

PRINCIPLES OF

ECONOMETRICS

5th
Edition

R. CARTER HILL

WILLIAM E. GRIFFITHS

GUAY C. LIM



WILEY

Principles of Econometrics

Fifth Edition

R. CARTER HILL

Louisiana State University

WILLIAM E. GRIFFITHS

University of Melbourne

GUAY C. LIM

University of Melbourne

WILEY

EDITORIAL DIRECTOR	Michael McDonald
EXECUTIVE EDITOR	Lise Johnson
CHANNEL MARKETING MANAGER	Michele Szczesniak
CONTENT ENABLEMENT SENIOR MANAGER	Leah Michael
CONTENT MANAGEMENT DIRECTOR	Lisa Wojcik
CONTENT MANAGER	Nichole Urban
SENIOR CONTENT SPECIALIST	Nicole Repasky
PRODUCTION EDITOR	Abidha Sulaiman
COVER PHOTO CREDIT	© liuzishan/iStockphoto

This book was set in STIX-Regular 10/12pt by SPi Global and printed and bound by Strategic Content Imaging.

This book is printed on acid free paper. ∞

Founded in 1807, John Wiley & Sons, Inc. has been a valued source of knowledge and understanding for more than 200 years, helping people around the world meet their needs and fulfill their aspirations. Our company is built on a foundation of principles that include responsibility to the communities we serve and where we live and work. In 2008, we launched a Corporate Citizenship Initiative, a global effort to address the environmental, social, economic, and ethical challenges we face in our business. Among the issues we are addressing are carbon impact, paper specifications and procurement, ethical conduct within our business and among our vendors, and community and charitable support. For more information, please visit our website: www.wiley.com/go/citizenship.

Copyright © 2018, 2011 and 2007 John Wiley & Sons, Inc. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923 (Web site: www.copyright.com). Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030-5774, (201) 748-6011, fax (201) 748-6008, or online at: www.wiley.com/go/permissions.

Evaluation copies are provided to qualified academics and professionals for review purposes only, for use in their courses during the next academic year. These copies are licensed and may not be sold or transferred to a third party. Upon completion of the review period, please return the evaluation copy to Wiley. Return instructions and a free of charge return shipping label are available at: www.wiley.com/go/returnlabel. If you have chosen to adopt this textbook for use in your course, please accept this book as your complimentary desk copy. Outside of the United States, please contact your local sales representative.

ISBN: 9781118452271 (PBK)

ISBN: 9781119320951 (EVALC)

Library of Congress Cataloging-in-Publication Data

Names: Hill, R. Carter, author. | Griffiths, William E., author. | Lim, G. C. (Guay C.), author.

Title: Principles of econometrics / R. Carter Hill, Louisiana State University, William E. Griffiths, University of Melbourne, Guay C. Lim, University of Melbourn.

Description: Fifth Edition. | Hoboken : Wiley, 2017. | Revised edition of the authors' Principles of econometrics, c2011. | Includes bibliographical references and index. |

Identifiers: LCCN 2017043094 (print) | LCCN 2017056927 (ebook) | ISBN 9781119342854 (pdf) | ISBN 9781119320944 (epub) | ISBN 9781118452271 (paperback) | ISBN 9781119320951 (eval copy)

Subjects: LCSH: Econometrics. | BISAC: BUSINESS & ECONOMICS / Econometrics.

Classification: LCC HB139 (ebook) | LCC HB139 .H548 2018 (print) | DDC 330.01/5195—dc23

LC record available at <https://lcn.loc.gov/2017043094>

The inside back cover will contain printing identification and country of origin if omitted from this page. In addition, if the ISBN on the back cover differs from the ISBN on this page, the one on the back cover is correct.

Carter Hill dedicates this work to his wife, Melissa Waters
Bill Griffiths dedicates this work to Jill, David, and Wendy Griffiths
Guay Lim dedicates this work to Tony Meagher

Brief Contents

PREFACE	v		
LIST OF EXAMPLES	xxi		
1	An Introduction to Econometrics	1	
Probability Primer	15		
2	The Simple Linear Regression Model	46	
3	Interval Estimation and Hypothesis Testing	112	
4	Prediction, Goodness-of-Fit, and Modeling Issues	152	
5	The Multiple Regression Model	196	
6	Further Inference in the Multiple Regression Model	260	
7	Using Indicator Variables	317	
8	Heteroskedasticity	368	
9	Regression with Time-Series Data: Stationary Variables	417	
10	Endogenous Regressors and Moment-Based Estimation	481	
11	Simultaneous Equations Models	531	
12	Regression with Time-Series Data: Nonstationary Variables	563	
13	Vector Error Correction and Vector Autoregressive Models	597	
14	Time-Varying Volatility and ARCH Models	614	
15	Panel Data Models	634	
16	Qualitative and Limited Dependent Variable Models	681	
APPENDIX A	Mathematical Tools	748	
APPENDIX B	Probability Concepts	768	
APPENDIX C	Review of Statistical Inference	812	
APPENDIX D	Statistical Tables	862	
INDEX		869	

Principles of Econometrics, Fifth Edition, is an introductory book for undergraduate students in economics and finance, as well as first-year graduate students in economics, finance, accounting, agricultural economics, marketing, public policy, sociology, law, forestry, and political science. We assume that students have taken courses in the principles of economics and elementary statistics. Matrix algebra is not used, and we introduce and develop calculus concepts in an Appendix. The title *Principles of Econometrics* emphasizes our belief that econometrics should be part of the economics curriculum, in the same way as the principles of microeconomics and the principles of macroeconomics. Those who have been studying and teaching econometrics as long as we have will remember that *Principles of Econometrics* was the title that Henri Theil used for his 1971 classic, which was also published by John Wiley & Sons. Our choice of the same title is not intended to signal that our book is similar in level and content. Theil's work was, and remains, a unique treatise on advanced graduate level econometrics. Our book is an introductory level econometrics text.

Book Objectives

Principles of Econometrics is designed to give students an understanding of why econometrics is necessary and to provide them with a working knowledge of basic econometric tools so that

- i. They can apply these tools to modeling, estimation, inference, and forecasting in the context of real-world economic problems.
- ii. They can evaluate critically the results and conclusions from others who use basic econometric tools.
- iii. They have a foundation and understanding for further study of econometrics.
- iv. They have an appreciation of the range of more advanced techniques that exist and that may be covered in later econometric courses.

The book is neither an econometrics cookbook nor is it in a theorem-proof format. It emphasizes motivation, understanding, and implementation. Motivation is achieved by introducing very simple economic models and asking economic questions that the student can answer. Understanding is aided by lucid description of techniques, clear interpretation, and appropriate applications. Learning is reinforced by doing, with clear worked examples in the text and exercises at the end of each chapter.

Overview of Contents

This fifth edition is a major revision in format and content. The chapters contain core material and exercises, while appendices contain more advanced material. Chapter examples are now identified and separated from other content so that they may be easily referenced. From the beginning, we recognize the observational nature of most economic data and modify modeling assumptions accordingly. Chapter 1 introduces econometrics and gives general guidelines for writing an empirical research paper and locating economic data sources. The Probability Primer preceding Chapter 2 summarizes essential properties of random variables and their probability distributions and reviews summation notation. The simple linear regression model is covered in Chapters 2–4, while the multiple regression model is treated in Chapters 5–7. Chapters 8 and 9 introduce econometric problems that are unique to cross-sectional data (heteroskedasticity) and time-series data

(dynamic models), respectively. Chapters 10 and 11 deal with endogenous regressors, the failure of least squares when a regressor is endogenous, and instrumental variables estimation, first in the general case, and then in the simultaneous equations model. In Chapter 12, the analysis of time-series data is extended to discussions of nonstationarity and cointegration. Chapter 13 introduces econometric issues specific to two special time-series models, the vector error correction, and vector autoregressive models, while Chapter 14 considers the analysis of volatility in data and the ARCH model. In Chapters 15 and 16, we introduce microeconomic models for panel data and qualitative and limited dependent variables. In appendices A, B, and C, we introduce math, probability, and statistical inference concepts that are used in the book.

