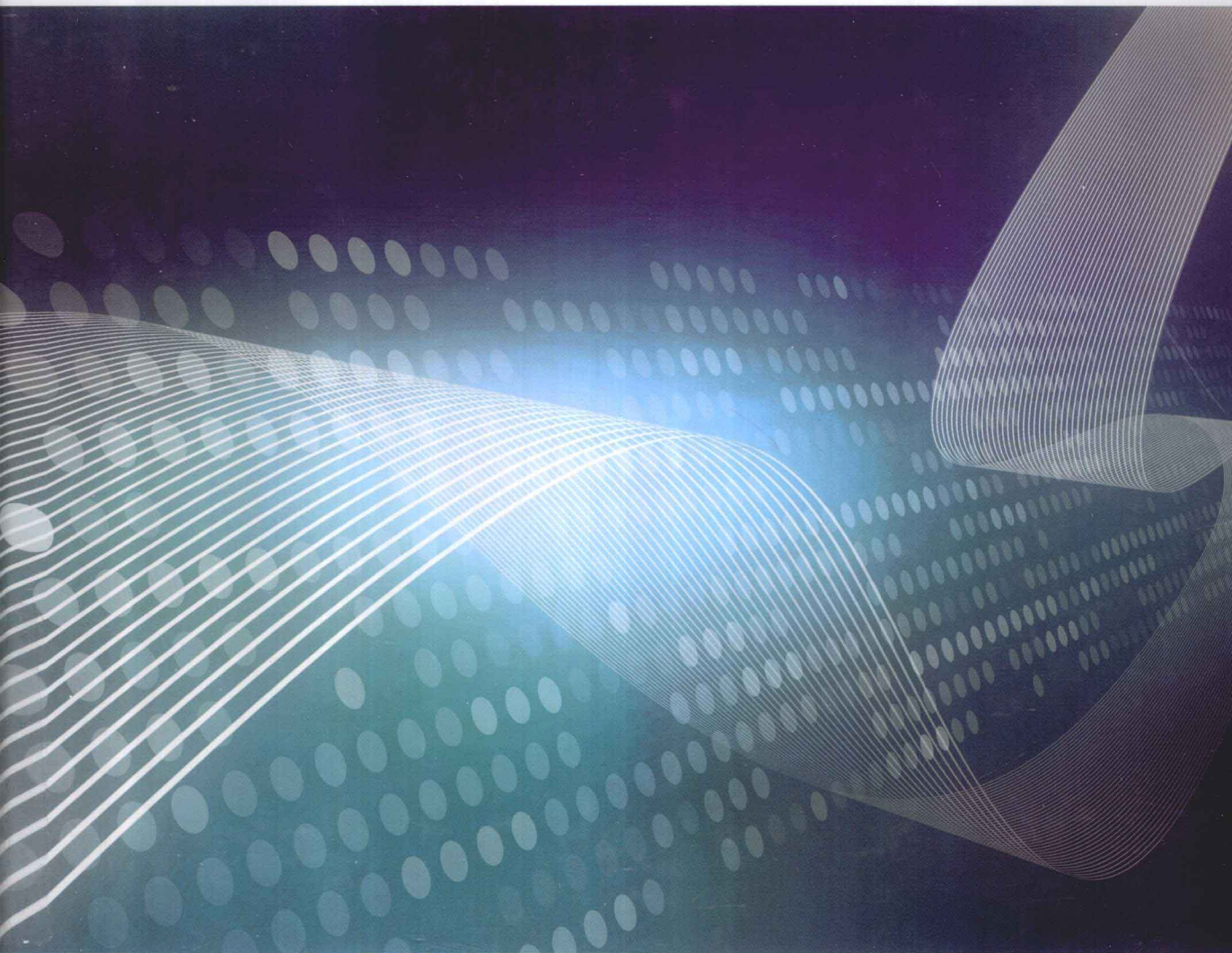


ECONOMETRIC ANALYSIS

SEVENTH EDITION



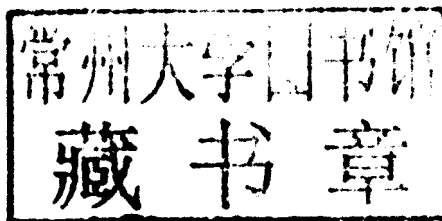
WILLIAM H. GREENE

SEVENTH EDITION
ECONOMETRIC ANALYSIS



William H. Greene

New York University



Prentice Hall

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montréal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

