics.*

### Applied Time Series Analysis and Forecasting

This book also can be used for a short course on applied time series and forecasting, for which a course on regression analysis is a prerequisite. Some time is spent reviewing the tools of basic regression analysis in Part II, depending on student preparation. The course then moves directly to time series forecasting (Chapter 15), estimation of dynamic causal effects (Chapter 16), and advanced topics in time series analysis (Chapter 17), including vector autoregressions. If there is time, the course can cover prediction using large data sets (Chapter 14 and Section 17.6). An important component of this course is hands-on forecasting exercises, available as the end-of-chapter Empirical Exercises for Chapters 15 and 17. *Prerequisites: Algebra II and basic introductory econometrics or the equivalent.*

### Introduction to Econometric Theory

This book is also suitable for an advanced undergraduate course in which the students have a strong mathematical preparation or for a master's level course in econometrics. The course briefly reviews the theory of statistics and probability as necessary (Part I). The course introduces regression analysis using the nonmathematical, applications-based treatment of Part II. This introduction is followed by the theoretical development in Chapters 18 and 19 (through Section 19.5). The course then takes up regression with a limited dependent variable (Chapter 11) and maximum likelihood estimation (Appendix 11.2). Next, the course optionally turns to instrumental variables regression and generalized method of moments (Chapter 12 and Section 19.7), time series methods (Chapter 15), the estimation of

causal effects using time series data and generalized least squares (Chapter 16 and Section 19.6), and/or to machine learning methods (Chapter 14 and Appendix 19.7). *Prerequisites: Calculus and introductory statistics. Chapter 18 assumes previous exposure to matrix algebra.*

## Instructor Teaching Resources

This program comes with the following teaching resources:

Supplements available to instructors at <a href="http://www.pearsonhighered.com">www.pearsonhighered.com</a>	Features of the Supplement
<b>Solutions Manual</b>	Solutions to the end-of-chapter content.
<b>Test Bank</b> Authored by Manfred Keil, Claremont McKenna College	1,000 multiple-choice questions, essays and longer questions, and mathematical and graphical problems with these annotations: <ul style="list-style-type: none"> <li>Type (Multiple-choice, essay, graphical)</li> </ul>
<b>Computerized TestGen</b>	TestGen allows instructors to: <ul style="list-style-type: none"> <li>Customize, save, and generate classroom tests</li> <li>Edit, add, or delete questions from the Test Item Files</li> <li>Analyze test results</li> <li>Organize a database of tests and student results.</li> </ul>
<b>PowerPoints</b>	Slides include all the graphs, tables, and equations in the text.  PowerPoints meet accessibility standards for students with disabilities. Features include, but not limited to: <ul style="list-style-type: none"> <li>Keyboard and Screen Reader access</li> <li>Alternative text for images</li> <li>High color contrast between background and foreground colors</li> </ul>
<b>Companion Website</b>	The Companion Website provides a wide range of additional resources for students and faculty. These resources include more and more in depth empirical exercises, data sets for the empirical exercises, replication files for empirical results reported in the text, and EViews tutorials.

## Acknowledgments

A great many people contributed to the first edition of this book. Our biggest debts of gratitude are to our colleagues at Harvard and Princeton who used early drafts of this book in their classrooms. At Harvard's Kennedy School of Government, Suzanne Cooper provided invaluable suggestions and detailed comments on multiple drafts. As a coteacher with one of the authors (Stock), she also helped vet much of the material in this book while it was being developed for a required course for master's students at the Kennedy School. We are also indebted to two other Kennedy School colleagues at the time, Alberto Abadie and Sue Dynarski, for their patient explanations of quasi-experiments and the field of program evaluation and for their detailed comments on early drafts of the text. At Princeton, Eli Tamer taught from an early draft and also provided helpful comments on the penultimate draft of the book.

We also owe much to many of our friends and colleagues in econometrics who spent time talking with us about the substance of this book and who collectively made so many helpful suggestions. Bruce Hansen (University of Wisconsin–Madison) and Bo Honore (Princeton) provided helpful feedback on very early outlines and preliminary versions of the core material in Part II. Joshua Angrist (MIT) and Guido Imbens (University of California, Berkeley) provided thoughtful suggestions about our treatment of materials on program evaluation. Our presentation of the material on time series has benefited from discussions with Yacine Ait-Sahalia (Princeton), Graham Elliott (University of California, San Diego), Andrew Harvey (Cambridge University), and Christopher Sims (Princeton). Finally, many people made helpful suggestions on parts of the manuscript close to their area of expertise: Don Andrews (Yale), John Bound (University of Michigan), Gregory Chow (Princeton), Thomas Downes (Tufts), David Drukker (StataCorp.), Jean Baldwin Grossman (Princeton), Eric Hanushek (Hoover Institution), James Heckman (University of Chicago), Han Hong (Princeton), Caroline Hoxby (Harvard), Alan Krueger (Princeton), Steven Levitt (University of Chicago), Richard Light (Harvard), David Neumark (Michigan State University), Joseph Newhouse (Harvard), Pierre Perron (Boston University), Kenneth Warner (University of Michigan), and Richard Zeckhauser (Harvard).