Summary of Changes and New Material

This edition includes a great deal of new material, including new examples and exercises using real data and some significant reorganizations. In this edition, we number examples for easy reference and offer 25–30 new exercises in each chapter. Important new features include

- Chapter 1 includes a discussion of data types and sources of economic data on the Internet. Tips on writing a research paper are given “up front” so that students can form ideas for a paper as the course develops.
- A Probability Primer precedes Chapter 2. This Primer reviews the concepts of random variables and how probabilities are calculated from probability density functions. Mathematical expectation and rules of expected values are summarized for discrete random variables. These rules are applied to develop the concepts of variance and covariance. Calculations of probabilities using the normal distribution are illustrated. New material includes sections on conditional expectation, conditional variance, iterated expectations, and the bivariate normal distribution.
- Chapter 2 now starts with a discussion of causality. We define the population regression function and discuss exogeneity in considerable detail. The properties of the ordinary least squares (OLS) estimator are examined within the framework of the new assumptions. New appendices have been added on the independent variable, covering the various assumptions that might be made about the sampling process, derivations of the properties of the OLS estimator, and Monte Carlo experiments to numerically illustrate estimator properties.
- In Chapter 3, we note that hypothesis test mechanics remain the same under the revised assumptions because test statistics are “pivotal,” meaning that their distributions under the null hypothesis do not depend on the data. In appendices, we add an extended discussion of test behavior under the alternative, introduce the noncentral t -distribution, and illustrate test power. We also include new Monte Carlo experiments illustrating test properties when the explanatory variable is random.
- Chapter 4 discusses in detail nonlinear relationships such as the log-log, log-linear, linear-log, and polynomial models. We have expanded the discussion of diagnostic residual plots and added sections on identifying influential observations. The familiar concepts of compound interest are used to motivate several log-linear models. We add an appendix on the concept of mean squared error and the minimum mean squared error predictor.
- Chapter 5 introduces multiple regression in the random- x framework. The Frisch–Waugh–Lovell (FWL) theorem is introduced as a way to help understand interpretation of the multiple regression model and used throughout the remainder of the book. Discussions of the properties of the OLS estimator, and interval estimates and t -tests, are updated. The large sample properties of the OLS estimator, and the delta method, are now introduced within the chapter rather than an appendix. Appendices provide further discussion and Monte Carlo

properties to illustrate the delta method. We provide a new appendix on bootstrapping and its uses.

- Chapter 6 adds a new section on large sample tests. We explain the use of control variables and the difference between causal and predictive models. We revise the discussion of collinearity and include a discussion of influential observations. We introduce nonlinear regression models and nonlinear least squares algorithms are discussed. Appendices are added to discuss the statistical power of F -tests and further uses of the Frisch–Waugh–Lovell theorem.
- Chapter 7 now includes an extensive section on treatment effects and causal modeling in Rubin’s potential outcomes framework. We explain and illustrate the interesting regression discontinuity design. An appendix includes a discussion of the important “overlap” assumption.
- Chapter 8 has been reorganized so that the heteroskedasticity robust variance of the OLS estimator appears before testing. We add a section on how model specification can ameliorate heteroskedasticity in some applications. We add appendices to explain the properties of the OLS residuals and another to explain alternative robust sandwich variance estimators. We present Monte Carlo experiments to illustrate the differences.
- Chapter 9 has been reorganized and streamlined. The initial section introduces the different ways that dynamic elements can be added to the regression model. These include using finite lag models, infinite lag models, and autoregressive errors. We carefully discuss autocorrelations, including testing for autocorrelation and representing autocorrelations using a correlogram. After introducing the concepts of stationarity and weak dependence, we discuss the general notions of forecasting and forecast intervals in the context of autoregressive distributed lag (ARDL) models. Following these introductory concepts, there are details of estimating and using alternative models, covering such topics as choosing lag lengths, testing for Granger causality, the Lagrange multiplier test for serial correlation, and using models for policy analysis. We provide very specific sets of assumptions for time-series regression models and outline how heteroskedastic and autocorrelation consistent, robust, standard errors are used. We discuss generalized least squares estimation of a time-series regression model and its relation to nonlinear least squares regression. A detailed discussion of the infinite lag model and how to use multiplier analysis is provided. An appendix contains details of the Durbin–Watson test.
- Chapter 10 on endogeneity problems has been streamlined because the concept of random explanatory variables is now introduced much earlier in the book. We provide further analysis of weak instruments and how weak instruments adversely affect the precision of IV estimation. The details of the Hausman test are now included in the chapter.
- Chapter 11 now adds Klein’s Model I as an example.
- Chapter 12 includes more details of deterministic trends and unit roots. The section on unit root testing has been restructured so that each Dickey–Fuller test is more fully explained and illustrated with an example. Numerical examples of ARDL models with nonstationary variables that are, and are not, cointegrated have been added.
- The data in Chapter 13 have been updated and new exercises added.
- Chapter 14 mentions further extensions of ARCH volatility models.
- Chapter 15 has been restructured to give priority to how panel data can be used to cope with the endogeneity caused by unobserved heterogeneity. We introduce the advantages of having panel data using the first difference estimator, and then discuss the within/fixed effects estimator. We provide an extended discussion of cluster robust standard errors in both the OLS and fixed effects model. We discuss the Mundlak version of the Hausman test for endogeneity. We give brief mention to how to extend the use of panel data in several ways.

- The Chapter 16 discussion of binary choice models is reorganized and expanded. It now includes brief discussions of advanced topics such as binary choice models with endogenous explanatory variables and binary choice models with panel data. We add new appendices on random utility models and latent variable models.
- Appendix A includes new sections on second derivatives and finding maxima and minima of univariate and bivariate functions.
- Appendix B includes new material on conditional expectations and conditional variances, including several useful decompositions. We include new sections on truncated random variables, including the truncated normal and Poisson distributions. To facilitate discussions of test power, we have new sections on the noncentral t -distribution, the noncentral Chi-square distribution, and the noncentral F -distribution. We have included an expanded new section on the log-normal distribution.
- Appendix C content does not change a great deal, but 20 new exercises are included.
- Statistical Tables for the Standard Normal cumulative distribution function, the t -distribution and Chi-square distribution critical values for selected percentiles, the F -distribution critical values for the 95th and 99th percentiles, and the Standard Normal density function values appear in Appendix D.
- A useful “cheat sheet” of essential formulas is provided at the authors’ website, www.principlesofeconometrics.com, rather than inside the covers as in the previous edition.

For Instructors: Suggested Course Plans

Principles of Econometrics, Fifth Edition is suitable for one or two semester courses at the undergraduate or first year graduate level. Some suitable plans for alternative courses are as follows:

- One-semester survey course: Sections P.1–P.6.2 and P.7; Sections 2.1–2.9; Chapters 3 and 4; Sections 5.1–5.6; Sections 6.1–6.5; Sections 7.1–7.3; Sections 8.1–8.4 and 8.6; Sections 9.1–9.4.2 and 9.5–9.5.1.
- One-semester survey course enhancements for Master’s or Ph.D.: Include Appendices for Chapters 2–9.
- Two-semester survey second course, cross-section emphasis: Section P.6; Section 2.10; Section 5.7; Section 6.6; Sections 7.4–7.6; Sections 8.5 and 8.6.3–8.6.5; Sections 10.1–10.4; Sections 15.1–15.4; Sections 16.1–16.2 and 16.6;
- Two-semester survey second course, time series emphasis: Section P.6; Section 2.10; Section 5.7; Section 6.6; Sections 7.4–7.6; Sections 8.5 and 8.6.3–8.6.5; Section 9.5; Sections 10.1–10.4; Sections 12.1–12.5; Sections 13.1–13.5; Sections 14.1–14.4;
- Two-semester survey course enhancements for Master’s or Ph.D.: Include Appendices from Chapters 10, Chapter 11, Appendices 15A–15B, Sections 16.3–16.5 and 16.7, Appendices 16A–16D, Book Appendices B and C.

Computer Supplement Books

There are several computer supplements to *Principles of Econometrics, Fifth Edition*. The supplements are not versions of the text and cannot substitute for the text. They use the examples in the text as a vehicle for learning the software. We show how to use the software to get the answers for each example in the text.

