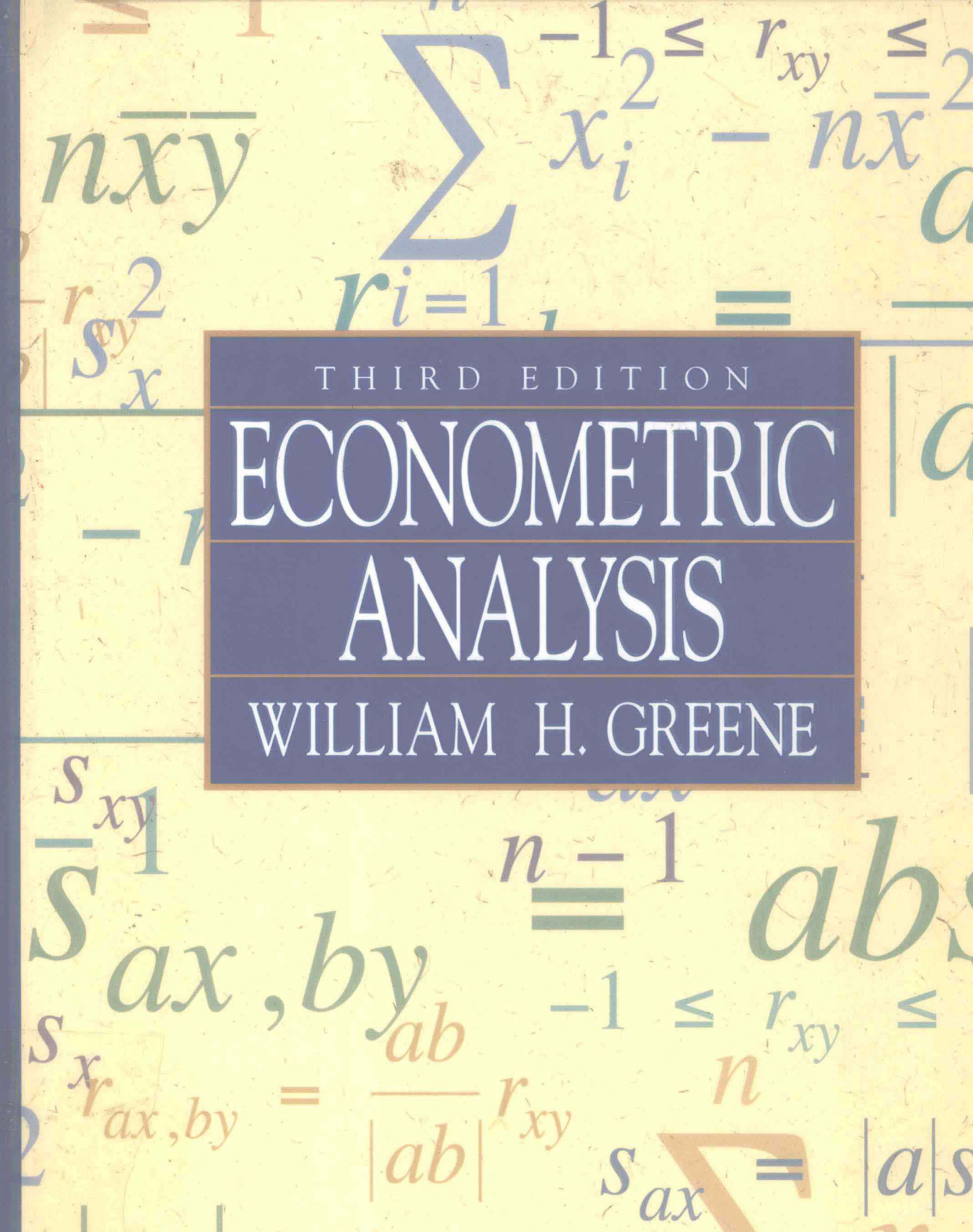
THIRD EDITION
ECONOMETRIC ANALYSIS
WILLIAM H. GREENE

Econometric Analysis

Third Edition

William H. Greene

New York University

Prentice Hall, Upper Saddle River, New Jersey 07458

Executive Editor: Leah Jewell
Editorial Assistant: Kristen Kaiser
Editor-in-Chief: James Boyd
Marketing Manager: Susan McLaughlin
Production Editor: Judith Leale
Production Coordinator: Renee Pelletier
Managing Editor: Carol Burgett
Manufacturing Buyer: Kenneth J. Clinton
Manufacturing Supervisor: Arnold Vila
Manufacturing Manager: Vincent Scelta
Cover Design: Bruce Kenselaar
Composition: Progressive Information Technologies

