SECOND EDITION

USING ECONOMETRICS

A PRACTICAL GUIDE

A. H. STUDENMUND

USING ECONOMETRICS

A Practical Guide

SECOND EDITION

A. H. STUDENMUND
Occidental College

with the assistance of

HENRY J. CASSIDY
Federal Home Loan Mortgage Corporation

HarperCollinsPublishers

Dedicated to

Scott Richard Studenmund

Sponsoring Editor: John Greenman
Project Coordination, Text and Cover Design: Ruttle, Shaw & Wetherill, Inc.
Production Manager: Michael Weinstein
Compositor: Ruttle, Shaw & Wetherill, Inc.
Printer and Binder: R. R. Donnelley & Sons Company
Cover Printer: New England Book Components, Inc.

Using Econometrics: A Practical Guide, Second Edition

Library of Congress Cataloging-in-Publication Data

Studenmund, A. H.
Using econometrics : a practical guide / A. H. Studenmund, with the assistance of Henry J. Cassidy. —2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-673-52125-7
1. Econometrics. 2. Regression analysis. I. Cassidy, Henry J.
II. Title.
HB139.S795 1992
330'.01'5195—dc20 91-29439
CIP

94 9 8 7 6

Acknowledgments
Table B-1: Reprinted with permission of Hafner Press, a division of Macmillan Publishing Company from *Statistical Methods for Research Workers*, 14th Edition, by Ronald A. Fisher. Copyright © 1970 by University of Adelaide.

Tables B-2 and B-3: Abridged from M. Merrington and C.M. Thompson, "Tables of percentage points of the inverted beta (*F*) distribution," *Biometrica,* Vol. 38, 1951, pp. 159-77. By permission of the Biometrica Trustees.

Table B-4: From N. E. Savin and Kenneth White, "The Durbin-Watson Test for Serial Correlation with Extreme Sample Sizes or Many Regressors," *Econometrica,* Nov. 1977, p. 1994. Reprinted with permission.

Tables B-5, and B-6: From J. Durbin and G. S. Watson, "Testing for serial correlation in least squares regressions," *Biometrika,* Vol. 38, 1951, pp. 159-77. By permission of the Biometrica Trustees.

Tables B-7 and B-8: Based on *Biometrica Tables for Statisticians,* Vol. 1, 3rd ed. (1966). By permission of the Biometrica Trustees.

Preface

"Econometric education is a lot like learning to fly a plane; you learn more from actually doing it than you learn from reading about it."

Using Econometrics represents a new approach to the understanding of elementary econometrics. It covers the topic of single-equation linear regression analysis in an easily understandable format that emphasizes real-world examples and exercises. As the subtitle, *A Practical Guide*, implies, the book is aimed not only at beginning econometrics students but also at regression users looking for a refresher and at experienced practitioners who want a convenient reference.

The material covered by this book is traditional, but there are four specific features that we feel distinguish *Using Econometrics:*

1. Our approach to the learning of econometrics is simple, intuitive, and easy to understand. We do not use matrix algebra, and we relegate proofs and calculus to the footnotes.
2. We include numerous examples and example-based exercises. We feel that the best way to get a solid grasp of applied econometrics is through an example-oriented approach.
3. Although most of this book is at a simpler level than other econometrics texts, Chapters 6 and 7 on specification choice are among the most complete in the field. We think that an understanding of specification issues is vital for regression users.
4. We include a new kind of learning tool, called an interactive regression learning exercise, that helps students simulate econometric

analysis by giving them feedback on various kinds of decisions without relying on computer time or much instructor supervision.

The formal prerequisites for using this book are few. Most important, readers are assumed to have some familiarity with macroeconomic and microeconomic theory; in addition, the book is easier to use if readers have had a statistics course (even if they have forgotten most of it), or if readers are not afraid of working with mathematical functions. While all the statistical concepts necessary for econometric study are covered in the text, they are covered only to the extent needed for an understanding of regression analysis. Because the prerequisites are few and the statistics material is self-contained, *Using Econometrics* can be used not only in undergraduate econometrics courses but also in MBA-level courses in quantitative methods. We also have been told that the book is a helpful supplement for graduate-level econometrics courses.

What's New in the Second Edition?

The first edition of *Using Econometrics* was the most-adopted new econometrics textbook of the 1980s. As a result, we've been careful to write a second edition that adds more coverage and fifty percent more data sets, exercises, and examples *without* changing the clarity and practicality that made the first edition a success. Twice as many interactive regression learning exercises also have been added. Most of the new materials we added were suggested by users of the text, although some were recommended by reviewers intent on keeping us up with changes in standard practices in applied econometrics.

Specific additions include:

1. Chapter 12 on distributed lag models
2. Chapter 13 on dummy dependent variable techniques
3. an appendix on additional specification criteria in Chapter 6
4. a section on Variance Inflation Factors in Chapter 8
5. a section on the Breusch-Pagan and White tests in Chapter 10

Also, our expanded instructor's manual includes answers to odd numbered exercises, lecture notes, sample examinations, and an additional interactive exercise.

Finally, we're pleased to be able to include the student version of ECSTAT with each copy of the text at virtually no additional cost.

ECSTAT is the extremely user-friendly PC-based regression software package that we use to produce the regression results in the text. While the book is not tied to the use of ECSTAT in any way, we think students will learn to appreciate ECSTAT's accuracy and simplicity. The ECSTAT diskette also includes the data sets published in the text and a manual for using the program.

Acknowledgments

If this book has a spiritual father, it's Henry Cassidy of the Federal Home Loan Mortgage Corporation. It was Henry who saw the need for a follow-on to Rao and Miller's legendary *Applied Econometrics* and who coauthored the first edition of *Using Econometrics* as an expansion of his own work of the same name. Henry also contributed quite a bit to this second edition, as he reviewed all the new material and provided a healthy dose of good humor and inspiration.

From the point of view of this edition, the contributions of two superb economists stand out. Carolyn Summers of the National Education Association continued her invaluable first-edition role as editorial consultant, galley proofreader, indexer, and general gadfly-about-town. Mary Hirschfeld here at Occidental College reviewed the entire text, suggested a number of crucial improvements in its substance, and provided me with extremely perceptive feedback from the point of view of teachers of econometrics.

The quality of the outside reviewers of this edition was also quite high. In particular, I owe a real debt of gratitude to Rob Engle of UCSD for taking the time to give me in-depth feedback on the new material. Others in this excellent group of reviewers were Dennis Byrne (University of Akron), William Brown (CSU Northridge), Edward Coulson (Penn State), William Dawes (SUNY Stony Brook), Cliff Huang (Vanderbilt), Elia Kacapyr (Ithaca College), John Warner (Clemson), Tom Witt (West Virginia), and Phanindra Wunnava (Middlebury College).

Especially helpful in the editorial and production process were Paula Cousin, Jack Greenman, Dave Murphy, Ilana Scheiner, and Robert Dohner (the author of ECSTAT). Others who provided timely and useful assistance in this edition were Sonmez Atesoglu (Clarkson University), Sandra Chadwick, James Keeler (Kenyon College), Bruce Gensemer (Kenyon College), Marlene Penfold, and Justin Meyer (Silver Oak). Finally, but perhaps most important, I'd like to thank my superb Occidental College colleagues and students for their feedback and encouragement. These especially included Ned Cull, Richard Falkenrath, Greg Gordon, Eric Hilt, Nichole Becker Kamman, Jim Kee, Mark Lieberman, Manuel Pastor, