PREFACE

THE FIFTH EDITION OF ECONOMETRIC ANALYSIS

Econometric Analysis is intended for a one-year graduate course in econometrics for social scientists. The prerequisites for this course should include calculus, mathematical statistics, and an introduction to econometrics at the level of, say, Gujarati's *Basic Econometrics* (McGraw-Hill, 1995) or Wooldridge's *Introductory Econometrics: A Modern Approach* [South-Western (2000)]. Self-contained (for our purposes) summaries of the matrix algebra, mathematical statistics, and statistical theory used later in the book are given in Appendices A through D. Appendix E contains a description of numerical methods that will be useful to practicing econometricians. The formal presentation of econometrics begins with discussion of a fundamental pillar, the linear multiple regression model, in Chapters 2 through 8. Chapters 9 through 15 present familiar extensions of the single linear equation model, including nonlinear regression, panel data models, the generalized regression model, and systems of equations. The linear model is usually not the sole technique used in most of the contemporary literature. In view of this, the (expanding) second half of this book is devoted to topics that will extend the linear regression model in many directions. Chapters 16 through 18 present the techniques and underlying theory of estimation in econometrics, including GMM and maximum likelihood estimation methods and simulation based techniques. We end in the last four chapters, 19 through 22, with discussions of current topics in applied econometrics, including time-series analysis and the analysis of discrete choice and limited dependent variable models.

This book has two objectives. The first is to introduce students to *applied econometrics*, including basic techniques in regression analysis and some of the rich variety of models that are used when the linear model proves inadequate or inappropriate. The second is to present students with sufficient *theoretical background* that they will recognize new variants of the models learned about here as merely natural extensions that fit within a common body of principles. Thus, I have spent what might seem to be a large amount of effort explaining the mechanics of GMM estimation, nonlinear least squares, and maximum likelihood estimation and GARCH models. To meet the second objective, this book also contains a fair amount of theoretical material, such as that on maximum likelihood estimation and on asymptotic results for regression models. Modern software has made complicated modeling very easy to do, and an understanding of the underlying theory is important.

I had several purposes in undertaking this revision. As in the past, readers continue to send me interesting ideas for my "next edition." It is impossible to use them all,

of course. Because the five volumes of the *Handbook of Econometrics* and two of the *Handbook of Applied Econometrics* already run to over 4,000 pages, it is also unnecessary. Nonetheless, this revision is appropriate for several reasons. First, there are new and interesting developments in the field, particularly in the areas of microeconometrics (panel data, models for discrete choice) and, of course, in time series, which continues its rapid development. Second, I have taken the opportunity to continue fine-tuning the text as the experience and shared wisdom of my readers accumulates in my files. For this revision, that adjustment has entailed a substantial rearrangement of the material—the main purpose of that was to allow me to add the new material in a more compact and orderly way than I could have with the table of contents in the 4th edition. The literature in econometrics has continued to evolve, and my third objective is to grow with it. This purpose is inherently difficult to accomplish in a textbook. Most of the literature is written by professionals for other professionals, and this textbook is written for students who are in the early stages of their training. But I do hope to provide a bridge to that literature, both theoretical and applied.

This book is a broad survey of the field of econometrics. This field grows continually, and such an effort becomes increasingly difficult. (A partial list of journals devoted at least in part, if not completely, to econometrics now includes the *Journal of Applied Econometrics*, *Journal of Econometrics*, *Econometric Theory*, *Econometric Reviews*, *Journal of Business and Economic Statistics*, *Empirical Economics*, and *Econometrica*.) Still, my view has always been that the serious student of the field must start somewhere, and one *can* successfully seek that objective in a single textbook. This text attempts to survey, at an entry level, enough of the fields in econometrics that a student can comfortably move from here to practice or more advanced study in one or more specialized areas. At the same time, I have tried to present the material in sufficient generality that the reader is also able to appreciate the important common foundation of all these fields and to use the tools that they all employ.

There are now quite a few recently published texts in econometrics. Several have gathered in compact, elegant treatises, the increasingly advanced and advancing theoretical background of econometrics. Others, such as this book, focus more attention on applications of econometrics. One feature that distinguishes this work from its predecessors is its greater emphasis on nonlinear models. [Davidson and MacKinnon (1993) is a noteworthy, but more advanced, exception.] Computer software now in wide use has made estimation of nonlinear models as routine as estimation of linear ones, and the recent literature reflects that progression. My purpose is to provide a textbook treatment that is in line with current practice. The book concludes with four lengthy chapters on time-series analysis, discrete choice models and limited dependent variable models. These nonlinear models are now the staples of the applied econometrics literature. This book also contains a fair amount of material that will extend beyond many first courses in econometrics, including, perhaps, the aforementioned chapters on limited dependent variables, the section in Chapter 22 on duration models, and some of the discussions of time series and panel data models. Once again, I have included these in the hope of providing a bridge to the professional literature in these areas.

I have had one overriding purpose that has motivated all five editions of this work. For the vast majority of readers of books such as this, whose ambition is to use, not develop econometrics, I believe that it is simply not sufficient to recite the theory of estimation, hypothesis testing and econometric analysis. Understanding the often subtle

background theory is extremely important. But, at the end of the day, my purpose in writing this work, and for my continuing efforts to update it in this now fifth edition, is to show readers how to *do* econometric analysis. I unabashedly accept the unflattering assessment of a correspondent who once likened this book to a "user's guide to econometrics."

