PEARSON Education

世纪高教·经济学英文版教材

Introduction to Econometrics

Second Edition

计量经济学

·第2版·

James H.Stock
[美] 詹姆斯·H.斯托克　哈佛大学经济系教授

Mark W.Watson
[美] 马克·W.沃森　普林斯顿大学经济系教授

著

世纪出版集团
上海人民出版社

PEARSON Education

世纪高教·经济学英文版教材

Introduction to Econometrics

Second Edition

计量经济学

·第2版·

James H.Stock
[美] 詹姆斯·H.斯托克　哈佛大学经济系教授

Mark W.Watson
[美] 马克·W.沃森　普林斯顿大学经济系教授

著

世纪出版集团
上海人民出版社

图书在版编目(CIP)数据

计量经济学 = Introduction to Econometrics:第2版:英文/(美)斯托克(Stock,J.H.),(美)沃森(Watson,M.W.)著.—上海:上海人民出版社,2007
ISBN 978-7-208-06924-4

Ⅰ.计… Ⅱ.①斯…②沃… Ⅲ.计量经济学—英文
Ⅳ.F224.0

中国版本图书馆CIP数据核字(2007)第033305号

责任编辑　钱　敏
封面设计　人马艺术设计工作室·储平

计量经济学(第2版)(英文版)
[美]詹姆斯·H.斯托克　马克·W.沃森　著

出　版　世纪出版集团　上海人民出版社
(200001　上海福建中路193号　www.ewen.cc)

出　品　上海世纪出版股份有限公司高等教育图书公司
www.hibooks.cn
世纪高教　(上海福建中路193号24层　021-63914988)

发　行　世纪出版集团发行中心
印　刷　上海商务联西印刷有限公司
开　本　787×1092毫米　1/16
印　张　53.75
插　页　2
字　数　1,000,000
版　次　2007年4月第1版
印　次　2007年4月第1次印刷
ISBN 978-7-208-06924-4/F·1566
定　价　78.00元

上海市版权局著作权合同登记号：图字 09-2007-222

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 ten minutes a day? Econometrics helps us to sort out sound ideas from crazy ones and to 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.

This textbook 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 theory is immediately relevant to the applications, this approach can make econometrics come alive.

The second edition benefits from the many constructive suggestions of teachers who used the first edition, while maintaining the philosophy that applications should drive the theory, not the other way around. The single greatest change in the second edition is a reorganization and expansion of the material on core regression analysis: Part II, which covers regression with cross-sectional data, has been expanded from four chapters to six. We have added new empirical examples (as boxes) drawn from economics and finance; some new optional sections on

classical regression theory; and many new exercises, both paper-and-pencil and computer-based empirical exercises using data sets newly placed on the textbook Web site. A more detailed description of changes to the second edition can be found on page xxxii.

Features of This Book

This textbook differs from others 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 textbook's companion Web site (**www.aw-bc.com/stock_watson**).

Contemporary Choice of Topics

Econometrics has come a long way in the past two decades. 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 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.*** An increasing number of 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.

- ***Forecasting.*** The chapter on forecasting (Chapter 14) 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 stochastic trends (unit roots), unit root tests, tests for 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.*** We make a clear distinction between two very different applications of time series regression: forecasting and estimation of dynamic causal effects. The chapter on causal inference using time series data (Chapter 15) 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 are large (hundreds of observations, often 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 textbook and other textbooks.

- ***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. Our experience is that 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 textbook Web site features data sets, software, and suggestions for empirical exercises of differing scopes. These web resources have been expanded considerably for the second edition.

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-IV of the text (which cover the substantive material) are accessible to students with only precalculus mathematics. Parts I-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.

This said, different students learn differently, and for the mathematically well-prepared students, learning can be enhanced by a more explicitly mathematical treatment. Part V therefore contains an introduction to econometric theory that

is appropriate for students with a stronger mathematical background. We believe that, when the mathematical chapters in Part V are used in conjunction with the material in Parts I-IV, this book is suitable for advanced undergraduate or master's level econometrics courses.