- *Using EViews for the Principles of Econometrics, Fifth Edition*, by William E. Griffiths, R. Carter Hill, and Guay C. Lim [ISBN 9781118469842]. This supplementary book presents the EViews 10 [www.eviews.com] software commands required for the examples in *Principles of Econometrics* in a clear and concise way. It includes many illustrations that are student friendly. It is useful not only for students and instructors who will be using this software as part of their econometrics course but also for those who wish to learn how to use EViews.
- *Using Stata for the Principles of Econometrics, Fifth Edition*, by Lee C. Adkins and R. Carter Hill [ISBN 9781118469873]. This supplementary book presents the Stata 15.0 [www.stata.com] software commands required for the examples in *Principles of Econometrics*. It is useful not only for students and instructors who will be using this software as part of their econometrics course but also for those who wish to learn how to use Stata. Screen shots illustrate the use of Stata's drop-down menus. Stata commands are explained and the use of "do-files" illustrated.
- *Using SAS for the Principles of Econometrics, Fifth Edition*, by Randall C. Campbell and R. Carter Hill [ISBN 9781118469880]. This supplementary book gives SAS 9.4 [www.sas.com] software commands for econometric tasks, following the general outline of *Principles of Econometrics, Fifth Edition*. It includes enough background material on econometrics so that instructors using any textbook can easily use this book as a supplement. The volume spans several levels of econometrics. It is suitable for undergraduate students who will use "canned" SAS statistical procedures, and for graduate students who will use advanced procedures as well as direct programming in SAS's matrix language; the latter is discussed in chapter appendices.
- *Using Excel for Principles of Econometrics, Fifth Edition*, by Genevieve Briand and R. Carter Hill [ISBN 9781118469835]. This supplement explains how to use Excel to reproduce most of the examples in *Principles of Econometrics*. Detailed instructions and screen shots are provided explaining both the computations and clarifying the operations of Excel. Templates are developed for common tasks.
- *Using GRETL for Principles of Econometrics, Fifth Edition*, by Lee C. Adkins. This free supplement, readable using Adobe Acrobat, explains how to use the freely available statistical software GRETL (download from <http://gretl.sourceforge.net>). Professor Adkins explains in detail, and using screen shots, how to use GRETL to replicate the examples in *Principles of Econometrics*. The manual is freely available at www.learneconometrics.com/gretl.html.
- *Using R for Principles of Econometrics, Fifth Edition*, by Constantin Colonescu and R. Carter Hill. This free supplement, readable using Adobe Acrobat, explains how to use the freely available statistical software *R* (download from <https://www.r-project.org/>). The supplement explains in detail, and using screen shots, how to use *R* to replicate the examples in *Principles of Econometrics, Fifth Edition*. The manual is freely available at <https://bookdown.org/ccolonescu/RPOE5/>.

Data Files

Data files for the book are provided in a variety of formats at the book website www.wiley.com/college/hill. These include

- ASCII format (*.dat). These are text files containing only data.
- Definition files (*.def). These are text files describing the data file contents, with a listing of variable names, variable definitions, and summary statistics.
- EViews (*.wf1) workfiles for each data file.
- Excel (*.xls) workbooks for each data file, including variable names in the first row.

- Comma separated values (*.csv) files that can be read into almost all software.
- Stata (*.dta) data files.
- SAS (*.sas7bdat) data files.
- GRETTL (*.gdt) data files.
- R (*.rdata) data files.

The author website www.principlesofeconometrics.com includes a complete list of the data files and where they are used in the book.

Additional Resources

The book website www.principlesofeconometrics.com includes

- Individual data files in each format as well as ZIP files containing data in compressed format.
- Book errata.
- Brief answers to odd number problems. These answers are also provided on the book website at www.wiley.com/college/hill.
- Additional examples with solutions. Some extra examples come with complete solutions so that students will know what a good answer looks like.
- Tips on writing research papers.

Resources for Instructors

For instructors, also available at the website www.wiley.com/college/hill are

- Complete solutions, in both Microsoft Word and *.pdf formats, to *all* exercises in the text.
- PowerPoint slides and PowerPoint Viewer.

Acknowledgments

Authors Hill and Griffiths want to acknowledge the gifts given to them over the past 40 years by mentor, friend, and colleague George Judge. Neither this book nor any of the other books we have shared in the writing of would have ever seen the light of day without his vision and inspiration.

We also wish to give thanks to the many students and faculty who have commented on the fourth edition of the text and contributed to the fifth edition. This list includes Alejandra Breve Ferrari, Alex James, Alyssa Wans, August Saibeni, Barry Rafferty, Bill Rising, Bob Martin, Brad Lewis, Bronson Fong, David Harris, David Iseral, Deborah Williams, Deokrye Baek, Diana Whistler, Emma Powers, Ercan Saridogan, Erdogan Cevher, Erika Haguette, Ethan Luedecke, Gareth Thomas, Gawon Yoon, Genevieve Briand, German Altgelt, Glenn Sueyoshi, Henry McCool, James Railton, Jana Ruimerman, Jeffery Parker, Joe Goss, John Jackson, Julie Leiby, Katharina Hauck, Katherine Ramirez, Kelley Pace, Lee Adkins, Matias Cattaneo, Max O’Krepki, Meagan McCollum, Micah West, Michelle Savolainen, Oystein Myrland, Patrick Scholten, Randy Campbell, Regina Riphahn, Sandamali Kankanamge, Sergio Pastorello, Shahrokh Towfighi, Tom Fomby, Tong Zeng, Victoria Pryor, Yann Nicolas, and Yuanbo Zhang. In the book errata we acknowledge those who have pointed out our errors.

R. Carter Hill
William E. Griffiths
Guay C. Lim

Table of Contents

PREFACE	v
LIST OF EXAMPLES	xxi

1	An Introduction to Econometrics	1
1.1	Why Study Econometrics?	1
1.2	What Is Econometrics About?	2
1.2.1	Some Examples	3
1.3	The Econometric Model	4
1.3.1	Causality and Prediction	5
1.4	How Are Data Generated?	5
1.4.1	Experimental Data	6
1.4.2	Quasi-Experimental Data	6
1.4.3	Nonexperimental Data	7
1.5	Economic Data Types	7
1.5.1	Time-Series Data	7
1.5.2	Cross-Section Data	8
1.5.3	Panel or Longitudinal Data	9
1.6	The Research Process	9
1.7	Writing an Empirical Research Paper	11
1.7.1	Writing a Research Proposal	11
1.7.2	A Format for Writing a Research Report	11
1.8	Sources of Economic Data	13
1.8.1	Links to Economic Data on the Internet	13
1.8.2	Interpreting Economic Data	14
1.8.3	Obtaining the Data	14
Probability Primer 15		
P.1	Random Variables	16
P.2	Probability Distributions	17
P.3	Joint, Marginal, and Conditional Probabilities	20
P.3.1	Marginal Distributions	20
P.3.2	Conditional Probability	21
P.3.3	Statistical Independence	21
P.4	A Digression: Summation Notation	22
P.5	Properties of Probability Distributions	23
P.5.1	Expected Value of a Random Variable	24
P.5.2	Conditional Expectation	25
P.5.3	Rules for Expected Values	25
P.5.4	Variance of a Random Variable	26
P.5.5	Expected Values of Several Random Variables	27

P.5.6	Covariance Between Two Random Variables	27
P.6	Conditioning	29
P.6.1	Conditional Expectation	30
P.6.2	Conditional Variance	31
P.6.3	Iterated Expectations	32
P.6.4	Variance Decomposition	33
P.6.5	Covariance Decomposition	34
P.7	The Normal Distribution	34
P.7.1	The Bivariate Normal Distribution	37
P.8	Exercises	39

2 The Simple Linear Regression Model 46

2.1	An Economic Model	47
2.2	An Econometric Model	49
2.2.1	Data Generating Process	51
2.2.2	The Random Error and Strict Exogeneity	52
2.2.3	The Regression Function	53
2.2.4	Random Error Variation	54
2.2.5	Variation in x	56
2.2.6	Error Normality	56
2.2.7	Generalizing the Exogeneity Assumption	56
2.2.8	Error Correlation	57
2.2.9	Summarizing the Assumptions	58
2.3	Estimating the Regression Parameters	59
2.3.1	The Least Squares Principle	61
2.3.2	Other Economic Models	65
2.4	Assessing the Least Squares Estimators	66
2.4.1	The Estimator b_2	67
2.4.2	The Expected Values of b_1 and b_2	68
2.4.3	Sampling Variation	69
2.4.4	The Variances and Covariance of b_1 and b_2	69
2.5	The Gauss–Markov Theorem	72
2.6	The Probability Distributions of the Least Squares Estimators	73
2.7	Estimating the Variance of the Error Term	74
2.7.1	Estimating the Variances and Covariance of the Least Squares Estimators	74
2.7.2	Interpreting the Standard Errors	76