For Margaret and Richard Greene

Editorial Director: *Sally Yagan*
Editor in Chief: *Donna Battista*
Acquisitions Editor: *Adrienne D'Ambrosio*
Editorial Project Manager: *Jill Kolongowski*
Director of Marketing: *Patrice Jones*
Senior Marketing Manager: *Lori DeShazo*
Managing Editor: *Nancy Fenton*
Production Project Manager: *Carla Thompson*
Manufacturing Director: *Evelyn Beaton*
Senior Manufacturing Buyer: *Carol Melville*
Creative Director: *Christy Mahon*
Cover Designer: *Pearson Central Design*

Cover Image: *Ralf Hiemisch/Getty Images*
Permissions Project Supervisor: *Michael Joyce*
Media Producer: *Melissa Honig*
Associate Production Project Manager:
Alison Eusden
Full-Service Project Management:
MPS Limited, a Macmillan Company
Composition: *MPS Limited, a Macmillan Company*
Printer/Binder: *Courier/Westford*
Cover Printer: *Lehigh-Phoenix*
Color/Hagerstown
Text Font: *10/12 Times*

Credits and acknowledgments for material borrowed from other sources and reproduced, with permission, in this textbook appear on appropriate page within text.

Copyright © 2012 Pearson Education, Inc., publishing as Prentice Hall, One Lake Street, Upper Saddle River, NJ 07458. All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright, and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use material from this work, please submit a written request to Pearson Education, Inc., Permissions Department, 501 Boylston Street, Suite 900, Boston, MA 02116, fax your request to 617-671-3447, or e-mail at <http://www.pearsoned.com/legal/permission.htm>.

Many of the designations by manufacturers and seller to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed in initial caps or all caps.

Library of Congress Cataloging-in-Publication Data

Greene, William H., 1951–
Econometric analysis / William H. Greene. — 7th ed.
p. cm.
ISBN 0-13-139538-6
1. Econometrics. I. Title.
HB139.G74 2012
330.01'5195—dc22

2010050532

10 9 8 7 6 5 4 3 2

Prentice Hall
is an imprint of

PEARSON

www.pearsonhighered.com

ISBN 10: 0-13-139538-6
ISBN 13: 978-0-13-139538-1

PEARSON SERIES IN ECONOMICS

Abel/Bernanke/Croushore

*Macroeconomics**

Bade/Parkin

*Foundations of Economics**

Berck/Helfand

The Economics of the Environment

Bierman/Fernandez

Game Theory with Economic

Applications

Blanchard

*Macroeconomics**

Blau/Ferber/Winkler

The Economics of Women,

Men and Work

Boardman/Greenberg/Vining/

Weimer

Cost-Benefit Analysis

Boyer

Principles of Transportation Economics

Branson

Macroeconomic Theory and Policy

Brock/Adams

The Structure of American Industry

Bruce

Public Finance and the American

Economy

Carlton/Perloff

Modern Industrial Organization

Case/Fair/Oster

*Principles of Economics**

Caves/Frankel/Jones

World Trade and Payments:

An Introduction

Chapman

Environmental Economics:

Theory, Application, and Policy

Cooter/Ulen

Law & Economics

Downs

An Economic Theory of Democracy

Ehrenberg/Smith

Modern Labor Economics

Ekelund/Ressler/Tollison

*Economics**

Farnham

Economics for Managers

Folland/Goodman/Stano

The Economics of Health and Health

Care

Fort

Sports Economics

Froyen

Macroeconomics

Fusfeld

The Age of the Economist

Gerber

*International Economics**

Gordon

*Macroeconomics**

Greene

Econometric Analysis

Gregory

Essentials of Economics

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

Hoffman/Averett

Women and the Economy:

Family, Work, and Pay

Holt

Markets, Games and Strategic

Behavior

Hubbard/O'Brien

*Economics**

Money, Banking, and the Financial

*System**

Hughes/Cain

American Economic History

Husted/Melvin

International Economics

Jehle/Reny

Advanced Microeconomic Theory

Johnson-Lans

A Health Economics Primer

Keat/Young

Managerial Economics

Klein

Mathematical Methods for Economics

Krugman/Obstfeld/Melitz

International Economics:

