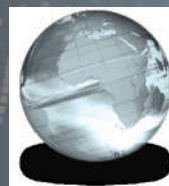


GLOBAL
EDITION



Introduction to Econometrics

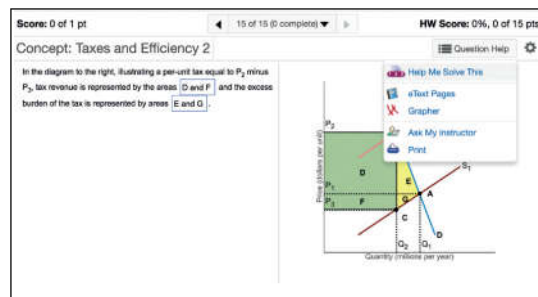
FOURTH EDITION

James H. Stock • Mark W. Watson



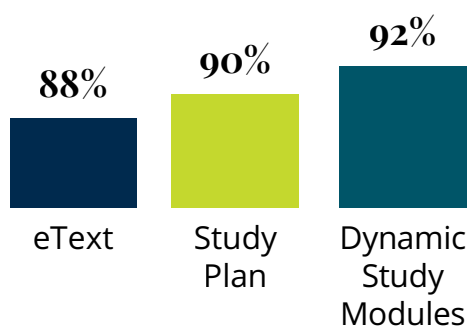
Question Help

MyLab Economics homework and practice questions are correlated to the textbook, and many generate algorithmically to give students unlimited opportunity for mastery of concepts. If students get stuck, Learning Aids including Help Me Solve This and eText Pages walk them through the problem and identify helpful information in the text, giving them assistance when they need it most.



"[MyLab Economics] provides ample practice and explanation of the concepts at hand."

— Heather Burkett, University of Nebraska at Omaha



% of students who found learning tool helpful

Dynamic Study Modules help students study chapter topics effectively on their own by continuously assessing their **knowledge application** and performance in real time. These are available as prebuilt Prepare assignments, and are accessible on smartphones, tablets, and computers.

Pearson eText enhances student learning—both in and outside the classroom. Worked examples, videos, and interactive tutorials bring learning to life, while algorithmic practice and self-assessment opportunities test students' understanding of the material. Accessible anytime, anywhere via MyLab or the app.

The **MyLab Gradebook** offers an easy way for students and instructors to view course performance. Item Analysis allows instructors to quickly see trends by analyzing details like the number of students who answered correctly/incorrectly, time on task, and median time spend on a question by question basis. And because it's correlated with the AACSB Standards, instructors can track students' progress toward outcomes that the organization has deemed important in preparing students to be **leaders**.

87% 

of students would tell their instructor to keep using MyLab Economics

For additional details visit: www.pearson.com/mylab/economics

The Pearson Series in Economics

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

*denotes **MyLab Economics** titles. Visit www.pearson.com/mylab/economics to learn more.

†denotes **Global Edition** titles.

Introduction to Econometrics

FOURTH EDITION

GLOBAL EDITION

James H. Stock

Harvard University

Mark W. Watson

Princeton University



Pearson

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Dubai • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • Sao Paulo • Mexico City • Madrid • Amsterdam • Munich • Paris • Milan

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
Project Editor, Global Edition: Paromita Banerjee
Project Editor, Global Edition: Punita Kaur Mann
Vice President, Product Marketing: Roxanne McCarley
Product Marketing Assistant: Marianela 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
Content Producer, Global Edition: Nikhil Rakshit

Operations Specialist: Carol Melville
Senior Manufacturing Controller, Global Edition: Kay Holman
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
Manager, Media Production, Global Edition: Vikram Kumar
Project Manager: Vikash Sharma, Cenveo Publisher Services
Interior Design: Cenveo Publisher Services
Cover Design: Lumina Datamatics
Cover Art: GarryKillian / Shutterstock

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

Pearson Education Limited

KAO Two
KAO Park
Harlow
CM17 9NA
United Kingdom

and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsonglobaleditions.com

© Pearson Education Limited 2020

The rights of James H. Stock and Mark W. Watson, to be identified as the authors of this work, have been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

Authorized adaptation from the United States edition, entitled *Introduction to Econometrics*, 4th Edition, ISBN 978-0-13-446199-1 by James H. Stock and Mark W. Watson, published by Pearson Education © 2020.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without either the prior written permission of the publisher or a license permitting restricted copying in the United Kingdom issued by the Copyright Licensing Agency Ltd, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

All trademarks used herein are the property of their respective owners. The use of any trademark in this text does not vest in the author or publisher any trademark ownership rights in such trademarks, nor does the use of such trademarks imply any affiliation with or endorsement of this book by such owners. 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/.