2.8	Estimating Nonlinear Relationships	77	3.2	Hypothesis Tests	118
2.8.1	Quadratic Functions	77	3.2.1	The Null Hypothesis	118
2.8.2	Using a Quadratic Model	77	3.2.2	The Alternative Hypothesis	118
2.8.3	A Log-Linear Function	79	3.2.3	The Test Statistic	119
2.8.4	Using a Log-Linear Model	80	3.2.4	The Rejection Region	119
2.8.5	Choosing a Functional Form	82	3.2.5	A Conclusion	120
2.9	Regression with Indicator Variables	82	3.3	Rejection Regions for Specific Alternatives	120
2.10	The Independent Variable	84	3.3.1	One-Tail Tests with Alternative “Greater Than” ($>$)	120
2.10.1	Random and Independent x	84	3.3.2	One-Tail Tests with Alternative “Less Than” ($<$)	121
2.10.2	Random and Strictly Exogenous x	86	3.3.3	Two-Tail Tests with Alternative “Not Equal To” (\neq)	122
2.10.3	Random Sampling	87	3.4	Examples of Hypothesis Tests	123
2.11	Exercises	89	3.5	The p-Value	126
2.11.1	Problems	89	3.6	Linear Combinations of Parameters	129
2.11.2	Computer Exercises	93	3.6.1	Testing a Linear Combination of Parameters	131
Appendix 2A	Derivation of the Least Squares Estimates	98	3.7	Exercises	133
Appendix 2B	Deviation from the Mean Form of b_2	99	3.7.1	Problems	133
Appendix 2C	b_2 Is a Linear Estimator	100	3.7.2	Computer Exercises	139
Appendix 2D	Derivation of Theoretical Expression for b_2	100	Appendix 3A	Derivation of the t-Distribution	144
Appendix 2E	Deriving the Conditional Variance of b_2	100	Appendix 3B	Distribution of the t-Statistic under H_1	145
Appendix 2F	Proof of the Gauss–Markov Theorem	102	Appendix 3C	Monte Carlo Simulation	147
Appendix 2G	Proofs of Results Introduced in Section 2.10	103	3C.1	Sampling Properties of Interval Estimators	148
2G.1	The Implications of Strict Exogeneity	103	3C.2	Sampling Properties of Hypothesis Tests	149
2G.2	The Random and Independent x Case	103	3C.3	Choosing the Number of Monte Carlo Samples	149
2G.3	The Random and Strictly Exogenous x Case	105	3C.4	Random- x Monte Carlo Results	150
2G.4	Random Sampling	106	4	Prediction, Goodness-of-Fit, and Modeling Issues	152
Appendix 2H	Monte Carlo Simulation	106	4.1	Least Squares Prediction	153
2H.1	The Regression Function	106	4.2	Measuring Goodness-of-Fit	156
2H.2	The Random Error	107	4.2.1	Correlation Analysis	158
2H.3	Theoretically True Values	107	4.2.2	Correlation Analysis and R^2	158
2H.4	Creating a Sample of Data	108	4.3	Modeling Issues	160
2H.5	Monte Carlo Objectives	109	4.3.1	The Effects of Scaling the Data	160
2H.6	Monte Carlo Results	109	4.3.2	Choosing a Functional Form	161
2H.7	Random- x Monte Carlo Results	110	4.3.3	A Linear-Log Food Expenditure Model	163
3	Interval Estimation and Hypothesis Testing	112	4.3.4	Using Diagnostic Residual Plots	165
3.1	Interval Estimation	113	4.3.5	Are the Regression Errors Normally Distributed?	167
3.1.1	The t -Distribution	113			
3.1.2	Obtaining Interval Estimates	115			
3.1.3	The Sampling Context	116			

- 4.3.6 Identifying Influential Observations 169
- 4.4 Polynomial Models 171**
 - 4.4.1 Quadratic and Cubic Equations 171
- 4.5 Log-Linear Models 173**
 - 4.5.1 Prediction in the Log-Linear Model 175
 - 4.5.2 A Generalized R^2 Measure 176
 - 4.5.3 Prediction Intervals in the Log-Linear Model 177
- 4.6 Log-Log Models 177**
- 4.7 Exercises 179**
 - 4.7.1 Problems 179
 - 4.7.2 Computer Exercises 185
- Appendix 4A Development of a Prediction Interval 192**
- Appendix 4B The Sum of Squares Decomposition 193**
- Appendix 4C Mean Squared Error: Estimation and Prediction 193**

5 The Multiple Regression Model 196

- 5.1 Introduction 197**
 - 5.1.1 The Economic Model 197
 - 5.1.2 The Econometric Model 198
 - 5.1.3 The General Model 202
 - 5.1.4 Assumptions of the Multiple Regression Model 203
- 5.2 Estimating the Parameters of the Multiple Regression Model 205**
 - 5.2.1 Least Squares Estimation Procedure 205
 - 5.2.2 Estimating the Error Variance σ^2 207
 - 5.2.3 Measuring Goodness-of-Fit 208
 - 5.2.4 Frisch–Waugh–Lovell (FWL) Theorem 209
- 5.3 Finite Sample Properties of the Least Squares Estimator 211**
 - 5.3.1 The Variances and Covariances of the Least Squares Estimators 212
 - 5.3.2 The Distribution of the Least Squares Estimators 214
- 5.4 Interval Estimation 216**
 - 5.4.1 Interval Estimation for a Single Coefficient 216
 - 5.4.2 Interval Estimation for a Linear Combination of Coefficients 217
- 5.5 Hypothesis Testing 218**
 - 5.5.1 Testing the Significance of a Single Coefficient 219

- 5.5.2 One-Tail Hypothesis Testing for a Single Coefficient 220
- 5.5.3 Hypothesis Testing for a Linear Combination of Coefficients 221
- 5.6 Nonlinear Relationships 222**
- 5.7 Large Sample Properties of the Least Squares Estimator 227**
 - 5.7.1 Consistency 227
 - 5.7.2 Asymptotic Normality 229
 - 5.7.3 Relaxing Assumptions 230
 - 5.7.4 Inference for a Nonlinear Function of Coefficients 232
- 5.8 Exercises 234**
 - 5.8.1 Problems 234
 - 5.8.2 Computer Exercises 240
- Appendix 5A Derivation of Least Squares Estimators 247**
- Appendix 5B The Delta Method 248**
 - 5B.1 Nonlinear Function of a Single Parameter 248
 - 5B.2 Nonlinear Function of Two Parameters 249
- Appendix 5C Monte Carlo Simulation 250**
 - 5C.1 Least Squares Estimation with Chi-Square Errors 250
 - 5C.2 Monte Carlo Simulation of the Delta Method 252
- Appendix 5D Bootstrapping 254**
 - 5D.1 Resampling 255
 - 5D.2 Bootstrap Bias Estimate 256
 - 5D.3 Bootstrap Standard Error 256
 - 5D.4 Bootstrap Percentile Interval Estimate 257
 - 5D.5 Asymptotic Refinement 258

6 Further Inference in the Multiple Regression Model 260

- 6.1 Testing Joint Hypotheses: The F -test 261**
 - 6.1.1 Testing the Significance of the Model 264
 - 6.1.2 The Relationship Between t - and F -Tests 265
 - 6.1.3 More General F -Tests 267
 - 6.1.4 Using Computer Software 268
 - 6.1.5 Large Sample Tests 269
- 6.2 The Use of Nonsample Information 271**
- 6.3 Model Specification 273**
 - 6.3.1 Causality versus Prediction 273
 - 6.3.2 Omitted Variables 275
 - 6.3.3 Irrelevant Variables 277

6.3.4	Control Variables	278
6.3.5	Choosing a Model	280
6.3.6	RESET	281
6.4	Prediction	282
6.4.1	Predictive Model Selection Criteria	285
6.5	Poor Data, Collinearity, and Insignificance	288
6.5.1	The Consequences of Collinearity	289
6.5.2	Identifying and Mitigating Collinearity	290
6.5.3	Investigating Influential Observations	293
6.6	Nonlinear Least Squares	294
6.7	Exercises	297
6.7.1	Problems	297
6.7.2	Computer Exercises	303
Appendix 6A	The Statistical Power of F-Tests	311
Appendix 6B	Further Results from the FWL Theorem	315

7 Using Indicator Variables 317

7.1	Indicator Variables	318
7.1.1	Intercept Indicator Variables	318
7.1.2	Slope-Indicator Variables	320
7.2	Applying Indicator Variables	323
7.2.1	Interactions Between Qualitative Factors	323
7.2.2	Qualitative Factors with Several Categories	324
7.2.3	Testing the Equivalence of Two Regressions	326
7.2.4	Controlling for Time	328
7.3	Log-Linear Models	329
7.3.1	A Rough Calculation	330
7.3.2	An Exact Calculation	330
7.4	The Linear Probability Model	331
7.5	Treatment Effects	332
7.5.1	The Difference Estimator	334
7.5.2	Analysis of the Difference Estimator	334
7.5.3	The Differences-in-Differences Estimator	338
7.6	Treatment Effects and Causal Modeling	342
7.6.1	The Nature of Causal Effects	342
7.6.2	Treatment Effect Models	343
7.6.3	Decomposing the Treatment Effect	344
7.6.4	Introducing Control Variables	345
7.6.5	The Overlap Assumption	347
7.6.6	Regression Discontinuity Designs	347