*Theory & Policy**

Laidler

The Demand for Money

Leeds/von Allmen

The Economics of Sports

Leeds/von Allmen/Schiming

*Economics**

Lipsey/Ragan/Storer

*Economics**

Lynn

Economic Development: Theory and

Practice for a Divided World

Miller

*Economics Today**

Understanding Modern

Economics

Miller/Benjamin

The Economics of Macro Issues

Miller/Benjamin/North

The Economics of Public Issues

Mills/Hamilton

Urban Economics

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

Nafziger

The Economics of Developing

Countries

O'Sullivan/Sheffrin/Perez

Economics: Principles, Applications and

*Tools**

Parkin

*Economics**

Perloff

*Microeconomics**

Microeconomics: Theory and

*Applications with Calculus**

Perman/Common/McGilvray/Ma

Natural Resources and

Environmental Economics

Phelps

Health Economics

Pindyck/Rubinfeld

*Microeconomics**

Riddell/Shackelford/Stamos/

Schneider

Economics: A Tool for Critically

Understanding Society

Ritter/Silber/Udell

Principles of Money, Banking &

*Financial Markets**

Roberts

The Choice: A Fable of Free Trade and

Protection

Rohlf

Introduction to Economic Reasoning

Ruffin/Gregory

Principles of Economics

Sargent

Rational Expectations and Inflation

Sawyer/Sprinkle

International Economics

Scherer

Industry Structure, Strategy, and Public

Policy

Schiller

The Economics of Poverty and

Discrimination

Sherman

Market Regulation

Silberberg

Principles of Microeconomics

Stock/Watson

Introduction to Econometrics

Introduction to Econometrics, Brief

Edition

Studenmund

Using Econometrics: A Practical Guide

Tietenberg/Lewis

Environmental and Natural Resource

Economics

Environmental Economics and Policy

Todaro/Smith

Economic Development

Waldman

Microeconomics

Waldman/Jensen

Industrial Organization: Theory and

Practice

Weil

Economic Growth

Williamson

Macroeconomics

PREFACE



ECONOMETRIC ANALYSIS

Econometric Analysis provides a broad introduction to the field of econometrics. This field grows continually—a list of journals devoted at least in part, if not completely, to econometrics now includes *The Journal of Applied Econometrics*, *The Journal of Econometrics*, *The Econometrics Journal*, *Econometric Theory*, *Econometric Reviews*, *Journal of Business and Economic Statistics*, *Empirical Economics*, *Foundations and Trends in Econometrics*, *The Review of Economics and Statistics*, and *Econometrica*. Constructing a textbook-style survey to introduce the topic at a graduate level has become increasingly ambitious. Nonetheless, I believe that one *can* successfully seek that objective in a single textbook. This text attempts to present, at an entry level, enough of the topics in econometrics that a student can comfortably move from here to practice or more advanced study in one or more specialized areas. The book is also intended as a bridge for students and analysts in the social sciences between an introduction to the field and the professional literature.

NEW TO THIS EDITION

This seventh edition is a major revision of *Econometric Analysis*. Among the most obvious changes are

- Reorganization of the early material that is taught in the first-semester course, including
 - All material on hypothesis testing and specification presented in a single chapter
 - New results on prediction
 - Greater and earlier emphasis on instrumental variables and endogeneity
 - Additional results on basic panel data models
- New applications and examples, with greater detail
- Greater emphasis on specific areas of application in the advanced material
- New material on simulation-based methods, especially bootstrapping and Monte Carlo studies
- Several examples that explain interaction effects
- Specific recent applications including quantile regression
- New applications in discrete choice modeling
- New material on endogeneity and its implications for model structure

PLAN OF THE BOOK

The arrangement of the book is as follows:

Part I begins the formal development of econometrics with its fundamental pillar, the *linear multiple regression model*. Estimation and inference with the linear least squares estimator are analyzed in Chapters 2 through 6. The *nonlinear regression model* is introduced in Chapter 7 along with quantile, semi- and nonparametric regression, all as extensions of the familiar linear model. *Instrumental variables estimation* is developed in Chapter 8.

Part II presents three major extensions of the regression model. Chapter 9 presents the consequences of relaxing one of the main assumptions of the linear model, homoscedastic nonautocorrelated disturbances, to introduce the *generalized regression model*. The focus here is on heteroscedasticity; autocorrelation is mentioned, but a detailed treatment is deferred to Chapter 20 in the context of time-series data. Chapter 10 introduces systems of regression equations, in principle, as the approach to modeling simultaneously a set of random variables and, in practical terms, as an extension of the generalized linear regression model. Finally, panel data methods, primarily fixed and random effects models of heterogeneity, are presented in Chapter 11.

The second half of the book is devoted to topics that will extend the linear regression model in many directions. Beginning with Chapter 12, we proceed to the more involved methods of analysis that contemporary researchers use in analysis of “real-world” data. Chapters 12 to 16 in Part III present different estimation methodologies. Chapter 12 presents an overview by making the distinctions between *parametric*, *semiparametric* and *nonparametric methods*. The leading application of semiparametric estimation in the current literature is the *generalized method of moments (GMM) estimator* presented in Chapter 13. This technique provides the platform for much of modern econometrics. *Maximum likelihood estimation* is developed in Chapter 14. Monte Carlo and simulation-based methods such as bootstrapping that have become a major component of current research are developed in Chapter 15. Finally, *Bayesian methods* are introduced in Chapter 16.