This eBook is a standalone product and may or may not include all assets that were part of the print version. It also does not provide access to other Pearson digital products like MyLab and Mastering. The publisher reserves the right to remove any material in this eBook at any time.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 10: 1-292-26445-4

ISBN 13: 978-1-292-26445-5

eBook ISBN 13: 978-1-292-26452-3

Typeset in Times NR MT Pro by Cenveo® Publisher Services

Brief Contents

PART ONE Introduction and Review

Chapter 1	Economic Questions and Data	43
Chapter 2	Review of Probability	55
Chapter 3	Review of Statistics	103

PART TWO Fundamentals of Regression Analysis

Chapter 4	Linear Regression with One Regressor	143
Chapter 5	Regression with a Single Regressor: Hypothesis Tests and Confidence Intervals	178
Chapter 6	Linear Regression with Multiple Regressors	211
Chapter 7	Hypothesis Tests and Confidence Intervals in Multiple Regression	247
Chapter 8	Nonlinear Regression Functions	277
Chapter 9	Assessing Studies Based on Multiple Regression	330

PART THREE Further Topics in Regression Analysis

Chapter 10	Regression with Panel Data	361
Chapter 11	Regression with a Binary Dependent Variable	392
Chapter 12	Instrumental Variables Regression	427
Chapter 13	Experiments and Quasi-Experiments	474
Chapter 14	Prediction with Many Regressors and Big Data	514

PART FOUR Regression Analysis of Economic Time Series Data

Chapter 15	Introduction to Time Series Regression and Forecasting	554
Chapter 16	Estimation of Dynamic Causal Effects	609
Chapter 17	Additional Topics in Time Series Regression	649

PART FIVE Regression Analysis of Economic Time Series Data

Chapter 18	The Theory of Linear Regression with One Regressor	687
Chapter 19	The Theory of Multiple Regression	713

This page intentionally left blank

Contents

Preface 27

PART ONE Introduction and Review

CHAPTER 1 Economic Questions and Data 43

- 1.1 Economic Questions We Examine 43
 - Question #1: Does Reducing Class Size Improve Elementary School Education? 43
 - Question #2: Is There Racial Discrimination in the Market for Home Loans? 44
 - Question #3: Does Healthcare Spending Improve Health Outcomes? 45
 - Question #4: By How Much Will U.S. GDP Grow Next Year? 46
 - Quantitative Questions, Quantitative Answers 47
- 1.2 Causal Effects and Idealized Experiments 47
 - Estimation of Causal Effects 48
 - Prediction, Forecasting, and Causality 48
- 1.3 Data: Sources and Types 49
 - Experimental versus Observational Data 49
 - Cross-Sectional Data 50
 - Time Series Data 51
 - Panel Data 52

CHAPTER 2 Review of Probability 55

- 2.1 Random Variables and Probability Distributions 56
 - Probabilities, the Sample Space, and Random Variables 56
 - Probability Distribution of a Discrete Random Variable 56
 - Probability Distribution of a Continuous Random Variable 58
- 2.2 Expected Values, Mean, and Variance 60
 - The Expected Value of a Random Variable 60
 - The Standard Deviation and Variance 61
 - Mean and Variance of a Linear Function of a Random Variable 62
 - Other Measures of the Shape of a Distribution 63
 - Standardized Random Variables 65
- 2.3 Two Random Variables 65
 - Joint and Marginal Distributions 65
 - Conditional Distributions 66
 - Independence 70
 - Covariance and Correlation 70
 - The Mean and Variance of Sums of Random Variables 71

- 2.4 The Normal, Chi-Squared, Student t , and F Distributions 75
 - The Normal Distribution 75
 - The Chi-Squared Distribution 80
 - The Student t Distribution 80
 - The F Distribution 80
- 2.5 Random Sampling and the Distribution of the Sample Average 81
 - Random Sampling 81
 - The Sampling Distribution of the Sample Average 82
- 2.6 Large-Sample Approximations to Sampling Distributions 85
 - The Law of Large Numbers and Consistency 85
 - The Central Limit Theorem 86
 - APPENDIX 2.1 Derivation of Results in Key Concept 2.3 100
 - APPENDIX 2.2 The Conditional Mean as the Minimum Mean Squared Error Predictor 101

CHAPTER 3 Review of Statistics 103

- 3.1 Estimation of the Population Mean 104
 - Estimators and Their Properties 104
 - Properties of \bar{Y} 106
 - The Importance of Random Sampling 108
- 3.2 Hypothesis Tests Concerning the Population Mean 109
 - Null and Alternative Hypotheses 109
 - The p -Value 110
 - Calculating the p -Value When σ_Y Is Known 111
 - The Sample Variance, Sample Standard Deviation, and Standard Error 112
 - Calculating the p -Value When σ_Y Is Unknown 113
 - The t -Statistic 113
 - Hypothesis Testing with a Prespecified Significance Level 114
 - One-Sided Alternatives 116
- 3.3 Confidence Intervals for the Population Mean 117
- 3.4 Comparing Means from Different Populations 119
 - Hypothesis Tests for the Difference Between Two Means 119
 - Confidence Intervals for the Difference Between Two Population Means 120
- 3.5 Differences-of-Means Estimation of Causal Effects Using Experimental Data 121
 - The Causal Effect as a Difference of Conditional Expectations 121
 - Estimation of the Causal Effect Using Differences of Means 121
- 3.6 Using the t -Statistic When the Sample Size Is Small 123
 - The t -Statistic and the Student t Distribution 125
 - Use of the Student t Distribution in Practice 126