7.7	Exercises	351
7.7.1	Problems	351
7.7.2	Computer Exercises	358

Appendix 7A	Details of Log-Linear Model Interpretation	366
--------------------	---	------------

Appendix 7B	Derivation of the Differences-in-Differences Estimator	366
--------------------	---	------------

Appendix 7C	The Overlap Assumption: Details	367
--------------------	--	------------

8 Heteroskedasticity 368

8.1	The Nature of Heteroskedasticity	369
8.2	Heteroskedasticity in the Multiple Regression Model	370
8.2.1	The Heteroskedastic Regression Model	371
8.2.2	Heteroskedasticity Consequences for the OLS Estimator	373
8.3	Heteroskedasticity Robust Variance Estimator	374
8.4	Generalized Least Squares: Known Form of Variance	375
8.4.1	Transforming the Model: Proportional Heteroskedasticity	375
8.4.2	Weighted Least Squares: Proportional Heteroskedasticity	377
8.5	Generalized Least Squares: Unknown Form of Variance	379
8.5.1	Estimating the Multiplicative Model	381
8.6	Detecting Heteroskedasticity	383
8.6.1	Residual Plots	384
8.6.2	The Goldfeld-Quandt Test	384
8.6.3	A General Test for Conditional Heteroskedasticity	385
8.6.4	The White Test	387
8.6.5	Model Specification and Heteroskedasticity	388
8.7	Heteroskedasticity in the Linear Probability Model	390
8.8	Exercises	391
8.8.1	Problems	391
8.8.2	Computer Exercises	401
Appendix 8A	Properties of the Least Squares Estimator	407
Appendix 8B	Lagrange Multiplier Tests for Heteroskedasticity	408

Appendix 8C	Properties of the Least Squares Residuals	410
8C.1	Details of Multiplicative Heteroskedasticity Model	411
Appendix 8D	Alternative Robust Sandwich Estimators	411
Appendix 8E	Monte Carlo Evidence: OLS, GLS, and FGLS	414

9 Regression with Time-Series Data: Stationary Variables 417

9.1	Introduction	418
9.1.1	Modeling Dynamic Relationships	420
9.1.2	Autocorrelations	424
9.2	Stationarity and Weak Dependence	427
9.3	Forecasting	430
9.3.1	Forecast Intervals and Standard Errors	433
9.3.2	Assumptions for Forecasting	435
9.3.3	Selecting Lag Lengths	436
9.3.4	Testing for Granger Causality	437
9.4	Testing for Serially Correlated Errors	438
9.4.1	Checking the Correlogram of the Least Squares Residuals	439
9.4.2	Lagrange Multiplier Test	440
9.4.3	Durbin–Watson Test	443
9.5	Time-Series Regressions for Policy Analysis	443
9.5.1	Finite Distributed Lags	445
9.5.2	HAC Standard Errors	448
9.5.3	Estimation with AR(1) Errors	452
9.5.4	Infinite Distributed Lags	456
9.6	Exercises	463
9.6.1	Problems	463
9.6.2	Computer Exercises	468

Appendix 9A	The Durbin–Watson Test	476
9A.1	The Durbin–Watson Bounds Test	478

Appendix 9B	Properties of an AR(1) Error	479
--------------------	-------------------------------------	------------

10 Endogenous Regressors and Moment-Based Estimation 481

10.1	Least Squares Estimation with Endogenous Regressors	482
10.1.1	Large Sample Properties of the OLS Estimator	483

10.1.2	Why Least Squares Estimation Fails	484
10.1.3	Proving the Inconsistency of OLS	486

10.2 Cases in Which x and e are Contemporaneously Correlated 487

10.2.1	Measurement Error	487
10.2.2	Simultaneous Equations Bias	488
10.2.3	Lagged-Dependent Variable Models with Serial Correlation	489
10.2.4	Omitted Variables	489

10.3 Estimators Based on the Method of Moments 490

10.3.1	Method of Moments Estimation of a Population Mean and Variance	490
10.3.2	Method of Moments Estimation in the Simple Regression Model	491
10.3.3	Instrumental Variables Estimation in the Simple Regression Model	492
10.3.4	The Importance of Using Strong Instruments	493
10.3.5	Proving the Consistency of the IV Estimator	494
10.3.6	IV Estimation Using Two-Stage Least Squares (2SLS)	495
10.3.7	Using Surplus Moment Conditions	496
10.3.8	Instrumental Variables Estimation in the Multiple Regression Model	498
10.3.9	Assessing Instrument Strength Using the First-Stage Model	500
10.3.10	Instrumental Variables Estimation in a General Model	502
10.3.11	Additional Issues When Using IV Estimation	504

10.4 Specification Tests 505

10.4.1	The Hausman Test for Endogeneity	505
10.4.2	The Logic of the Hausman Test	507
10.4.3	Testing Instrument Validity	508

10.5 Exercises 510

10.5.1	Problems	510
10.5.2	Computer Exercises	516

Appendix 10A Testing for Weak Instruments 520

10A.1	A Test for Weak Identification	521
10A.2	Testing for Weak Identification: Conclusions	525

Appendix 10B Monte Carlo Simulation 525

10B.1	Illustrations Using Simulated Data	526
10B.2	The Sampling Properties of IV/2SLS	528

11 Simultaneous Equations Models 531

- 11.1 A Supply and Demand Model 532
- 11.2 The Reduced-Form Equations 534
- 11.3 The Failure of Least Squares Estimation 535
 - 11.3.1 Proving the Failure of OLS 535
- 11.4 The Identification Problem 536
- 11.5 Two-Stage Least Squares Estimation 538
 - 11.5.1 The General Two-Stage Least Squares Estimation Procedure 539
 - 11.5.2 The Properties of the Two-Stage Least Squares Estimator 540
- 11.6 Exercises 545
 - 11.6.1 Problems 545
 - 11.6.2 Computer Exercises 551
- Appendix 11A 2SLS Alternatives 557
 - 11A.1 The k -Class of Estimators 557
 - 11A.2 The LIML Estimator 558
 - 11A.3 Monte Carlo Simulation Results 562

12 Regression with Time-Series Data: Nonstationary Variables 563

- 12.1 Stationary and Nonstationary Variables 564
 - 12.1.1 Trend Stationary Variables 567
 - 12.1.2 The First-Order Autoregressive Model 570
 - 12.1.3 Random Walk Models 572
- 12.2 Consequences of Stochastic Trends 574
- 12.3 Unit Root Tests for Stationarity 576
 - 12.3.1 Unit Roots 576
 - 12.3.2 Dickey–Fuller Tests 577
 - 12.3.3 Dickey–Fuller Test with Intercept and No Trend 577
 - 12.3.4 Dickey–Fuller Test with Intercept and Trend 579
 - 12.3.5 Dickey–Fuller Test with No Intercept and No Trend 580
 - 12.3.6 Order of Integration 581
 - 12.3.7 Other Unit Root Tests 582
- 12.4 Cointegration 582
 - 12.4.1 The Error Correction Model 584
- 12.5 Regression When There Is No Cointegration 585
- 12.6 Summary 587

- 12.7 Exercises 588
 - 12.7.1 Problems 588
 - 12.7.2 Computer Exercises 592

13 Vector Error Correction and Vector Autoregressive Models 597

- 13.1 VEC and VAR Models 598
- 13.2 Estimating a Vector Error Correction Model 600
- 13.3 Estimating a VAR Model 601
- 13.4 Impulse Responses and Variance Decompositions 603
 - 13.4.1 Impulse Response Functions 603
 - 13.4.2 Forecast Error Variance Decompositions 605
- 13.5 Exercises 607
 - 13.5.1 Problems 607
 - 13.5.2 Computer Exercises 608

Appendix 13A The Identification Problem 612

14 Time-Varying Volatility and ARCH Models 614

- 14.1 The ARCH Model 615
- 14.2 Time-Varying Volatility 616
- 14.3 Testing, Estimating, and Forecasting 620
- 14.4 Extensions 622
 - 14.4.1 The GARCH Model—Generalized ARCH 622
 - 14.4.2 Allowing for an Asymmetric Effect 623
 - 14.4.3 GARCH-in-Mean and Time-Varying Risk Premium 624
 - 14.4.4 Other Developments 625
- 14.5 Exercises 626
 - 14.5.1 Problems 626
 - 14.5.2 Computer Exercises 627

15 Panel Data Models 634

- 15.1 The Panel Data Regression Function 636
 - 15.1.1 Further Discussion of Unobserved Heterogeneity 638
 - 15.1.2 The Panel Data Regression Exogeneity Assumption 639