Parts IV and V develop two major subfields of econometric methods, *microeconomics*, which is typically based on cross-section and panel data, and *macroeconometrics*, which is usually associated with analysis of time-series data. In Part IV, Chapters 17 to 19 are concerned with models of discrete choice, censoring, truncation, sample selection, duration, treatment effects, and the analysis of counts of events. In Part V, Chapters 20 and 21, we consider two topics in time-series analysis, models of serial correlation and regression models for nonstationary data—the usual substance of macroeconomic analysis.

REVISIONS

I have substantially rearranged the early part of the book to produce what I hope is a more natural sequence of topics for the graduate econometrics course. Chapter 4 is now devoted entirely to point and interval estimation, including prediction and forecasting. Finite sample, then asymptotic properties of least squares are developed in detail. All

of the material on hypothesis testing and specification search is moved into Chapter 5, rather than fragmented over several chapters as in the sixth edition. I have also brought the material on instrumental variables much farther forward in the text, from after the development of the generalized regression model in the sixth edition to Chapter 8 in this one, immediately after full development of the linear regression model. This accords with the greater emphasis on this method in recent applications. A very large number of other rearrangements of the material will also be evident. Chapter 7 now contains a range of advanced extensions of the linear regression model, including nonlinear, quantile, partially linear, and nonparametric regression. This is also a point at which the differences between parametric, semiparametric, and nonparametric methods can be examined. One conspicuous modification is the excision of the long chapter on linear simultaneous equations models. Some of the material from this chapter appears elsewhere. Two-stage least squares now appears with instrumental variables estimation. Remaining parts of this chapter that are of lesser importance in recent treatments, such as rank and order conditions for identification of linear models and 3SLS and FIML estimation, have been deleted or greatly reduced and placed in context elsewhere in the text. The material on discrete choice models has been rearranged to orient the topics to the behavioral foundations. Chapter 17 now broadly introduces discrete choice and random utility models, and then builds on variants of the binary choice model. The analysis is continued in Chapter 18 with unordered, then ordered choice models and, finally, models for counts. The last chapter of the section studies models for continuous variables in the contexts of particular data-generating mechanisms and behavioral contexts.

I have added new material and some different examples and applications at numerous points. Topics that have been expanded or given greater emphasis include treatment effects, bootstrapping, simulation-based estimation, robust estimation, missing and faulty data, and a variety of different new methods of discrete choice analysis in microeconometrics. I have also added or expanded material on techniques recently of interest, such as quantile regression and stochastic frontier models.

I note a few specific highlights of the revision: In general terms, I have increased the focus on robust methods a bit. I have placed discussions of specification tests at several points, consistent with the trend in the literature to examine more closely the fragility of heavily parametric models. A few of the specific new applications are as follows:

- Chapter 15 on simulation-based estimation has been considerably expanded. It now includes substantially more material on bootstrapping standard errors and confidence intervals. The Krinsky and Robb (1986) approach to asymptotic inference has been placed here as well.
- A great deal of attention has been focused in recent papers on how to understand interaction effects in nonlinear models. Chapter 7 contains a lengthy application of interaction effects in a nonlinear (exponential) regression model. The issue is revisited in Chapter 17.
- As an exercise that will challenge the student's facility with asymptotic distribution theory, I have added a detailed proof of the Murphy and Topel (2002) result for two-step estimation in Chapter 14.
- Sources and treatment of endogeneity appear at various points, for example an application of inverse probability weighting to deal with attrition in Chapter 17.

The seventh edition is a major revision of *Econometric Analysis* both in terms of organization of the material and in terms of new ideas and treatments. I hope that readers will find the changes helpful.

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

<i>EViews</i>	www.eviews.com	(QMS, Irvine, CA)
<i>Gauss</i>	www.aptech.com	(Aptech Systems, Kent, WA)
<i>LIMDEP</i>	www.limdep.com	(Econometric Software, Plainview, NY)
<i>MATLAB</i>	www.mathworks.com	(Mathworks, Natick, MA)
<i>NLOGIT</i>	www.nlogit.com	(Econometric Software, Plainview, NY)
<i>R</i>	www.r-project.org/	(The R Project for Statistical Computing)
<i>RATS</i>	www.estima.com	(Estima, Evanston, IL)
<i>SAS</i>	www.sas.com	(SAS, Cary, NC)
<i>Shazam</i>	econometrics.com	(Northwest Econometrics Ltd., Gibsons, Canada)
<i>Stata</i>	www.stata.com	(Stata, College Station, TX)
<i>TSP</i>	www.tspintl.com	(TSP International, Stanford, CA)

A more extensive list of computer software used for econometric analysis can be found at the resource Web site, <http://www.oswego.edu/~economic/econsoftware.htm>.