Library of Congress Cataloging-in-Publication Data

Greene, William H. 1951–
Econometric analysis / William H. Greene. — 3rd ed.
p. cm.
Includes bibliographical references and indexes.
ISBN 0-02-346602-2
1. Econometrics.
HB139.G74 1997 96-32586
330′ .01′5195 — dc20 CIP

Prentice-Hall International (UK) Limited, London
Prentice-Hall of Australia Pty. Limited, Sydney
Prentice-Hall Canada, Inc., Toronto
Prentice-Hall Hispanoamericana, S.A., Mexico
Prentice-Hall of India Private Limited, New Delhi
Prentice-Hall of Japan, Inc., Tokyo
Simon & Schuster Asia Pte. Ltd., Singapore
Editora Prentice-Hall do Brasil, Ltda., Rio de Janeiro

Printed in the United States of America

10 9 8 7 6 5 4

Econometric Analysis

For
Lesley, Elizabeth, Allison, and Julianna

Preface

This third edition of *Econometric Analysis* is intended for a one-year graduate course in econometrics for social scientists. The prerequisites for the course should include calculus, basic mathematical statistics, and an introduction to the paradigm of econometrics at the level of, say, Gujarati's *Basic Econometrics* (McGraw-Hill, 1995), Maddala's *Introduction to Econometrics* [Macmillan (now Prentice Hall), 1992], or Griffiths, Hill, and Judge's *Learning and Practicing Econometrics* (John Wiley and Sons, 1993). Self-contained (for our purposes) summaries of the matrix algebra, statistical theory, and mathematical statistics used later in the book are given in Chapters 2 through 4. Chapter 5 is mostly new and contains a description of numerical methods that will be useful to practicing econometricians. The formal presentation of econometrics begins in Chapters 6 through 10 with discussion of the fundamental building block, the linear multiple regression model. Chapters 11 through 16 present familiar extensions of the single linear equation model, including nonlinear regression, panel data models, the generalized regression model, and systems of equations. We end in the last four chapters with discussions of current topics in applied econometrics including GMM estimation methods, Lagrange multiplier tests, time-series analysis, and the analysis of qualitative 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 objective is to present students with sufficient *theoretical background* that they will recognize new variants of the models that they learn about here as merely natural extensions that fit within a common body of principles. This is why 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, for example, of 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.

Readers of my second edition will see that this work is a major revision. I had three purposes in undertaking this revision. The first was to respond to the many readers who generously wrote to me with interesting suggestions for my "next edition." Had I followed all their suggestions, I would have (a lifetime of work later) produced an enormous reference work in econometrics. But the four volumes of the *Handbook of Econometrics* already run to over 3000 pages, and volumes 1 and 2 of the *Handbook of Applied Econometrics* will soon appear, so this is hardly necessary. Nonetheless, when I was given the opportunity to do this revision by my publisher, I found the invitation irresistible. This has involved extending some topics that were already present in the second edition and adding many new ones. In fact, a revision of the earlier material on time series was essential. Second, it became obvious to me that the second edition could greatly benefit from some reorganization and some editing. I hope that this will make the flow of the book more logical. Finally, the literature in econometrics has continued to evolve, and my third objective is to continue to grow with it. This is inherently difficult to do with a textbook. Most of the literature is written by professionals for other professionals, and this is a textbook written for students who are in the early stages of their training. But I do hope to provide a bridge to that literature.

There are four components to this revision.

1. New topics in computational methods. I have added a new chapter on numerical methods that includes discussions of digital computing, random number generation, bootstrap sampling, Monte Carlo integration, importance sampling, the Gibbs sampler, numerical quadrature, and optimization methods. The last of these was covered in the previous edition, but I have added a discussion of the EM algorithm to show students an important application of the link between the theoretical specification and the computation of numerical estimates.