- 3.7 Scatterplots, the Sample Covariance, and the Sample Correlation 127
 - Scatterplots 127
 - Sample Covariance and Correlation 127
 - APPENDIX 3.1 The U.S. Current Population Survey 141
 - APPENDIX 3.2 Two Proofs That \bar{Y} Is the Least Squares Estimator of μ_Y 141
 - APPENDIX 3.3 A Proof That the Sample Variance Is Consistent 142

PART TWO Fundamentals of Regression Analysis

CHAPTER 4 Linear Regression with One Regressor 143

- 4.1 The Linear Regression Model 144
- 4.2 Estimating the Coefficients of the Linear Regression Model 147
 - The Ordinary Least Squares Estimator 148
 - OLS Estimates of the Relationship Between Test Scores and the Student-Teacher Ratio 149
 - Why Use the OLS Estimator? 151
- 4.3 Measures of Fit and Prediction Accuracy 153
 - The R^2 153
 - The Standard Error of the Regression 154
 - Prediction Using OLS 155
 - Application to the Test Score Data 155
- 4.4 The Least Squares Assumptions for Causal Inference 156
 - Assumption 1: The Conditional Distribution of u_i Given X_i Has a Mean of Zero 157
 - Assumption 2: $(X_i, Y_i), i = 1, \dots, n$, Are Independently and Identically Distributed 158
 - Assumption 3: Large Outliers Are Unlikely 159
 - Use of the Least Squares Assumptions 160
- 4.5 The Sampling Distribution of the OLS Estimators 161
- 4.6 Conclusion 164
 - APPENDIX 4.1 The California Test Score Data Set 172
 - APPENDIX 4.2 Derivation of the OLS Estimators 172
 - APPENDIX 4.3 Sampling Distribution of the OLS Estimator 173
 - APPENDIX 4.4 The Least Squares Assumptions for Prediction 176

CHAPTER 5 Regression with a Single Regressor: Hypothesis Tests and Confidence Intervals 178

- 5.1 Testing Hypotheses About One of the Regression Coefficients 178
 - Two-Sided Hypotheses Concerning β_1 179
 - One-Sided Hypotheses Concerning β_1 182
 - Testing Hypotheses About the Intercept β_0 184
- 5.2 Confidence Intervals for a Regression Coefficient 184

- 5.3 Regression When X Is a Binary Variable 186
 - Interpretation of the Regression Coefficients 186
- 5.4 Heteroskedasticity and Homoskedasticity 188
 - What Are Heteroskedasticity and Homoskedasticity? 188
 - Mathematical Implications of Homoskedasticity 190
 - What Does This Mean in Practice? 192
- 5.5 The Theoretical Foundations of Ordinary Least Squares 194
 - Linear Conditionally Unbiased Estimators and the Gauss–Markov Theorem 194
 - Regression Estimators Other Than OLS 195
- 5.6 Using the t -Statistic in Regression When the Sample Size Is Small 196
 - The t -Statistic and the Student t Distribution 196
 - Use of the Student t Distribution in Practice 197
- 5.7 Conclusion 197
 - APPENDIX 5.1 Formulas for OLS Standard Errors 206
 - APPENDIX 5.2 The Gauss–Markov Conditions and a Proof of the Gauss–Markov Theorem 207

CHAPTER 6 Linear Regression with Multiple Regressors 211

- 6.1 Omitted Variable Bias 211
 - Definition of Omitted Variable Bias 212
 - A Formula for Omitted Variable Bias 214
 - Addressing Omitted Variable Bias by Dividing the Data into Groups 215
- 6.2 The Multiple Regression Model 217
 - The Population Regression Line 217
 - The Population Multiple Regression Model 218
- 6.3 The OLS Estimator in Multiple Regression 220
 - The OLS Estimator 220
 - Application to Test Scores and the Student–Teacher Ratio 221
- 6.4 Measures of Fit in Multiple Regression 222
 - The Standard Error of the Regression (SER) 222
 - The R^2 223
 - The Adjusted R^2 223
 - Application to Test Scores 224
- 6.5 The Least Squares Assumptions for Causal Inference in Multiple Regression 225
 - Assumption 1: The Conditional Distribution of u_i Given $X_{1i}, X_{2i}, \dots, X_{ki}$ Has a Mean of 0 225
 - Assumption 2: $(X_{1i}, X_{2i}, \dots, X_{ki}, Y_i), i = 1, \dots, n$, Are i.i.d. 225
 - Assumption 3: Large Outliers Are Unlikely 225
 - Assumption 4: No Perfect Multicollinearity 226

- 6.6 The Distribution of the OLS Estimators in Multiple Regression 227
- 6.7 Multicollinearity 228
 - Examples of Perfect Multicollinearity 228
 - Imperfect Multicollinearity 230
- 6.8 Control Variables and Conditional Mean Independence 231
 - Control Variables and Conditional Mean Independence 232
- 6.9 Conclusion 234
 - APPENDIX 6.1 Derivation of Equation (6.1) 242
 - APPENDIX 6.2 Distribution of the OLS Estimators When There Are Two Regressors and Homoskedastic Errors 243
 - APPENDIX 6.3 The Frisch–Waugh Theorem 243
 - APPENDIX 6.4 The Least Squares Assumptions for Prediction with Multiple Regressors 244
 - APPENDIX 6.5 Distribution of OLS Estimators in Multiple Regression with Control Variables 245