With only a few exceptions, the computations described in this book can be carried out with any of the packages listed. *NLOGIT* was used for the computations in the applications. This text contains no instruction on using any particular program or language. (The author's Web site for the text does provide some code and data for replication of the numerical examples.) Many authors have produced *RATS*, *LIMDEP/NLOGIT*, *EViews*, *SAS*, or *Stata* code for some of our applications, including, in a few cases, the documentation for their computer programs. There are also quite a few volumes now specifically devoted to econometrics associated with particular packages, such as Cameron and Trivedi's (2009) companion to their treatise on microeconomics.

The data sets used in the examples are also available on the Web site for the text, <http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm>. Throughout the text, these data sets are referred to "Table Fn.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 in generic ASCII and portable formats on the Web site with the other supplementary materials for the text. There are now thousands of interesting Web sites containing software, data sets, papers, and commentary on econometrics. It would be hopeless to attempt any kind of a survey here. One code/data site that is particularly agreeably structured and well targeted for readers of this book is

the data archive for the *Journal of Applied Econometrics*. They have archived all the nonconfidential data sets used in their publications since 1988 (with some gaps before 1995). This useful site can be found at <http://qed.econ.queensu.ca/jae/>. Several of the examples in the text use the *JAE* data sets. Where we have done so, we direct the reader to the *JAE*'s Web site, rather than our own, for replication. Other journals have begun to ask their authors to provide code and data to encourage replication. Another vast, easy-to-navigate site for aggregate data on the U.S. economy is www.economagic.com.

ACKNOWLEDGMENTS

It is a pleasure to express my appreciation to those who have influenced this work. I remain grateful to Arthur Goldberger (dec.), Arnold Zellner (dec.), Dennis Aigner, Bill Becker, and Laurits Christensen for their encouragement and guidance. After seven editions of this book, the number of individuals who have significantly improved it through their comments, criticisms, and encouragement has become far too large for me to thank each of them individually. I am grateful for their help and I hope that all of them see their contribution to this edition. I would like to acknowledge the many reviewers of my work whose careful reading has vastly improved the book through this edition: Scott Atkinson, University of Georgia; Badi Baltagi, Syracuse University; Neal Beck, New York University; William E. Becker (Ret.), Indiana University; Eric J. Belasko, Texas Tech University; Anil Bera, University of Illinois; John Burkett, University of Rhode Island; Leonard Carlson, Emory University; Frank Chaloupka, University of Illinois at Chicago; Chris Cornwell, University of Georgia; Craig Depken II, University of Texas at Arlington; Frank Diebold, University of Pennsylvania; Edward Dwyer, Clemson University; Michael Ellis, Wesleyan University; Martin Evans, Georgetown University; Vahagn Galstyan, Trinity College Dublin; Paul Glewwe, University of Minnesota; Ed Greenberg, Washington University at St. Louis; Miguel Herce, University of North Carolina; Joseph Hilbe, Arizona State University; Dr. Uwe Jensen, Christian-Albrecht University; K. Rao Kadiyala, Purdue University; William Lott, University of Connecticut; Thomas L. Marsh, Washington State University; Edward Mathis, Villanova University; Mary McGarvey, University of Nebraska–Lincoln; Ed Melnick, New York University; Thad Mirer, State University of New York at Albany; Cyril Pasche, University of Geneva; Paul Ruud, University of California at Berkeley; Sherrie Rhine, Federal Deposit Insurance Corp.; Terry G. Seaks (Ret.), 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; Dmitrios Thomakos, Columbia University; Paul Wachtel, New York University; Mary Beth Walker, Georgia State University; Mark Watson, Harvard University; and Kenneth West, University of Wisconsin. My numerous discussions with Bruce McCullough of Drexel University 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. Professor Pedro Bacao, University of Coimbra, Portugal, and Mark Strahan of Sand Hill Econometrics did likewise with the sixth edition.

I've had great support and encouragement over the years from many people close to me, especially my family, and many not so close. None has been more gratifying than the mail I've received from readers from the world over who have shared my enthusiasm for this exciting field and for this work that has taught them and me econometrics since the first edition in 1990.