15.1.3	Using OLS to Estimate the Panel Data Regression	639
15.2	The Fixed Effects Estimator	640
15.2.1	The Difference Estimator: $T = 2$	640
15.2.2	The Within Estimator: $T = 2$	642
15.2.3	The Within Estimator: $T > 2$	643
15.2.4	The Least Squares Dummy Variable Model	644
15.3	Panel Data Regression Error Assumptions	646
15.3.1	OLS Estimation with Cluster-Robust Standard Errors	648
15.3.2	Fixed Effects Estimation with Cluster-Robust Standard Errors	650
15.4	The Random Effects Estimator	651
15.4.1	Testing for Random Effects	653
15.4.2	A Hausman Test for Endogeneity in the Random Effects Model	654
15.4.3	A Regression-Based Hausman Test	656
15.4.4	The Hausman–Taylor Estimator	658
15.4.5	Summarizing Panel Data Assumptions	660
15.4.6	Summarizing and Extending Panel Data Model Estimation	661
15.5	Exercises	663
15.5.1	Problems	663
15.5.2	Computer Exercises	670
Appendix 15A	Cluster-Robust Standard Errors: Some Details	677
Appendix 15B	Estimation of Error Components	679
16	Qualitative and Limited Dependent Variable Models	681
16.1	Introducing Models with Binary Dependent Variables	682
16.1.1	The Linear Probability Model	683
16.2	Modeling Binary Choices	685
16.2.1	The Probit Model for Binary Choice	686
16.2.2	Interpreting the Probit Model	687
16.2.3	Maximum Likelihood Estimation of the Probit Model	690
16.2.4	The Logit Model for Binary Choices	693
16.2.5	Wald Hypothesis Tests	695
16.2.6	Likelihood Ratio Hypothesis Tests	696
16.2.7	Robust Inference in Probit and Logit Models	698
16.2.8	Binary Choice Models with a Continuous Endogenous Variable	698
16.2.9	Binary Choice Models with a Binary Endogenous Variable	699
16.2.10	Binary Endogenous Explanatory Variables	700
16.2.11	Binary Choice Models and Panel Data	701
16.3	Multinomial Logit	702
16.3.1	Multinomial Logit Choice Probabilities	703
16.3.2	Maximum Likelihood Estimation	703
16.3.3	Multinomial Logit Postestimation Analysis	704
16.4	Conditional Logit	707
16.4.1	Conditional Logit Choice Probabilities	707
16.4.2	Conditional Logit Postestimation Analysis	708
16.5	Ordered Choice Models	709
16.5.1	Ordinal Probit Choice Probabilities	710
16.5.2	Ordered Probit Estimation and Interpretation	711
16.6	Models for Count Data	713
16.6.1	Maximum Likelihood Estimation of the Poisson Regression Model	713
16.6.2	Interpreting the Poisson Regression Model	714
16.7	Limited Dependent Variables	717
16.7.1	Maximum Likelihood Estimation of the Simple Linear Regression Model	717
16.7.2	Truncated Regression	718
16.7.3	Censored Samples and Regression	718
16.7.4	Tobit Model Interpretation	720
16.7.5	Sample Selection	723
16.8	Exercises	725
16.8.1	Problems	725
16.8.2	Computer Exercises	733
Appendix 16A	Probit Marginal Effects: Details	739
16A.1	Standard Error of Marginal Effect at a Given Point	739
16A.2	Standard Error of Average Marginal Effect	740
Appendix 16B	Random Utility Models	741
16B.1	Binary Choice Model	741
16B.2	Probit or Logit?	742
Appendix 16C	Using Latent Variables	743
16C.1	Tobit (Tobit Type I)	743
16C.2	Heckit (Tobit Type II)	744
Appendix 16D	A Tobit Monte Carlo Experiment	745

A Mathematical Tools 748

A.1 Some Basics 749

- A.1.1 Numbers 749
- A.1.2 Exponents 749
- A.1.3 Scientific Notation 749
- A.1.4 Logarithms and the Number e 750
- A.1.5 Decimals and Percentages 751
- A.1.6 Logarithms and Percentages 751

A.2 Linear Relationships 752

- A.2.1 Slopes and Derivatives 753
- A.2.2 Elasticity 753

A.3 Nonlinear Relationships 753

- A.3.1 Rules for Derivatives 754
- A.3.2 Elasticity of a Nonlinear Relationship 757
- A.3.3 Second Derivatives 757
- A.3.4 Maxima and Minima 758
- A.3.5 Partial Derivatives 759
- A.3.6 Maxima and Minima of Bivariate Functions 760

A.4 Integrals 762

- A.4.1 Computing the Area Under a Curve 762

A.5 Exercises 764

B Probability Concepts 768

B.1 Discrete Random Variables 769

- B.1.1 Expected Value of a Discrete Random Variable 769
- B.1.2 Variance of a Discrete Random Variable 770
- B.1.3 Joint, Marginal, and Conditional Distributions 771
- B.1.4 Expectations Involving Several Random Variables 772
- B.1.5 Covariance and Correlation 773
- B.1.6 Conditional Expectations 774
- B.1.7 Iterated Expectations 774
- B.1.8 Variance Decomposition 774
- B.1.9 Covariance Decomposition 777

B.2 Working with Continuous Random Variables 778

- B.2.1 Probability Calculations 779
- B.2.2 Properties of Continuous Random Variables 780
- B.2.3 Joint, Marginal, and Conditional Probability Distributions 781
- B.2.4 Using Iterated Expectations with Continuous Random Variables 785

B.2.5 Distributions of Functions of Random Variables 787

B.2.6 Truncated Random Variables 789

B.3 Some Important Probability Distributions 789

- B.3.1 The Bernoulli Distribution 790
- B.3.2 The Binomial Distribution 790
- B.3.3 The Poisson Distribution 791
- B.3.4 The Uniform Distribution 792
- B.3.5 The Normal Distribution 793
- B.3.6 The Chi-Square Distribution 794
- B.3.7 The t -Distribution 796
- B.3.8 The F -Distribution 797
- B.3.9 The Log-Normal Distribution 799

B.4 Random Numbers 800

B.4.1 Uniform Random Numbers 805

B.5 Exercises 806

C Review of Statistical Inference 812

C.1 A Sample of Data 813

C.2 An Econometric Model 814

C.3 Estimating the Mean of a Population 815

- C.3.1 The Expected Value of \bar{Y} 816
- C.3.2 The Variance of \bar{Y} 817
- C.3.3 The Sampling Distribution of \bar{Y} 817
- C.3.4 The Central Limit Theorem 818
- C.3.5 Best Linear Unbiased Estimation 820

C.4 Estimating the Population Variance and Other Moments 820

- C.4.1 Estimating the Population Variance 821
- C.4.2 Estimating Higher Moments 821

C.5 Interval Estimation 822

- C.5.1 Interval Estimation: σ^2 Known 822
- C.5.2 Interval Estimation: σ^2 Unknown 825

C.6 Hypothesis Tests About a Population Mean 826

- C.6.1 Components of Hypothesis Tests 826
- C.6.2 One-Tail Tests with Alternative “Greater Than” ($>$) 828
- C.6.3 One-Tail Tests with Alternative “Less Than” ($<$) 829
- C.6.4 Two-Tail Tests with Alternative “Not Equal To” (\neq) 829
- C.6.5 The p -Value 831
- C.6.6 A Comment on Stating Null and Alternative Hypotheses 832

- C.6.7** Type I and Type II Errors 833
 - C.6.8** A Relationship Between Hypothesis Testing and Confidence Intervals 833
- C.7 Some Other Useful Tests 834**
 - C.7.1** Testing the Population Variance 834
 - C.7.2** Testing the Equality of Two Population Means 834
 - C.7.3** Testing the Ratio of Two Population Variances 835
 - C.7.4** Testing the Normality of a Population 836
- C.8 Introduction to Maximum Likelihood Estimation 837**
 - C.8.1** Inference with Maximum Likelihood Estimators 840
 - C.8.2** The Variance of the Maximum Likelihood Estimator 841
 - C.8.3** The Distribution of the Sample Proportion 842
 - C.8.4** Asymptotic Test Procedures 843
- C.9 Algebraic Supplements 848**
 - C.9.1** Derivation of Least Squares Estimator 848
 - C.9.2** Best Linear Unbiased Estimation 849

- C.10 Kernel Density Estimator 851**
- C.11 Exercises 854**
 - C.11.1** Problems 854
 - C.11.2** Computer Exercises 857