CHAPTER 7 Hypothesis Tests and Confidence Intervals in Multiple Regression 247

- 7.1 Hypothesis Tests and Confidence Intervals for a Single Coefficient 247
 - Standard Errors for the OLS Estimators 247
 - Hypothesis Tests for a Single Coefficient 248
 - Confidence Intervals for a Single Coefficient 249
 - Application to Test Scores and the Student–Teacher Ratio 249
- 7.2 Tests of Joint Hypotheses 251
 - Testing Hypotheses on Two or More Coefficients 252
 - The F -Statistic 253
 - Application to Test Scores and the Student–Teacher Ratio 255
 - The Homoskedasticity-Only F -Statistic 256
- 7.3 Testing Single Restrictions Involving Multiple Coefficients 258
- 7.4 Confidence Sets for Multiple Coefficients 259
- 7.5 Model Specification for Multiple Regression 260
 - Model Specification and Choosing Control Variables 261
 - Interpreting the R^2 and the Adjusted R^2 in Practice 262
- 7.6 Analysis of the Test Score Data Set 262
- 7.7 Conclusion 268
 - APPENDIX 7.1 The Bonferroni Test of a Joint Hypothesis 274

CHAPTER 8	Nonlinear Regression Functions	277
8.1	A General Strategy for Modeling Nonlinear Regression Functions	279
	Test Scores and District Income	279
	The Effect on Y of a Change in X in Nonlinear Specifications	282
	A General Approach to Modeling Nonlinearities Using Multiple Regression	285
8.2	Nonlinear Functions of a Single Independent Variable	286
	Polynomials	286
	Logarithms	288
	Polynomial and Logarithmic Models of Test Scores and District Income	296
8.3	Interactions Between Independent Variables	297
	Interactions Between Two Binary Variables	298
	Interactions Between a Continuous and a Binary Variable	300
	Interactions Between Two Continuous Variables	305
8.4	Nonlinear Effects on Test Scores of the Student–Teacher Ratio	310
	Discussion of Regression Results	310
	Summary of Findings	314
8.5	Conclusion	315
	APPENDIX 8.1 Regression Functions That Are Nonlinear in the Parameters	325
	APPENDIX 8.2 Slopes and Elasticities for Nonlinear Regression Functions	328
CHAPTER 9	Assessing Studies Based on Multiple Regression	330
9.1	Internal and External Validity	330
	Threats to Internal Validity	331
	Threats to External Validity	332
9.2	Threats to Internal Validity of Multiple Regression Analysis	333
	Omitted Variable Bias	334
	Misspecification of the Functional Form of the Regression Function	336
	Measurement Error and Errors-in-Variables Bias	336
	Missing Data and Sample Selection	339
	Simultaneous Causality	341
	Sources of Inconsistency of OLS Standard Errors	343
9.3	Internal and External Validity When the Regression Is Used for Prediction	344
9.4	Example: Test Scores and Class Size	345
	External Validity	346
	Internal Validity	352
	Discussion and Implications	353
9.5	Conclusion	354
	APPENDIX 9.1 The Massachusetts Elementary School Testing Data	360

PART THREE Further Topics in Regression Analysis

CHAPTER 10 Regression with Panel Data 361

- 10.1 Panel Data 362
 - Example: Traffic Deaths and Alcohol Taxes 362
- 10.2 Panel Data with Two Time Periods: “Before and After” Comparisons 365
- 10.3 Fixed Effects Regression 367
 - The Fixed Effects Regression Model 367
 - Estimation and Inference 369
 - Application to Traffic Deaths 370
- 10.4 Regression with Time Fixed Effects 371
 - Time Effects Only 371
 - Both Entity and Time Fixed Effects 372
- 10.5 The Fixed Effects Regression Assumptions and Standard Errors for Fixed Effects Regression 374
 - The Fixed Effects Regression Assumptions 374
 - Standard Errors for Fixed Effects Regression 376
- 10.6 Drunk Driving Laws and Traffic Deaths 377
- 10.7 Conclusion 381
 - [APPENDIX 10.1 The State Traffic Fatality Data Set](#) 387
 - [APPENDIX 10.2 Standard Errors for Fixed Effects Regression](#) 388

CHAPTER 11 Regression with a Binary Dependent Variable 392

- 11.1 Binary Dependent Variables and the Linear Probability Model 393
 - Binary Dependent Variables 393
 - The Linear Probability Model 395
- 11.2 Probit and Logit Regression 397
 - Probit Regression 397
 - Logit Regression 401
 - Comparing the Linear Probability, Probit, and Logit Models 403
- 11.3 Estimation and Inference in the Logit and Probit Models 404
 - Nonlinear Least Squares Estimation 404
 - Maximum Likelihood Estimation 405
 - Measures of Fit 406
- 11.4 Application to the Boston HMDA Data 407
- 11.5 Conclusion 413
 - [APPENDIX 11.1 The Boston HMDA Data Set](#) 421
 - [APPENDIX 11.2 Maximum Likelihood Estimation](#) 421
 - [APPENDIX 11.3 Other Limited Dependent Variable Models](#) 424