Finally, I would also like to thank the many people at Prentice Hall who have put this book together with me: Adrienne D'Ambrosio, Jill Kolongowski, Carla Thompson, Nancy Fenton, Alison Eusden, Joe Vetere, and Martha Wetherill and the composition team at Macmillan Publishing Solutions.

William H. Greene
August 2010

BRIEF CONTENTS



	Examples and Applications	xxv
	Preface	xxxiii
Part I	The Linear Regression Model	
Chapter 1	Econometrics	1
Chapter 2	The Linear Regression Model	11
Chapter 3	Least Squares	26
Chapter 4	The Least Squares Estimator	51
Chapter 5	Hypothesis Tests and Model Selection	108
Chapter 6	Functional Form and Structural Change	149
Chapter 7	Nonlinear, Semiparametric, and Nonparametric Regression Models	181
Chapter 8	Endogeneity and Instrumental Variable Estimation	219
Part II	Generalized Regression Model and Equation Systems	
Chapter 9	The Generalized Regression Model and Heteroscedasticity	257
Chapter 10	Systems of Equations	290
Chapter 11	Models for Panel Data	343
Part III	Estimation Methodology	
Chapter 12	Estimation Frameworks in Econometrics	432
Chapter 13	Minimum Distance Estimation and the Generalized Method of Moments	455
Chapter 14	Maximum Likelihood Estimation	509
Chapter 15	Simulation-Based Estimation and Inference and Random Parameter Models	603
Chapter 16	Bayesian Estimation and Inference	655
Part IV	Cross Sections, Panel Data, and Microeconometrics	
Chapter 17	Discrete Choice	681
Chapter 18	Discrete Choices and Event Counts	760
Chapter 19	Limited Dependent Variables—Truncation, Censoring, and Sample Selection	833

Part V Time Series and Macroeconometrics

Chapter 20	Serial Correlation	903
Chapter 21	Nonstationary Data	942

Part VI Appendices

Appendix A	Matrix Algebra	973
Appendix B	Probability and Distribution Theory	1015
Appendix C	Estimation and Inference	1047
Appendix D	Large-Sample Distribution Theory	1066
Appendix E	Computation and Optimization	1089
Appendix F	Data Sets Used in Applications	1109
References		1115
Combined Author and Subject Index		1171

CONTENTS



Examples and Applications xxv

Preface xxxiii

PART I The Linear Regression Model

CHAPTER 1 Econometrics **1**

- 1.1 Introduction 1
- 1.2 The Paradigm of Econometrics 1
- 1.3 The Practice of Econometrics 3
- 1.4 Econometric Modeling 4
- 1.5 Plan of the Book 7
- 1.6 Preliminaries 9
 - 1.6.1 Numerical Examples* 9
 - 1.6.2 Software and Replication* 9
 - 1.6.3 Notational Conventions* 9

CHAPTER 2 The Linear Regression Model **11**

- 2.1 Introduction 11
- 2.2 The Linear Regression Model 12
- 2.3 Assumptions of the Linear Regression Model 15
 - 2.3.1 Linearity of the Regression Model* 15
 - 2.3.2 Full Rank* 19
 - 2.3.3 Regression* 20
 - 2.3.4 Spherical Disturbances* 21
 - 2.3.5 Data Generating Process for the Regressors* 23
 - 2.3.6 Normality* 23
 - 2.3.7 Independence* 24
- 2.4 Summary and Conclusions 25

CHAPTER 3 Least Squares **26**

- 3.1 Introduction 26
- 3.2 Least Squares Regression 26
 - 3.2.1 The Least Squares Coefficient Vector* 27

3.2.2	<i>Application: An Investment Equation</i>	28
3.2.3	<i>Algebraic Aspects of the Least Squares Solution</i>	30
3.2.4	<i>Projection</i>	31
3.3	Partitioned Regression and Partial Regression	32
3.4	Partial Regression and Partial Correlation Coefficients	36
3.5	Goodness of Fit and the Analysis of Variance	39
3.5.1	<i>The Adjusted R-Squared and a Measure of Fit</i>	42
3.5.2	<i>R-Squared and the Constant Term in the Model</i>	44
3.5.3	<i>Comparing Models</i>	45
3.6	Linearly Transformed Regression	46
3.7	Summary and Conclusions	47