D Statistical Tables 862

- Table D.1** Cumulative Probabilities for the Standard Normal Distribution $\Phi(z) = P(Z \leq z)$ 862
- Table D.2** Percentiles of the t -distribution 863
- Table D.3** Percentiles of the Chi-square Distribution 864
- Table D.4** 95th Percentile for the F -distribution 865
- Table D.5** 99th Percentile for the F -distribution 866
- Table D.6** Standard Normal pdf Values $\phi(z)$ 867

- INDEX 869

List of Examples

- Example P.1** Using a *cdf* 19
- Example P.2** Calculating a Conditional Probability 21
- Example P.3** Calculating an Expected Value 24
- Example P.4** Calculating a Conditional Expectation 25
- Example P.5** Calculating a Variance 26
- Example P.6** Calculating a Correlation 28
- Example P.7** Conditional Expectation 30
- Example P.8** Conditional Variance 31
- Example P.9** Iterated Expectation 32
- Example P.10** Covariance Decomposition 34
- Example P.11** Normal Distribution Probability Calculation 36
- Example 2.1** A Failure of the Exogeneity Assumption 53
- Example 2.2** Strict Exogeneity in the Household Food Expenditure Model 54
- Example 2.3** Food Expenditure Model Data 59
- Example 2.4a** Estimates for the Food Expenditure Function 63
- Example 2.4b** Using the Estimates 64
- Example 2.5** Calculations for the Food Expenditure Data 75
- Example 2.6** Baton Rouge House Data 78
- Example 2.7** Baton Rouge House Data, Log-Linear Model 81
- Example 3.1** Interval Estimate for Food Expenditure Data 116
- Example 3.2** Right-Tail Test of Significance 123
- Example 3.3** Right-Tail Test of an Economic Hypothesis 124
- Example 3.4** Left-Tail Test of an Economic Hypothesis 125
- Example 3.5** Two-Tail Test of an Economic Hypothesis 125
- Example 3.6** Two-Tail Test of Significance 126
- Example 3.3(continued)** p -Value for a Right-Tail Test 127
- Example 3.4(continued)** p -Value for a Left-Tail Test 128
- Example 3.5(continued)** p -Value for a Two-Tail Test 129
- Example 3.6(continued)** p -Value for a Two-Tail Test of Significance 129
- Example 3.7** Estimating Expected Food Expenditure 130
- Example 3.8** An Interval Estimate of Expected Food Expenditure 131
- Example 3.9** Testing Expected Food Expenditure 132
- Example 4.1** Prediction in the Food Expenditure Model 156
- Example 4.2** Goodness-of-Fit in the Food Expenditure Model 159
- Example 4.3** Reporting Regression Results 159
- Example 4.4** Using the Linear-Log Model for Food Expenditure 164
- Example 4.5** Heteroskedasticity in the Food Expenditure Model 167
- Example 4.6** Testing Normality in the Food Expenditure Model 168
- Example 4.7** Influential Observations in the Food Expenditure Data 171
- Example 4.8** An Empirical Example of a Cubic Equation 172
- Example 4.9** A Growth Model 174
- Example 4.10** A Wage Equation 175
- Example 4.11** Prediction in a Log-Linear Model 176
- Example 4.12** Prediction Intervals for a Log-Linear Model 177
- Example 4.13** A Log-Log Poultry Demand Equation 178
- Example 5.1** Data for Hamburger Chain 200
- Example 5.2** OLS Estimates for Hamburger Chain Data 206
- Example 5.3** Error Variance Estimate for Hamburger Chain Data 208

- Example 5.4** R^2 for Hamburger Chain Data 209
- Example 5.5** Variances, Covariances, and Standard Errors for Hamburger Chain Data 214
- Example 5.6** Interval Estimates for Coefficients in Hamburger Sales Equation 216
- Example 5.7** Interval Estimate for a Change in Sales 218
- Example 5.8** Testing the Significance of Price 219
- Example 5.9** Testing the Significance of Advertising Expenditure 220
- Example 5.10** Testing for Elastic Demand 220
- Example 5.11** Testing Advertising Effectiveness 221
- Example 5.12** Testing the Effect of Changes in Price and Advertising 222
- Example 5.13** Cost and Product Curves 223
- Example 5.14** Extending the Model for Burger Barn Sales 224
- Example 5.15** An Interaction Variable in a Wage Equation 225
- Example 5.16** A Log-Quadratic Wage Equation 226
- Example 5.17** The Optimal Level of Advertising 232
- Example 5.18** How Much Experience Maximizes Wages? 233
- Example 5.19** An Interval Estimate for $\exp(\beta_2/10)$ 249
- Example 5.20** An Interval Estimate for β_1/β_2 250
- Example 5.21** Bootstrapping for Nonlinear Functions $g_1(\beta_2) = \exp(\beta_2/10)$ and $g_2(\beta_1, \beta_2) = \beta_1/\beta_2$ 258
- Example 6.1** Testing the Effect of Advertising 262
- Example 6.2** The F -Test Procedure 263
- Example 6.3** Overall Significance of Burger Barns Equation 265
- Example 6.4** When are t - and F -tests equivalent? 266
- Example 6.5** Testing Optimal Advertising 267
- Example 6.6** A One-Tail Test 268
- Example 6.7** Two ($J = 2$) Complex Hypotheses 268
- Examples 6.2 and 6.5** Revisited 270
- Example 6.8** A Nonlinear Hypothesis 270
- Example 6.9** Restricted Least Squares 272
- Example 6.10** Family Income Equation 275
- Example 6.11** Adding Children Aged Less Than 6 Years 277
- Example 6.12** Adding Irrelevant Variables 277
- Example 6.13** A Control Variable for Ability 279
- Example 6.14** Applying RESET to Family Income Equation 282
- Example 6.15** Forecasting SALES for the Burger Barn 284
- Example 6.16** Predicting House Prices 287
- Example 6.17** Collinearity in a Rice Production Function 291
- Example 6.18** Influential Observations in the House Price Equation 293
- Example 6.19** Nonlinear Least Squares Estimates for Simple Model 295
- Example 6.20** A Logistic Growth Curve 296
- Example 7.1** The University Effect on House Prices 321
- Example 7.2** The Effects of Race and Sex on Wage 323
- Example 7.3** A Wage Equation with Regional Indicators 325
- Example 7.4** Testing the Equivalence of Two Regressions: The Chow Test 327
- Example 7.5** Indicator Variables in a Log-Linear Model: The Rough Approximation 330
- Example 7.6** Indicator Variables in a Log-Linear Model: An Exact Calculation 330
- Example 7.7** The Linear Probability Model: An Example from Marketing 332
- Example 7.8** An Application of Difference Estimation: Project STAR 335
- Example 7.9** The Difference Estimator with Additional Controls 336
- Example 7.10** The Difference Estimator with Fixed Effects 337
- Example 7.11** Linear Probability Model Check of Random Assignment 338

- Example 7.12** Estimating the Effect of a Minimum Wage Change: The DID Estimator 340
- Example 7.13** Estimating the Effect of a Minimum Wage Change: Using Panel Data 341
- Example 8.1** Heteroskedasticity in the Food Expenditure Model 372
- Example 8.2** Robust Standard Errors in the Food Expenditure Model 374
- Example 8.3** Applying GLS/WLS to the Food Expenditure Data 378
- Example 8.4** Multiplicative Heteroskedasticity in the Food Expenditure Model 382
- Example 8.5** A Heteroskedastic Partition 383
- Example 8.6** The Goldfeld–Quandt Test with Partitioned Data 384
- Example 8.7** The Goldfeld–Quandt Test in the Food Expenditure Model 385
- Example 8.8** Variance Stabilizing Log-transformation 389
- Example 8.9** The Marketing Example Revisited 391
- Example 8.10** Alternative Robust Standard Errors in the Food Expenditure Model 413
- Example 9.1** Plotting the Unemployment Rate and the GDP Growth Rate for the United States 419
- Example 9.2** Sample Autocorrelations for Unemployment 426
- Example 9.3** Sample Autocorrelations for GDP Growth Rate 427
- Example 9.4** Are the Unemployment and Growth Rate Series Stationary and Weakly Dependent? 428
- Example 9.5** Forecasting Unemployment with an AR(2) Model 431
- Example 9.6** Forecast Intervals for Unemployment from the AR(2) Model 434
- Example 9.7** Forecasting Unemployment with an ARDL(2, 1) Model 434
- Example 9.8** Choosing Lag Lengths in an ARDL(p , q) Unemployment Equation 437
- Example 9.9** Does the Growth Rate Granger Cause Unemployment? 438
- Example 9.10** Checking the Residual Correlogram for the ARDL(2, 1) Unemployment Equation 439
- Example 9.11** Checking the Residual Correlogram for an ARDL(1, 1) Unemployment Equation 440
- Example 9.12** LM Test for Serial Correlation in the ARDL Unemployment Equation 443
- Example 9.13** Okun’s Law 446
- Example 9.14** A Phillips Curve 450
- Example 9.15** The Phillips Curve with AR(1) Errors 455
- Example 9.16** A Consumption Function 458
- Example 9.17** Deriving Multipliers for an Infinite Lag Okun’s Law Model 459
- Example 9.18** Computing the Multiplier Estimates for the Infinite Lag Okun’s Law Model 460
- Example 9.19** Testing for Consistency of Least Squares Estimation of Okun’s Law 462
- Example 9.20** Durbin–Watson Bounds Test for Phillips Curve 479
- Example 10.1** Least Squares Estimation of a Wage Equation 489
- Example 10.2** IV Estimation of a Simple Wage Equation 495
- Example 10.3** 2SLS Estimation of a Simple Wage Equation 496
- Example 10.4** Using Surplus Instruments in the Simple Wage Equation 497
- Example 10.5** IV/2SLS Estimation in the Wage Equation 499
- Example 10.6** Checking Instrument Strength in the Wage Equation 502
- Example 10.7** Specification Tests for the Wage Equation 509
- Example 10.8** Testing for Weak Instruments 523
- Example 11.1** Supply and Demand for Truffles 541
- Example 11.2** Supply and Demand at the Fulton Fish Market 542
- Example 11.3** Klein’s Model I 544
- Example 11.4** Testing for Weak Instruments Using LIML 560