CHAPTER 12	Instrumental Variables Regression	427
12.1	The IV Estimator with a Single Regressor and a Single Instrument	428
	The IV Model and Assumptions	428
	The Two Stage Least Squares Estimator	429
	Why Does IV Regression Work?	429
	The Sampling Distribution of the TOLS Estimator	434
	Application to the Demand for Cigarettes	435
12.2	The General IV Regression Model	437
	TOLS in the General IV Model	439
	Instrument Relevance and Exogeneity in the General IV Model	440
	The IV Regression Assumptions and Sampling Distribution of the TOLS Estimator	441
	Inference Using the TOLS Estimator	442
	Application to the Demand for Cigarettes	443
12.3	Checking Instrument Validity	444
	Assumption 1: Instrument Relevance	444
	Assumption 2: Instrument Exogeneity	446
12.4	Application to the Demand for Cigarettes	450
12.5	Where Do Valid Instruments Come From?	454
	Three Examples	455
12.6	Conclusion	459
	APPENDIX 12.1 The Cigarette Consumption Panel Data Set	467
	APPENDIX 12.2 Derivation of the Formula for the TOLS Estimator in Equation (12.4)	467
	APPENDIX 12.3 Large-Sample Distribution of the TOLS Estimator	468
	APPENDIX 12.4 Large-Sample Distribution of the TOLS Estimator When the Instrument Is Not Valid	469
	APPENDIX 12.5 Instrumental Variables Analysis with Weak Instruments	470
	APPENDIX 12.6 TOLS with Control Variables	472
CHAPTER 13	Experiments and Quasi-Experiments	474
13.1	Potential Outcomes, Causal Effects, and Idealized Experiments	475
	Potential Outcomes and the Average Causal Effect	475
	Econometric Methods for Analyzing Experimental Data	476
13.2	Threats to Validity of Experiments	478
	Threats to Internal Validity	478
	Threats to External Validity	481
13.3	Experimental Estimates of the Effect of Class Size Reductions	482
	Experimental Design	482
	Analysis of the STAR Data	483
	Comparison of the Observational and Experimental Estimates of Class Size Effects	488

13.4	Quasi-Experiments	490
	Examples	490
	The Differences-in-Differences Estimator	492
	Instrumental Variables Estimators	494
	Regression Discontinuity Estimators	495
13.5	Potential Problems with Quasi-Experiments	496
	Threats to Internal Validity	496
	Threats to External Validity	498
13.6	Experimental and Quasi-Experimental Estimates in Heterogeneous Populations	498
	OLS with Heterogeneous Causal Effects	499
	IV Regression with Heterogeneous Causal Effects	500
13.7	Conclusion	503
	APPENDIX 13.1 The Project STAR Data Set	510
	APPENDIX 13.2 IV Estimation When the Causal Effect Varies Across Individuals	511
	APPENDIX 13.3 The Potential Outcomes Framework for Analyzing Data from Experiments	512
CHAPTER 14	Prediction with Many Regressors and Big Data	514
14.1	What Is “Big Data”?	515
14.2	The Many-Predictor Problem and OLS	516
	The Mean Squared Prediction Error	518
	The First Least Squares Assumption for Prediction	519
	The Predictive Regression Model with Standardized Regressors	519
	The MSPE of OLS and the Principle of Shrinkage	521
	Estimation of the MSPE	522
14.3	Ridge Regression	524
	Shrinkage via Penalization and Ridge Regression	524
	Estimation of the Ridge Shrinkage Parameter by Cross Validation	525
	Application to School Test Scores	526
14.4	The Lasso	527
	Shrinkage Using the Lasso	528
	Application to School Test Scores	531
14.5	Principal Components	532
	Principals Components with Two Variables	532
	Principal Components with k Variables	534
	Application to School Test Scores	536
14.6	Predicting School Test Scores with Many Predictors	537

- 14.7 Conclusion 542
 - APPENDIX 14.1 The California School Test Score Data Set 551
 - APPENDIX 14.2 Derivation of Equation (14.4) for $k = 1$ 551
 - APPENDIX 14.3 The Ridge Regression Estimator When $k = 1$ 551
 - APPENDIX 14.4 The Lasso Estimator When $k = 1$ 552
 - APPENDIX 14.5 Computing Out-of-Sample Predictions in the Standardized Regression Model 552

PART FOUR Regression Analysis of Economic Time Series Data

CHAPTER 15 Introduction to Time Series Regression and Forecasting 554

- 15.1 Introduction to Time Series Data and Serial Correlation 555
 - Real GDP in the United States 555
 - Lags, First Differences, Logarithms, and Growth Rates 555
 - Autocorrelation 558
 - Other Examples of Economic Time Series 560
- 15.2 Stationarity and the Mean Squared Forecast Error 561
 - Stationarity 561
 - Forecasts and Forecast Errors 562
 - The Mean Squared Forecast Error 563
- 15.3 Autoregressions 565
 - The First-Order Autoregressive Model 565
 - The p^{th} -Order Autoregressive Model 567
- 15.4 Time Series Regression with Additional Predictors and the Autoregressive Distributed Lag Model 568
 - Forecasting GDP Growth Using the Term Spread 569
 - The Autoregressive Distributed Lag Model 570
 - The Least Squares Assumptions for Forecasting with Multiple Predictors 571
- 15.5 Estimation of the MSFE and Forecast Intervals 573
 - Estimation of the MSFE 573
 - Forecast Uncertainty and Forecast Intervals 576
- 15.6 Estimating the Lag Length Using Information Criteria 578
 - Determining the Order of an Autoregression 578
 - Lag Length Selection in Time Series Regression with Multiple Predictors 581
- 15.7 Nonstationarity I: Trends 582
 - What Is a Trend? 582
 - Problems Caused by Stochastic Trends 584
 - Detecting Stochastic Trends: Testing for a Unit AR Root 586
 - Avoiding the Problems Caused by Stochastic Trends 588