CHAPTER 4 The Least Squares Estimator 51

4.1	Introduction	51
4.2	Motivating Least Squares	52
4.2.1	<i>The Population Orthogonality Conditions</i>	52
4.2.2	<i>Minimum Mean Squared Error Predictor</i>	53
4.2.3	<i>Minimum Variance Linear Unbiased Estimation</i>	54
4.3	Finite Sample Properties of Least Squares	54
4.3.1	<i>Unbiased Estimation</i>	55
4.3.2	<i>Bias Caused by Omission of Relevant Variables</i>	56
4.3.3	<i>Inclusion of Irrelevant Variables</i>	58
4.3.4	<i>The Variance of the Least Squares Estimator</i>	58
4.3.5	<i>The Gauss–Markov Theorem</i>	60
4.3.6	<i>The Implications of Stochastic Regressors</i>	60
4.3.7	<i>Estimating the Variance of the Least Squares Estimator</i>	61
4.3.8	<i>The Normality Assumption</i>	63
4.4	Large Sample Properties of the Least Squares Estimator	63
4.4.1	<i>Consistency of the Least Squares Estimator of β</i>	63
4.4.2	<i>Asymptotic Normality of the Least Squares Estimator</i>	65
4.4.3	<i>Consistency of s^2 and the Estimator of Asy. Var[\mathbf{b}]</i>	67
4.4.4	<i>Asymptotic Distribution of a Function of \mathbf{b}: The Delta Method</i>	68
4.4.5	<i>Asymptotic Efficiency</i>	69
4.4.6	<i>Maximum Likelihood Estimation</i>	73
4.5	Interval Estimation	75
4.5.1	<i>Forming a Confidence Interval for a Coefficient</i>	76
4.5.2	<i>Confidence Intervals Based on Large Samples</i>	78
4.5.3	<i>Confidence Interval for a Linear Combination of Coefficients: The Oaxaca Decomposition</i>	79
4.6	Prediction and Forecasting	80
4.6.1	<i>Prediction Intervals</i>	81
4.6.2	<i>Predicting y When the Regression Model Describes Log y</i>	81

4.6.3	<i>Prediction Interval for y When the Regression Model Describes Log y</i>	83
4.6.4	<i>Forecasting</i>	87
4.7	Data Problems	88
4.7.1	<i>Multicollinearity</i>	89
4.7.2	<i>Pretest Estimation</i>	91
4.7.3	<i>Principal Components</i>	92
4.7.4	<i>Missing Values and Data Imputation</i>	94
4.7.5	<i>Measurement Error</i>	97
4.7.6	<i>Outliers and Influential Observations</i>	99
4.8	Summary and Conclusions	102
CHAPTER 5 Hypothesis Tests and Model Selection		108
5.1	Introduction	108
5.2	Hypothesis Testing Methodology	108
5.2.1	<i>Restrictions and Hypotheses</i>	109
5.2.2	<i>Nested Models</i>	110
5.2.3	<i>Testing Procedures—Neyman–Pearson Methodology</i>	111
5.2.4	<i>Size, Power, and Consistency of a Test</i>	111
5.2.5	<i>A Methodological Dilemma: Bayesian versus Classical Testing</i>	112
5.3	Two Approaches to Testing Hypotheses	112
5.4	Wald Tests Based on the Distance Measure	115
5.4.1	<i>Testing a Hypothesis about a Coefficient</i>	115
5.4.2	<i>The F Statistic and the Least Squares Discrepancy</i>	117
5.5	Testing Restrictions Using the Fit of the Regression	121
5.5.1	<i>The Restricted Least Squares Estimator</i>	121
5.5.2	<i>The Loss of Fit from Restricted Least Squares</i>	122
5.5.3	<i>Testing the Significance of the Regression</i>	126
5.5.4	<i>Solving Out the Restrictions and a Caution about Using R^2</i>	126
5.6	Nonnormal Disturbances and Large-Sample Tests	127
5.7	Testing Nonlinear Restrictions	131
5.8	Choosing between Nonnested Models	134
5.8.1	<i>Testing Nonnested Hypotheses</i>	134
5.8.2	<i>An Encompassing Model</i>	135
5.8.3	<i>Comprehensive Approach—The J Test</i>	136
5.9	A Specification Test	137
5.10	Model Building—A General to Simple Strategy	138
5.10.1	<i>Model Selection Criteria</i>	139
5.10.2	<i>Model Selection</i>	140
5.10.3	<i>Classical Model Selection</i>	140
5.10.4	<i>Bayesian Model Averaging</i>	141
5.11	Summary and Conclusions	143