Changes to the Second Edition

The changes introduced in the second edition fall into three categories: more empirical examples; expanded theoretical material, especially in the treatment of the core regression topics; and additional student exercises.

More empirical examples. The second edition retains the empirical examples from the first edition and adds a significant number of new ones. These additional examples include estimation of the returns to education; inference about the gender gap in earnings; the difficulty of forecasting the stock market; and modeling the volatility clustering in stock returns. The data sets for these empirical examples are posted on the course Web site. The second edition also includes more general-interest boxes, for example how sample selection bias ("survivorship bias") can produce misleading conclusions about whether actively managed mutual funds actually beat the market.

Expanded theoretical material. The philosophy of this and the previous edition is that the modeling assumptions should be motivated by empirical applications. For this reason, our three basic least squares assumptions that underpin regression with a single regressor include neither normality nor homoskedasticity, both of which are arguably the exception in econometric applications. This leads directly to large-sample inference using heteroskedasticity-robust standard errors. Our experience is that students do not find this difficult—in fact, what they find difficult is the traditional approach of introducing the homoskedasticity and normality assumptions, learning how to use t- and F-tables, then being told that what they just learned is not reliable in applications because of the failure of these assumptions and that these "problems" must be "fixed." But not all instructors share this view, and some find it useful to introduce the homoskedastic normal regression model. Moreover, even if homoskedasticity is the exception instead of the rule, assuming homoskedasticity permits discussing the Gauss-Markov theorem, a key motivation for using ordinary least squares (OLS).

For these reasons, the treatment of the core regression material has been significantly expanded in the second edition, and now includes sections on the theoretical motivation for OLS (the Gauss-Markov theorem), small-sample inference in the homoskedastic normal model, and multicollinearity and the dummy vari-

able trap. To accommodate these new sections, the new empirical examples, the new general-interest boxes, and the many new exercises, the core regression chapters have been expanded from two to four: The linear regression model with a single regressor and OLS (Chapter 4); inference in regression with a single regressor (Chapter 5); the multiple regression model and OLS (Chapter 6); and inference in the multiple regression model (Chapter 7). This expanded and reorganized treatment of the core regression material constitutes the single greatest change in the second edition.

The second edition also includes some additional topics requested by some instructors. One such addition is specification and estimation of models that are nonlinear in the parameters (Appendix 8.1). Another is how to compute standard errors in panel data regression when the error term is serially correlated for a given entity (clustered standard errors; Section 10.5 and Appendix 10.2). A third addition is an introduction to current best practices for detecting and handling weak instruments (Appendix 12.5), and a fourth addition is a treatment, in a new final section of the last chapter (Section 18.7), of efficient estimation in the heteroskedastic linear IV regression model using generalized method of moments.

Additional student exercises. The second edition contains many new exercises, both "paper and pencil" and empirical exercises that involve the use of data bases, supplied on the course Web site, and regression software. The data section of the course Web site has been significantly enhanced by the addition of numerous databases.

Contents and Organization

There are five parts to the textbook. This textbook 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 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."

Part IV

Part IV takes up regression with time series data. Chapter 14 focuses on forecasting and introduces various modern tools for analyzing time series regressions such as unit root tests and tests for stability. Chapter 15 discusses the use of time series data to estimate causal relations. Chapter 16 presents some more advanced tools for time series analysis, including models of conditional heteroskedasticity.

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 17 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.

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	14.1–14.4	14.5–14.8	15	17
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]	[b]						
15	X[a]	X[a]	[b]			X			
16	X[a]	X[a]	[b]			X	X	X	
17	X	X	X						
18	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 15) 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 14.1–14.4.

[a]Chapters 10–16 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 14–16 (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.

Chapter 18 presents and studies the multiple regression model, instrumental variables regression, and generalized method of moments estimation of the linear model, 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 textbook 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.