15.8	Nonstationarity II: Breaks	589
	What Is a Break?	589
	Testing for Breaks	589
	Detecting Breaks Using Pseudo Out-of-Sample Forecasts	594
	Avoiding the Problems Caused by Breaks	595
15.9	Conclusion	596
	APPENDIX 15.1 Time Series Data Used in Chapter 15	604
	APPENDIX 15.2 Stationarity in the AR(1) Model	605
	APPENDIX 15.3 Lag Operator Notation	606
	APPENDIX 15.4 ARMA Models	607
	APPENDIX 15.5 Consistency of the BIC Lag Length Estimator	607
CHAPTER 16	Estimation of Dynamic Causal Effects	609
16.1	An Initial Taste of the Orange Juice Data	610
16.2	Dynamic Causal Effects	612
	Causal Effects and Time Series Data	612
	Two Types of Exogeneity	615
16.3	Estimation of Dynamic Causal Effects with Exogenous Regressors	617
	The Distributed Lag Model Assumptions	617
	Autocorrelated u_t , Standard Errors, and Inference	618
	Dynamic Multipliers and Cumulative Dynamic Multipliers	618
16.4	Heteroskedasticity- and Autocorrelation-Consistent Standard Errors	620
	Distribution of the OLS Estimator with Autocorrelated Errors	620
	HAC Standard Errors	621
16.5	Estimation of Dynamic Causal Effects with Strictly Exogenous Regressors	624
	The Distributed Lag Model with AR(1) Errors	625
	OLS Estimation of the ADL Model	627
	GLS Estimation	628
16.6	Orange Juice Prices and Cold Weather	630
16.7	Is Exogeneity Plausible? Some Examples	637
	U.S. Income and Australian Exports	637
	Oil Prices and Inflation	637
	Monetary Policy and Inflation	638
	The Growth Rate of GDP and the Term Spread	638
16.8	Conclusion	639
	APPENDIX 16.1 The Orange Juice Data Set	646
	APPENDIX 16.2 The ADL Model and Generalized Least Squares in Lag Operator Notation	647

CHAPTER 17 Additional Topics in Time Series Regression 649

- 17.1 Vector Autoregressions 649
 - The VAR Model 650
 - A VAR Model of the Growth Rate of GDP and the Term Spread 653
- 17.2 Multi-period Forecasts 654
 - Iterated Multi-period Forecasts 654
 - Direct Multi-period Forecasts 656
 - Which Method Should You Use? 658
- 17.3 Orders of Integration and the Nonnormality of Unit Root Test Statistics 658
 - Other Models of Trends and Orders of Integration 659
 - Why Do Unit Root Tests Have Nonnormal Distributions? 661
- 17.4 Cointegration 663
 - Cointegration and Error Correction 663
 - How Can You Tell Whether Two Variables Are Cointegrated? 664
 - Estimation of Cointegrating Coefficients 665
 - Extension to Multiple Cointegrated Variables 666
- 17.5 Volatility Clustering and Autoregressive Conditional Heteroskedasticity 667
 - Volatility Clustering 667
 - Realized Volatility 668
 - Autoregressive Conditional Heteroskedasticity 669
 - Application to Stock Price Volatility 670
- 17.6 Forecasting with Many Predictors Using Dynamic Factor Models and Principal Components 671
 - The Dynamic Factor Model 672
 - The DFM: Estimation and Forecasting 673
 - Application to U.S. Macroeconomic Data 676
- 17.7 Conclusion 682
 - [APPENDIX 17.1](#) The Quarterly U.S. Macro Data Set 686

PART FIVE Regression Analysis of Economic Time Series Data

CHAPTER 18 The Theory of Linear Regression with One Regressor 687

- 18.1 The Extended Least Squares Assumptions and the OLS Estimator 688
 - The Extended Least Squares Assumptions 688
 - The OLS Estimator 689
- 18.2 Fundamentals of Asymptotic Distribution Theory 690
 - Convergence in Probability and the Law of Large Numbers 690
 - The Central Limit Theorem and Convergence in Distribution 692

Slutsky's Theorem and the Continuous Mapping Theorem 693
 Application to the t -Statistic Based on the Sample Mean 694

18.3 Asymptotic Distribution of the OLS Estimator and t -Statistic 695
 Consistency and Asymptotic Normality of the OLS Estimators 695
 Consistency of Heteroskedasticity-Robust Standard Errors 695
 Asymptotic Normality of the Heteroskedasticity-Robust t -Statistic 696

18.4 Exact Sampling Distributions When the Errors Are Normally Distributed 697
 Distribution of $\hat{\beta}_1$ with Normal Errors 697
 Distribution of the Homoskedasticity-Only t -Statistic 698