2. Estimation methods. I have expanded the treatments of both the maximum likelihood and GMM estimators. Chapter 6 now contains considerably more discussion of asymptotic results for the linear model. This discussion is fairly general so that it, with the earlier material on ML and GMM estimation, can be extended to, for example, discrete choice models without further development. Theoretical background and several applications of two-step estimation have also been added. Finally, the discussion of Bayesian estimation of the linear model has been expanded and moved to Chapter 6 with the other basic material on the multiple linear regression model.

3. Specific models and techniques. This edition contains new material on several models that are used in the econometrics literature. These include time-series models including unit roots, VARs, and cointegration, Box–Jenkins ARMA models (mainly as a stepping-stone to the literature on unit roots and VARs, not as a method of building forecasting models as such), the nested logit model, various specifications for panel data including dynamic panels and ro-

bust estimation in a panel data setting, new material on binary choice including a nonparametric, kernel estimator, and newly developed models for count data.

4. Reorganization. To improve the flow of the book, I have rearranged or moved several chapters. The new chapter on numerical methods follows Chapter 4 on statistical inference. This chapter also includes all of what was Chapter 12 on optimization in the second edition, so that all of the "tools" are now gathered at the beginning of the book. Chapter 5 on the bivariate regression model has been eliminated, since most of it was repeated, more compactly, in Chapter 6. The discussion of the linear regression model begins in Chapter 6 with multiple regression. This chapter now also includes all of the previous Chapter 10 on asymptotics as well as the part of the previous Chapter 8 on instrumental variables estimation. There is now an extensive treatment of asymptotic theory in Chapter 6. Some of the material on hypothesis testing in the old Chapter 6 that overlapped with Chapter 7 has been moved to Chapter 7 to improve the organization of this material. The material on the linear regression model has been consolidated a bit. Chapter 6 is now quite long, but the arrangement is somewhat more natural than it was in the second edition. The previous chapter on covariance structures and panel data is now specialized in panel data models, mostly fixed and random effects. Some new material on robust estimation of covariance matrices and dynamic models has been added to this chapter. All the material on covariance structures has been moved to the next chapter on sets of regression equations, where the model of groupwise covariance structures and the random coefficients model have been reoriented as special cases of the seemingly unrelated regressions model. The chapter on simultaneous equations now follows, I believe more naturally, immediately after the chapter on sets of regressions. The two chapters on time-series models, followed by the two chapters on qualitative and limited dependent variables, now appear as the latter chapters in the book, more or less in the fashion of a topics section. As before, the heading "advanced topics" is inappropriate here. These chapters document the mainstream of two major areas of research.

Many authors have used the term *explosion* in the literature to describe the most clearly visible evolution in econometrics, the new methods in time-series analysis and empirical macroeconomics. It has seemed to some observers that the term was more apt than the authors might have intended; for some time, the literature seemed to be huge, but fragmented and everywhere at once. But the field seems to have found an equilibrium, and the current body of literature documents an exciting new paradigm. I have included in this edition a considerable expansion of the sections on unit roots, VARs, and cointegration. But this constitutes only the barest introduction to this literature, and I have no illusions that this will satisfy critics of the earlier editions who looked here for a comprehensive, intermediate-level introduction. That more than one correspondent thought that perhaps a second volume might be a feasible approach for this text suggests how difficult that would be to provide. Fortunately, a half dozen new books are now available for the reader who wishes to continue studies in this area. A good place to begin is Enders (1995), followed by Hamilton (1994). There are too many survey papers to list fully—another mini-

growth industry—but two that are noteworthy are Stock (1994) and Watson (1994). Others are listed in Chapters 17 and 18.

Another sea of change in econometrics is the appearance of robust methods for estimation and testing. This is dominated by robust covariance matrix estimates, GMM estimation, and the redemption of the Wald statistic. It has been but a decade or less since pure significance tests were rendered déclassé under the assault of a wave of Lagrange multiplier tests. But recent developments in GMM estimation and distribution free estimation have brought the Wald statistic back from its ignominious fate. In this edition, I have attempted to apply many of the newer methods in estimation, including two-step and GMM estimation and LM and conditional moment testing, in a variety of settings. Likewise, the reader will find numerous applications of robust covariance estimation and hypothesis testing using several methods.