SOFTWARE AND DATA

There are many computer programs that are widely used for the computations described in this book. All were written by econometricians or statisticians, and in general, all are regularly updated to incorporate new developments in applied econometrics. A sampling of the most widely used packages and Internet home pages where you can find information about them are:

E-Views	`www.eviews.com`	(QMS, Irvine, Calif.)
Gauss	`www.aptech.com`	(Aptech Systems, Kent, Wash.)
LIMDEP	`www.limdep.com`	(Econometric Software, Plainview, N.Y.)
RATS	`www.estima.com`	(Estima, Evanston, Ill.)
SAS	`www.sas.com`	(SAS, Cary, N.C.)
Shazam	`shazam.econ.ubc.ca`	(Ken White, UBC, Vancouver, B.C.)
Stata	`www.stata.com`	(Stata, College Station, Tex.)
TSP	`www.tspintl.com`	(TSP International, Stanford, Calif.)

Programs vary in size, complexity, cost, the amount of programming required of the user, and so on. Journals such as *The American Statistician*, *The Journal of Applied Econometrics*, and *The Journal of Economic Surveys* regularly publish reviews of individual packages and comparative surveys of packages, usually with reference to particular functionality such as panel data analysis or forecasting.

With only a few exceptions, the computations described in this book can be carried out with any of these packages. We hesitate to link this text to any of them in particular. We have placed for general access a customized version of *LIMDEP*, which was also written by the author, on the website for this text, `www.prenhall.com/greene`. *LIMDEP* programs used for many of the computations are posted on the sites as well.

The data sets used in the examples are also on the website. Throughout the text, these data sets are referred to "TableFn.m," for example Table F4.1. The F refers to Appendix F at the back of the text, which contains descriptions of the data sets. The actual data are posted on the website with the other supplementary materials for the text. (The data sets are also replicated in the system format of most of the commonly used econometrics computer programs, including in addition to *LIMDEP*, *SAS*, *TSP*, *SPSS*, *E-Views*, and *Stata*, so that you can easily import them into whatever program you might be using.)

I should also note, there are now thousands of interesting websites containing software, data sets, papers, and commentary on econometrics. It would be hopeless to attempt any kind of a survey here. But, I do note one which is particularly agreeably structured and well targeted for readers of this book, the data archive for the

Journal of Applied Econometrics. This journal publishes many papers that are precisely at the right level for readers of this text. They have archived all the nonconfidential data sets used in their publications since 1994. This useful archive can be found at http://qed.econ.queensu.ca/jae/.

ACKNOWLEDGEMENTS

It is a pleasure to express my appreciation to those who have influenced this work. I am grateful to Arthur Goldberger and Arnold Zellner for their encouragement, guidance, and always interesting correspondence. Dennis Aigner and Laurits Christensen were also influential in shaping my views on econometrics. Some collaborators to the earlier editions whose contributions remain in this one include Aline Quester, David Hensher, and Donald Waldman. The number of students and colleagues whose suggestions have helped to produce what you find here is far too large to allow me to thank them all individually. I would like to acknowledge the many reviewers of my work whose careful reading has vastly improved the book: Badi Baltagi, University of Houston: Neal Beck, University of California at San Diego; Diane Belleville, Columbia University; Anil Bera, University of Illinois; John Burkett, University of Rhode Island; Leonard Carlson, Emory University; Frank Chaloupka, City University of New York; Chris Cornwell, University of Georgia; Mitali Das, Columbia University; Craig Depken II, University of Texas at Arlington; Edward Dwyer, Clemson University; Michael Ellis, Wesleyan University; Martin Evans, New York University; Ed Greenberg, Washington University at St. Louis; Miguel Herce, University of North Carolina; K. Rao Kadiyala, Purdue University; Tong Li, Indiana University; Lubomir Litov, New York University; William Lott, University of Connecticut; Edward Mathis, Villanova University; Mary McGarvey, University of Nebraska-Lincoln; Ed Melnick, New York University; Thad Mirer, State University of New York at Albany; Paul Ruud, University of California at Berkeley; Sherrie Rhine, Chicago Federal Reserve Board; Terry G. Seaks, University of North Carolina at Greensboro; Donald Snyder, California State University at Los Angeles; Steven Stern, University of Virginia; Houston Stokes, University of Illinois at Chicago; Dimitrios Thomakos, Florida International University; Paul Wachtel, New York University; Mark Watson, Harvard University; and Kenneth West, University of Wisconsin. My numerous discussions with B. D. McCullough have improved Appendix E and at the same time increased my appreciation for numerical analysis. I am especially grateful to Jan Kiviet of the University of Amsterdam, who subjected my third edition to a microscopic examination and provided literally scores of suggestions, virtually all of which appear herein. Chapters 19 and 20 have also benefited from previous reviews by Frank Diebold, B. D. McCullough, Mary McGarvey, and Nagesh Revankar. I would also like to thank Rod Banister, Gladys Soto, Cindy Regan, Mike Reynolds, Marie McHale, Lisa Amato, and Torie Anderson at Prentice Hall for their contributions to the completion of this book. As always, I owe the greatest debt to my wife, Lynne, and to my daughters, Lesley, Allison, Elizabeth, and Julianna.

William H. Greene