18.5 Weighted Least Squares 699
 WLS with Known Heteroskedasticity 700
 WLS with Heteroskedasticity of Known Functional Form 701
 Heteroskedasticity-Robust Standard Errors or WLS? 703

APPENDIX 18.1 The Normal and Related Distributions and Moments of Continuous Random Variables 709

APPENDIX 18.2 Two Inequalities 711

CHAPTER 19 The Theory of Multiple Regression 713

19.1 The Linear Multiple Regression Model and OLS Estimator in Matrix Form 714
 The Multiple Regression Model in Matrix Notation 714
 The Extended Least Squares Assumptions 715
 The OLS Estimator 716

19.2 Asymptotic Distribution of the OLS Estimator and t -Statistic 717
 The Multivariate Central Limit Theorem 718
 Asymptotic Normality of $\hat{\beta}$ 718
 Heteroskedasticity-Robust Standard Errors 719
 Confidence Intervals for Predicted Effects 720
 Asymptotic Distribution of the t -Statistic 720

19.3 Tests of Joint Hypotheses 721
 Joint Hypotheses in Matrix Notation 721
 Asymptotic Distribution of the F -Statistic 721
 Confidence Sets for Multiple Coefficients 722

19.4 Distribution of Regression Statistics with Normal Errors 722
 Matrix Representations of OLS Regression Statistics 723
 Distribution of $\hat{\beta}$ with Independent Normal Errors 724
 Distribution of $s_{\hat{\beta}_j}^2$ 724
 Homoskedasticity-Only Standard Errors 724
 Distribution of the t -Statistic 725
 Distribution of the F -Statistic 725

19.5	Efficiency of the OLS Estimator with Homoskedastic Errors	726
	The Gauss–Markov Conditions for Multiple Regression	726
	Linear Conditionally Unbiased Estimators	726
	The Gauss–Markov Theorem for Multiple Regression	727
19.6	Generalized Least Squares	728
	The GLS Assumptions	729
	GLS When Ω Is Known	730
	GLS When Ω Contains Unknown Parameters	731
	The Conditional Mean Zero Assumption and GLS	731
19.7	Instrumental Variables and Generalized Method of Moments Estimation	733
	The IV Estimator in Matrix Form	733
	Asymptotic Distribution of the TSLS Estimator	734
	Properties of TSLS When the Errors Are Homoskedastic	735
	Generalized Method of Moments Estimation in Linear Models	738
	APPENDIX 19.1 Summary of Matrix Algebra	748
	APPENDIX 19.2 Multivariate Distributions	752
	APPENDIX 19.3 Derivation of the Asymptotic Distribution of $\hat{\beta}$	753
	APPENDIX 19.4 Derivations of Exact Distributions of OLS Test Statistics with Normal Errors	754
	APPENDIX 19.5 Proof of the Gauss–Markov Theorem for Multiple Regression	755
	APPENDIX 19.6 Proof of Selected Results for IV and GMM Estimation	756
	APPENDIX 19.7 Regression with Many Predictors: MSPE, Ridge Regression, and Principal Components Analysis	758
	Appendix	763
	References	771
	Glossary	775
	Index	785

Key Concepts

PART ONE Introduction and Review

- 1.1 Cross-Sectional, Time Series, and Panel Data 53
- 2.1 Expected Value and the Mean 60
- 2.2 Variance and Standard Deviation 61
- 2.3 Means, Variances, and Covariances of Sums of Random Variables 74
- 2.4 Computing Probabilities and Involving Normal Random Variables 76
- 2.5 Simple Random Sampling and i.i.d. Random Variables 82
- 2.6 Convergence in Probability, Consistency, and the Law of Large Numbers 86
- 2.7 The Central Limit Theorem 89
- 3.1 Estimators and Estimates 105
- 3.2 Bias, Consistency, and Efficiency 105
- 3.3 Efficiency of \bar{Y} : \bar{Y} Is BLUE 107
- 3.4 The Standard Error of \bar{Y} 113
- 3.5 The Terminology of Hypothesis Testing 115
- 3.6 Testing the Hypothesis $E(Y) = \mu_{Y,0}$ Against the Alternative $E(Y) \neq \mu_{Y,0}$ 116
- 3.7 Confidence Intervals for the Population Mean 118

PART TWO Fundamentals of Regression Analysis

- 4.1 Terminology for the Linear Regression Model with a Single Regressor 146
- 4.2 The OLS Estimator, Predicted Values, and Residuals 150
- 4.3 The Least Squares Assumptions for Causal Inference 160
- 4.4 Large-Sample Distributions of $\hat{\beta}_0$ and $\hat{\beta}_1$ 162
- 5.1 General Form of the t -Statistic 179
- 5.2 Testing the Hypothesis $\beta_1 = \beta_{1,0}$ Against the Alternative $\beta_1 \neq \beta_{1,0}$ 181
- 5.3 Confidence Interval for β_1 185
- 5.4 Heteroskedasticity and Homoskedasticity 190
- 5.5 The Gauss–Markov Theorem for $\hat{\beta}_1$ 195
- 6.1 Omitted Variable Bias in Regression with a Single Regressor 213
- 6.2 The Multiple Regression Model 219
- 6.3 The OLS Estimators, Predicted Values, and Residuals in the Multiple Regression Model 221
- 6.4 The Least Squares Assumptions for Causal Inference in the Multiple Regression Model 227
- 6.5 Large-Sample Distribution of $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ 228
- 6.6 The Least Squares Assumptions for Causal Inference in the Multiple Regression Model with Control Variables 233
- 7.1 Testing the Hypothesis $\beta_j = \beta_{j,0}$ Against the Alternative $\beta_j \neq \beta_{j,0}$ 249