Many people were very generous in providing us with data. The California test score data were constructed with the assistance of Les Axelrod of the Standards and Assessments Division, California Department of Education. We are grateful to Charlie DePascale, Student Assessment Services, Massachusetts Department of Education, for his help with aspects of the Massachusetts test score data set. Christopher Ruhm (University of North Carolina, Greensboro) graciously provided us with his data set on drunk driving laws and traffic fatalities. The research department at the Federal Reserve Bank of Boston deserves thanks for putting together its data on racial discrimination in mortgage lending; we particularly thank Geoffrey Tootell for providing us with the updated version of the data set we use in Chapter 9 and Lynn Browne for explaining its policy context. We thank Jonathan Gruber (MIT) for sharing his data on cigarette sales, which we analyze in Chapter 12, and

Alan Krueger (Princeton) for his help with the Tennessee STAR data that we analyze in Chapter 13.

We thank several people for carefully checking the page proof for errors. Kerry Griffin and Yair Listokin read the entire manuscript, and Andrew Fraker, Ori Heffetz, Amber Henry, Hong Li, Alessandro Tarozzi, and Matt Watson worked through several chapters.

In the first edition, we benefited from the help of an exceptional development editor, Jane Tufts, whose creativity, hard work, and attention to detail improved the book in many ways, large and small. Pearson provided us with first-rate support, starting with our excellent editor, Sylvia Mallory, and extending through the entire publishing team. Jane and Sylvia patiently taught us a lot about writing, organization, and presentation, and their efforts are evident on every page of this book. We extend our thanks to the superb Pearson team, who worked with us on the second edition: Adrienne D'Ambrosio (senior acquisitions editor), Bridget Page (associate media producer), Charles Spaulding (senior designer), Nancy Fenton (managing editor) and her selection of Nancy Freihofer and Thompson Steele Inc. who handled the entire production process, Heather McNally (supplements coordinator), and Denise Clinton (editor-in-chief). Finally, we had the benefit of Kay Ueno's skilled editing in the second edition. We are also grateful to the excellent third edition Pearson team of Adrienne D'Ambrosio, Nancy Fenton, and Jill Kolongowski, as well as Rose Kernan, the project manager with Cengage Publisher Services. We also wish to thank the Pearson team who worked on the fourth edition: David Alexander, Christine Donovan, Nicole Nedwidek, and Rose Kernan, project manager with Cengage Publisher Services.

We also received a great deal of help and suggestions from faculty, students, and researchers as we prepared the third edition and its update. The changes made in the third edition incorporate or reflect suggestions, corrections, comments, data, and help provided by a number of researchers and instructors: Donald Andrews (Yale University), Jushan Bai (Columbia), James Cobbe (Florida State University), Susan Dynarski (University of Michigan), Nicole Eichelberger (Texas Tech University), Boyd Fjeldsted (University of Utah), Martina Grunow, Daniel Hamermesh (University of Texas–Austin), Keisuke Hirano (University of Arizona), Bo Honore (Princeton University), Guido Imbens (Harvard University), Manfred Keil (Claremont McKenna College), David Laibson (Harvard University), David Lee (Princeton University), Brigitte Madrian (Harvard University), Jorge Marquez (University of Maryland), Karen Bennett Mathis (Florida Department of Citrus), Alan Mehlenbacher (University of Victoria), Ulrich Müller (Princeton University), Serena Ng (Columbia University), Harry Patrinos (World Bank), Zhuan Pei (Brandeis University), Peter Summers (Texas Tech University), Andrey Vasnov (University of Sydney), and Douglas Young (Montana State University). We also benefited from student input from F. Hoces dela Guardia and Carrie Wilson.

Thoughtful reviews for the third edition were prepared for Pearson by Steve DeLoach (Elon University), Jeffrey DeSimone (University of Texas at Arlington),



Gary V. Engelhardt (Syracuse University), Luca Flabbi (Georgetown University), Steffen Habermalz (Northwestern University), Carolyn J. Heinrich (University of Wisconsin–Madison), Emma M. Iglesias-Vazquez (Michigan State University), Carlos Lamarche (University of Oklahoma), Vicki A. McCracken (Washington State University), Claudiney M. Pereira (Tulane University), and John T. Warner (Clemson University). We also received very helpful input on draft revisions of Chapters 7 and 10 from John Berdell (DePaul University), Janet Kohlhasse (University of Houston), Aprajit Mahajan (Stanford University), Xia Meng (Brandeis University), and Chan Shen (Georgetown University). We thank Christopher Stock for helping with the third edition cover.

In the fourth edition, we benefited from detailed comments on our prior treatment of causal analysis by Judea Pearl (UCLA) and Bryant Chen. Gary Chamberlain (Harvard), Guido Imbens (Stanford) and Jan Speiss (Stanford) provided thoughtful comments and guidance on Chapter 14. We received additional comments and/or corrections from Carlos C. Bautista (University of the Philippines), Brian Bethune (Tufts), Amitabh Chandra (Harvard Kennedy School), Julia Chang (University of New South Wales), Maia Güell (University of Edinburgh), Greg Mankiw (Harvard), Alan Mehlenbacher (University of Victoria), Franco Peracchi (Tor Vergata University), Peter Siminski (University of Wollongong), Jinhua Wang (University of Cambridge), and Michael Wolf (University of Zurich). We also benefited from a review panel that focused on the new Chapter 14, comprised of Chrystie Burr (University of Colorado-Boulder), Bentley Coffey (University of South Carolina), and Galin Todorov (Florida Atlantic University).

Above all, we are indebted to our families for their endurance throughout this project. Writing this book took a long time, and for them, the project must have seemed endless. They, more than anyone else, bore the burden of this commitment, and for their help and support we are deeply grateful.