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 of Part II). The course proceeds to cover regression with panel data (Chapter 10), regression with a limited dependent variable (Chapter 11), and/or instrumental variables regression (Chapter 12), as time permits. The course concludes with 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. *Prerequisites: Algebra II and introductory statistics.*

Introductory Econometrics with Time Series and Forecasting Applications

Like the standard introductory course, this course covers all of Part I (as needed) and all of 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 Part IV, covering forecasting (Chapter 14) and estimation of dynamic causal effects (Chapter 15). If time permits, the course can include some advanced topics in time series analysis such as volatility clustering and conditional heteroskedasticity (Section 16.5). *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 Part IV and works through forecasting (Chapter 14), estimation of dynamic causal effects (Chapter 15), and advanced topics in time series analysis (Chapter 16), including vector autoregressions and conditional heteroskedasticity. An important component of this course is hands-on forecasting exercises, available to instructors on the book's accompanying Web site. *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 17 and 18 (through section 18.5). 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 18.7), time series methods (Chapter 14), and/or the estimation of causal effects using time series data and generalized least squares (Chapter 15 and Section 18.6). *Prerequisites: calculus and introductory statistics. Chapter 18 assumes previous exposure to matrix algebra.*

Pedagogical Features

The textbook has a variety of pedagogical features aimed at helping students to understand, to retain, and to apply the essential ideas. *Chapter introductions* provide a real-world grounding and motivation, as well as a brief road map 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 numbered *Summary* concluding each chapter serves as a helpful framework for reviewing the main points of coverage. The questions in the *Review the Concepts* section check students' understanding of the core content, *Exercises* give more intensive practice working with the concepts and techniques introduced in the chapter, and *Empirical Exercises* allow the students to apply what they have learned to answer real-world empirical questions. At the end of the textbook, the *References* section lists sources for further reading, the *Appendix* provides statistical tables, and a *Glossary* conveniently defines all the key terms in the book.

Supplements to Accompany the Textbook

The online supplements accompanying the Second Edition of *Introduction to Econometrics* include the Solutions Manual, Test Bank (by Manfred W. Keil of Claremont McKenna College), and PowerPoint Lecture Notes with text figures,

tables, and Key Concepts. The Solutions Manual includes solutions to all the end-of-chapter exercises, while the Test Bank, offered in Test Generator Software (TestGen with QuizMaster), provides a rich supply of easily edited test problems and questions of various types to meet specific course needs. These resources are available for download from the Instructor's Resource Center at **www.aw-bc.com/irc.** If instructors prefer their supplements on a CD-ROM, our Instructor's Resource Disk, available for Windows and Macintosh, contains the PowerPoint Lecture Notes, the Test Bank, and the Solutions Manual.

In addition, a Companion Web site, found at **www.aw-bc.com/stock_watson**, provides a wide range of additional resources for students and faculty. These include data sets for all the text examples, replication files for empirical results reported in the text, data sets for the end-of-chapter *Empirical Exercises*, EViews and STATA tutorials for students, and an Excel add-in for OLS regressions.

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 co-teacher with one of the authors (Stock), she also helped to 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, 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 (Stata, Corp.), Jean Baldwin Grossman (Princeton), Eric Hanushek (the 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 their 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 10, and Alan Krueger (Princeton) for his help with the Tennessee STAR data that we analyze in Chapter 11.

We are also grateful for the many constructive, detailed, and thoughtful comments we received from those who reviewed various drafts for Addison-Wesley:

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.

Michael Abbott, Queen's University, Canada
Richard J. Agnello, University of Delaware
Clopper Almon, University of Maryland
Joshua Angrist, Massachusetts Institute of Technology
Swarnjit S. Arora, University of Wisconsin, Milwaukee
Christopher F. Baum, Boston College
McKinley L. Blackburn, University of South Carolina
Alok Bohara, University of New Mexico
Chi-Young Choi, University of New Hampshire
Dennis Coates, University of Maryland, Baltimore County
Tim Conley, Graduate School of Business, University of Chicago
Douglas Dalenberg, University of Montana