CHAPTER 6	Functional Form and Structural Change	149
6.1	Introduction	149
6.2	Using Binary Variables	149
6.2.1	<i>Binary Variables in Regression</i>	149
6.2.2	<i>Several Categories</i>	152
6.2.3	<i>Several Groupings</i>	152
6.2.4	<i>Threshold Effects and Categorical Variables</i>	154
6.2.5	<i>Treatment Effects and Differences in Differences Regression</i>	155
6.3	Nonlinearity in the Variables	158
6.3.1	<i>Piecewise Linear Regression</i>	158
6.3.2	<i>Functional Forms</i>	160
6.3.3	<i>Interaction Effects</i>	161
6.3.4	<i>Identifying Nonlinearity</i>	162
6.3.5	<i>Intrinsically Linear Models</i>	165
6.4	Modeling and Testing for a Structural Break	168
6.4.1	<i>Different Parameter Vectors</i>	168
6.4.2	<i>Insufficient Observations</i>	169
6.4.3	<i>Change in a Subset of Coefficients</i>	170
6.4.4	<i>Tests of Structural Break with Unequal Variances</i>	171
6.4.5	<i>Predictive Test of Model Stability</i>	174
6.5	Summary and Conclusions	175
CHAPTER 7	Nonlinear, Semiparametric, and Nonparametric Regression Models	181
7.1	Introduction	181
7.2	Nonlinear Regression Models	182
7.2.1	<i>Assumptions of the Nonlinear Regression Model</i>	182
7.2.2	<i>The Nonlinear Least Squares Estimator</i>	184
7.2.3	<i>Large Sample Properties of the Nonlinear Least Squares Estimator</i>	186
7.2.4	<i>Hypothesis Testing and Parametric Restrictions</i>	189
7.2.5	<i>Applications</i>	191
7.2.6	<i>Computing the Nonlinear Least Squares Estimator</i>	200
7.3	Median and Quantile Regression	202
7.3.1	<i>Least Absolute Deviations Estimation</i>	203
7.3.2	<i>Quantile Regression Models</i>	207
7.4	Partially Linear Regression	210
7.5	Nonparametric Regression	212
7.6	Summary and Conclusions	215
CHAPTER 8	Endogeneity and Instrumental Variable Estimation	219
8.1	Introduction	219
8.2	Assumptions of the Extended Model	223

8.3	Estimation	224
	8.3.1	<i>Least Squares</i> 225
	8.3.2	<i>The Instrumental Variables Estimator</i> 225
	8.3.3	<i>Motivating the Instrumental Variables Estimator</i> 227
	8.3.4	<i>Two-Stage Least Squares</i> 230
8.4	Two Specification Tests	233
	8.4.1	<i>The Hausman and Wu Specification Tests</i> 234
	8.4.2	<i>A Test for Overidentification</i> 238
8.5	Measurement Error	239
	8.5.1	<i>Least Squares Attenuation</i> 240
	8.5.2	<i>Instrumental Variables Estimation</i> 242
	8.5.3	<i>Proxy Variables</i> 242
8.6	Nonlinear Instrumental Variables Estimation	246
8.7	Weak Instruments	249
8.8	Natural Experiments and the Search for Causal Effects	251
8.9	Summary and Conclusions	254

PART II Generalized Regression Model and Equation Systems

CHAPTER 9 The Generalized Regression Model and Heteroscedasticity 257

9.1	Introduction	257
9.2	Inefficient Estimation by Least Squares and Instrumental Variables	258
	9.2.1	<i>Finite-Sample Properties of Ordinary Least Squares</i> 259
	9.2.2	<i>Asymptotic Properties of Ordinary Least Squares</i> 259
	9.2.3	<i>Robust Estimation of Asymptotic Covariance Matrices</i> 261
	9.2.4	<i>Instrumental Variable Estimation</i> 262
9.3	Efficient Estimation by Generalized Least Squares	264
	9.3.1	<i>Generalized Least Squares (GLS)</i> 264
	9.3.2	<i>Feasible Generalized Least Squares (FGLS)</i> 266
9.4	Heteroscedasticity and Weighted Least Squares	268
	9.4.1	<i>Ordinary Least Squares Estimation</i> 269
	9.4.2	<i>Inefficiency of Ordinary Least Squares</i> 270
	9.4.3	<i>The Estimated Covariance Matrix of \mathbf{b}</i> 270
	9.4.4	<i>Estimating the Appropriate Covariance Matrix for Ordinary Least Squares</i> 272
9.5	Testing for Heteroscedasticity	275
	9.5.1	<i>White's General Test</i> 275
	9.5.2	<i>The Breusch-Pagan/Godfrey LM Test</i> 276
9.6	Weighted Least Squares	277
	9.6.1	<i>Weighted Least Squares with Known Ω</i> 278
	9.6.2	<i>Estimation When Ω Contains Unknown Parameters</i> 279