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/Erfle
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 to learn more.

Introduction to Econometrics

FOURTH EDITION

James H. Stock
Harvard University

Mark W. Watson
Princeton University

Pearson

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: 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
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, Cenveo Publisher Services
Interior Design: Cenveo Publisher Services
Cover Design: Studio Montage
Cover Art: Courtsey 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.

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://lccn.loc.gov/2018035117

Pearson

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

Brief Contents

Contents

Key Concepts

PART ONE Introduction and Review

PART TWO Fundamentals of Regression Analysis

PART THREE Further Topics in Regression Analysis

PART FOUR Regression Analysis of Economic Time Series Data

General Interest Boxes

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.

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**) 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

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								
	Part I	Part II		Part III		Part IV			Part V
Chapter	1–3	4–7, 9	8	10.1, 10.2	12.1, 12.2	15.1–15.4	15.5–15.8	16	18
10	X[a]	X[a]	X						
11	X[a]	X[a]	X						
12.1, 12.2	X[a]	X[a]	X						
12.3–12.6	X[a]	X[a]	X	X	X				
13	X[a]	X[a]	X	X	X				
14	X[a]	X[a]	X						
15	X[a]	X[a]	b						
16	X[a]	X[a]	b			X			
17	X[a]	X[a]	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.

[a]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.

[b]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 www.pearsonhighered.com	Features of the Supplement
Solutions Manual	Solutions to the end-of-chapter content.
Test Bank Authored by Manfred Keil, Claremont McKenna College	1,000 multiple-choice questions, essays and longer questions, and mathematical and graphical problems with these annotations: • Type (Multiple-choice, essay, graphical)
Computerized TestGen	TestGen allows instructors to: • Customize, save, and generate classroom tests • Edit, add, or delete questions from the Test Item Files • Analyze test results • Organize a database of tests and student results.
PowerPoints	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: • Keyboard and Screen Reader access • Alternative text for images • High color contrast between background and foreground colors
Companion Website	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 Cenveo 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 Cenveo 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 Kohlhase (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.