I have attempted to keep the mathematical level consistent throughout. This has meant liberal use of matrix algebra, but it has required little in the way of advanced distribution theory. I give proofs only when they are particularly revealing about some underlying principle that will appear in other contexts or provide students with a useful tool for their work. White's proof of the limiting distribution of the Wald statistic in Chapter 6 is an example. In contrast, a proof of the central limit theorem, although obviously of great utility in its own right, is a one-shot deal. For those who are teaching at a relatively high level and who desire more of a theorem/proof format, I suggest Peter Schmidt's *Econometrics* (Marcel Dekker, 1976) as a very handy adjunct. Also, Davidson and MacKinnon (1993) present many of the topics that we cover at a higher, more theoretical level, but with an appealing operational flavor.

One feature that distinguishes this work from its predecessors is its greater emphasis on nonlinear models, including full chapters on nonlinear regression and nonlinear optimization. [Davidson and MacKinnon (1993) are 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. The purpose of these chapters is to bring the textbook treatment in line with current practice. I have also included two long chapters on limited dependent variable models. These nonlinear models are now common in the applied literature. I have written these chapters because there is still no other source that presents these topics at a level elementary enough to initiate the newcomer but complete enough to enable a diligent student to use the information to undertake a serious empirical study. This book 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 20 on duration models, and some of the discussion of time series. Once again, I have included these in the hope of providing a heretofore missing bridge to the professional literature in these areas.

With the current state of microcomputers, students can be given realistic data sets and challenging empirical analyses as a routine part of their econo-

metrics training. To this end, I have included in this book a large number of data sets, many of which have been used in studies already in the literature. In addition, the appendix to Chapter 16 contains a yearly data set on a number of macroeconomic variables. These could be used, for example, to update Klein's Model I or, for the more ambitious, to estimate a new model.

There are many computer programs that students can use in an econometrics course. The most important features are ease of use and flexibility; the same program can easily be used for many different types of analyses. Most of these programs are available in both mainframe and microcomputer versions. The programs vary in size and complexity, cost, and the amount of programming required of the user. Journals such as the *American Statistician* and the *Journal of Applied Econometrics* present frequent surveys and reviews of econometric software. An intriguing development of recent years has been the worldwide dissemination of routines written for other packages such as Gauss and LIMDEP via various list servers and the World Wide Web. (For example, the CodEc project offers a freely distributed library of documentation and many kinds of programs to users of the Internet.)

It is a pleasure to express my appreciation to those who have influenced this work. My gratitude to Arthur Goldberger continues, for his encouragement, guidance, and always interesting correspondence. Dennis Aigner and Laurits Christensen were also influential in shaping my views on econometrics. 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. The several collaborators to the second edition, whose contributions I have retained in this one, include Aline Quester, David Hensher, Donald Waldman, Martin Evans, and Paul Wachtel. This work has also benefited at several stages from the careful reading of many reviewers, including Badi Baltagi, University of Houston; Diane Belleville of New York University; Leonard Carlson, Emory University; Chris Cornwell, University of Georgia; Michael Ellis, Wesleyan University; K. Rao Kadiyala, Purdue University; William Lott, University of Connecticut; Edward Mathis, Villanova University; Thad Mirer, State University of New York at Albany; Terry G. Seaks, University of North Carolina at Greensboro; Donald Snyder, California State University at Los Angeles; Houston Stokes, University of Illinois at Chicago; Mark Watson, Harvard University; Kenneth West, University of Wisconsin; Ed Greenberg, Washington University at St. Louis; Edward Dwyer, Clemson University; Frank Chaloupka, City University of New York; Arnold Zellner, University of Chicago; Dimitrios Thomakos at Columbia University; Paul Ruud, University of California at Berkeley; Neal Beck, University of California at San Diego; Anil Bera, University of Illinois; and Miguel Herce, University of North Carolina. I would also like to thank Leah Jewell at Prentice Hall, and Donna King at Progressive Publishing Alternatives, 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

Examples