- Example 11.5** Testing for Weak Instruments with Fuller-Modified LIML 561
- Example 12.1** Plots of Some U.S. Economic Time Series 564
- Example 12.2** A Deterministic Trend for Wheat Yield 569
- Example 12.3** A Regression with Two Random Walks 575
- Example 12.4** Checking the Two Interest Rate Series for Stationarity 579
- Example 12.5** Is GDP Trend Stationary? 580
- Example 12.6** Is Wheat Yield Trend Stationary? 580
- Example 12.7** The Order of Integration of the Two Interest Rate Series 581
- Example 12.8** Are the Federal Funds Rate and Bond Rate Cointegrated? 583
- Example 12.9** An Error Correction Model for the Bond and Federal Funds Rates 585
- Example 12.10** A Consumption Function in First Differences 586
- Example 13.1** VEC Model for GDP 600
- Example 13.2** VAR Model for Consumption and Income 602
- Example 14.1** Characteristics of Financial Variables 617
- Example 14.2** Simulating Time-Varying Volatility 618
- Example 14.3** Testing for ARCH in BrightenYourDay (BYD) Lighting 620
- Example 14.4** ARCH Model Estimates for BrightenYourDay (BYD) Lighting 621
- Example 14.5** Forecasting BrightenYourDay (BYD) Volatility 621
- Example 14.6** A GARCH Model for BrightenYourDay 623
- Example 14.7** A T-GARCH Model for BYD 624
- Example 14.8** A GARCH-in-Mean Model for BYD 625
- Example 15.1** A Microeconomic Panel 636
- Example 15.1** Revisited 637
- Example 15.2** Using $T = 2$ Differenced Observations for a Production Function 641
- Example 15.3** Using $T = 2$ Differenced Observations for a Wage Equation 641
- Example 15.4** Using the Within Transformation with $T = 2$ Observations for a Production Function 642
- Example 15.5** Using the Within Transformation with $T = 3$ Observations for a Production Function 644
- Example 15.6** Using the Fixed Effects Estimator with $T = 3$ Observations for a Production Function 646
- Example 15.7** Using Pooled OLS with Cluster-Robust Standard Errors for a Production Function 650
- Example 15.8** Using Fixed Effects and Cluster-Robust Standard Errors for a Production Function 651
- Example 15.9** Random Effects Estimation of a Production Function 652
- Example 15.10** Random Effects Estimation of a Wage Equation 652
- Example 15.11** Testing for Random Effects in a Production Function 654
- Example 15.12** Testing for Endogenous Random Effects in a Production Function 656
- Example 15.13** Testing for Endogenous Random Effects in a Wage Equation 656
- Example 15.14** The Mundlak Approach for a Production Function 657
- Example 15.15** The Mundlak Approach for a Wage Equation 658
- Example 15.16** The Hausman–Taylor Estimator for a Wage Equation 659
- Example 16.1** A Transportation Problem 683
- Example 16.2** A Transportation Problem: The Linear Probability Model 684
- Example 16.3** Probit Maximum Likelihood: A Small Example 690
- Example 16.4** The Transportation Data: Probit 691
- Example 16.5** The Transportation Data: More Postestimation Analysis 692
- Example 16.6** An Empirical Example from Marketing 694

- Example 16.7** Coke Choice Model: Wald Hypothesis Tests 696
- Example 16.8** Coke Choice Model: Likelihood Ratio Hypothesis Tests 697
- Example 16.9** Estimating the Effect of Education on Labor Force Participation 699
- Example 16.10** Women's Labor Force Participation and Having More Than Two Children 700
- Example 16.11** Effect of War Veteran Status on Wages 701
- Example 16.12** Postsecondary Education Multinomial Choice 705
- Example 16.13** Conditional Logit Soft Drink Choice 708
- Example 16.14** Postsecondary Education Ordered Choice Model 712
- Example 16.15** A Count Model for the Number of Doctor Visits 715
- Example 16.16** Tobit Model of Hours Worked 721
- Example 16.17** Heckit Model of Wages 724
- Example A.1** Slope of a Linear Function 755
- Example A.2** Slope of a Quadratic Function 755
- Example A.3** Taylor Series Approximation 756
- Example A.4** Second Derivative of a Linear Function 758
- Example A.5** Second Derivative of a Quadratic Function 758
- Example A.6** Finding the Minimum of a Quadratic Function 759
- Example A.7** Maximizing a Profit Function 761
- Example A.8** Minimizing a Sum of Squared Differences 761
- Example A.9** Area Under a Curve 763
- Example B.1** Variance Decomposition: Numerical Example 776
- Example B.2** Probability Calculation Using Geometry 779
- Example B.3** Probability Calculation Using Integration 780
- Example B.4** Expected Value of a Continuous Random Variable 780
- Example B.5** Variance of a Continuous Random Variable 781
- Example B.6** Computing a Joint Probability 783
- Example B.7** Another Joint Probability Calculation 783
- Example B.8** Finding and Using a Marginal *pdf* 784
- Example B.9** Finding and Using a Conditional *pdf* 784
- Example B.10** Computing a Correlation 785
- Example B.11** Using Iterated Expectation 786
- Example B.12** Change of Variable: Continuous Case 788
- Example B.13** Change of Variable: Continuous Case 789
- Example B.14** An Inverse Transformation 801
- Example B.15** The Inversion Method: An Example 802
- Example B.16** Linear Congruential Generator Example 805
- Example C.1** Histogram of Hip Width Data 814
- Example C.2** Sample Mean of Hip Width Data 815
- Example C.3** Sampling Variation of Sample Means of Hip Width Data 816
- Example C.4** The Effect of Sample Size on Sample Mean Precision 818
- Example C.5** Illustrating the Central Limit Theorem 819
- Example C.6** Sample Moments of the Hip Data 822
- Example C.7** Using the Hip Data Estimates 822
- Example C.8** Simulating the Hip Data: Interval Estimates 824
- Example C.9** Simulating the Hip Data: Continued 825
- Example C.10** Interval Estimation Using the Hip Data 826
- Example C.11** One-tail Test Using the Hip Data 830
- Example C.12** Two-tail Test Using the Hip Data 830
- Example C.13** One-tail Test p -value: The Hip Data 831
- Example C.14** Two-Tail Test p -Value: The Hip Data 832
- Example C.15** Testing the Normality of the Hip Data 836

Example C.16 The “Wheel of Fortune” Game:
 $p = 1/4$ or $3/4$ 837

Example C.17 The “Wheel of Fortune” Game:
 $0 < p < 1$ 838

Example C.18 The “Wheel of Fortune” Game:
Maximizing the Log-likelihood 839

Example C.19 Estimating a Population
Proportion 840

Example C.20 Testing a Population
Proportion 843

Example C.21 Likelihood Ratio Test of the
Population Proportion 845

Example C.22 Wald Test of the Population
Proportion 846

Example C.23 Lagrange Multiplier Test of the
Population Proportion 848

Example C.24 Hip Data: Minimizing the Sum of
Squares Function 849