- 7.2 Confidence Intervals for a Single Coefficient in Multiple Regression 250
- 7.3 R^2 and \bar{R}^2 : What They Tell You—and What They Don't 263
- 8.1 The Expected Change in Y from a Change in X_1 in the Nonlinear Regression Model [Equation (8.3)] 283
- 8.2 Logarithms in Regression: Three Cases 295
- 8.3 A Method for Interpreting Coefficients in Regressions with Binary Variables 299
- 8.4 Interactions Between Binary and Continuous Variables 302
- 8.5 Interactions in Multiple Regression 306
- 9.1 Internal and External Validity 331
- 9.2 Omitted Variable Bias: Should I Include More Variables in My Regression? 335
- 9.3 Functional Form Misspecification 336
- 9.4 Errors-in-Variables Bias 338
- 9.5 Sample Selection Bias 340
- 9.6 Simultaneous Causality Bias 343
- 9.7 Threats to the Internal Validity of a Multiple Regression Study 344

PART THREE Further Topics in Regression Analysis

- 10.1 Notation for Panel Data 362
- 10.2 The Fixed Effects Regression Model 369
- 10.3 The Fixed Effects Regression Assumptions 375
- 11.1 The Linear Probability Model 396
- 11.2 The Probit Model, Predicted Probabilities, and Estimated Effects 400
- 11.3 Logit Regression 402
- 12.1 The General Instrumental Variables Regression Model and Terminology 438
- 12.2 Two Stage Least Squares 440
- 12.3 The Two Conditions for Valid Instruments 441
- 12.4 The IV Regression Assumptions 442
- 12.5 A Rule of Thumb for Checking for Weak Instruments 446
- 12.6 The Overidentifying Restrictions Test (The J -Statistic) 449
- 14.1 m -Fold Cross Validation 523
- 14.2 The Principal Components of X 535

PART FOUR Regression Analysis of Economic Time Series Data

- 15.1 Lags, First Differences, Logarithms, and Growth Rates 557
- 15.2 Autocorrelation (Serial Correlation) and Autocovariance 559
- 15.3 Stationarity 562
- 15.4 Autoregressions 568
- 15.5 The Autoregressive Distributed Lag Model 571
- 15.6 The Least Squares Assumptions for Forecasting with Time Series Data 572
- 15.7 Pseudo Out-of-Sample Forecasts 575
- 15.8 The QLR Test for Coefficient Stability 592
- 16.1 The Distributed Lag Model and Exogeneity 616

- 16.2 The Distributed Lag Model Assumptions 618
- 16.3 HAC Standard Errors 624
- 17.1 Vector Autoregressions 650
- 17.2 Iterated Multi-period Forecasts 656
- 17.3 Direct Multi-period Forecasts 658
- 17.4 Orders of Integration, Differencing, and Stationarity 660
- 17.5 Cointegration 664

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

- 18.1 The Extended Least Squares Assumptions for Regression with a Single Regressor 689
- 19.1 The Extended Least Squares Assumptions in the Multiple Regression Model 715
- 19.2 The Multivariate Central Limit Theorem 718
- 19.3 Gauss–Markov Theorem for Multiple Regression 727
- 19.4 The GLS Assumptions 729

This page intentionally left blank

General Interest Boxes

The Distribution of Adulthood Earnings in the United Kingdom by Childhood Socioeconomic Circumstances	72
The Unpegging of the Swiss Franc	77
Financial Diversification and Portfolios	84
Off the Mark!	108
Social Class or Education? Childhood Circumstances and Adult Earnings Revisited	122
A Way to Increase Voter Turnout	124
The “Beta” of a Stock	152
The Economic Value of a Year of Education: Homoskedasticity or Heteroskedasticity?	193
Is Coffee Good for Your Health?	214
The Effect of Ageing on Healthcare Expenditures: A Red Herring?	304
The Demand for Economics Journals	307
Do Stock Mutual Funds Outperform the Market?	341
James Heckman and Daniel McFadden, Nobel Laureates	414
When Was Instrumental Variables Regression Invented?	430
The First IV Regression	447
The Externalities of Smoking	451
The Hawthorne Effect	480
Conditional Cash Transfers in Rural Mexico to Increase School Enrollment	483
Text as Data	543
Can You Beat the Market?	564
The River of Blood	577
Orange Trees on the March	635
NEWS FLASH: Commodity Traders Send Shivers Through Disney World	636
Nobel Laureates in Time Series Econometrics	680

This page intentionally left blank

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. Does healthcare spending actually improve health outcomes? 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? Does heavy intake of coffee lower the risk of disease or death? 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.

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
- Coverage of 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.pearsonglobaleditions.com) 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 4th 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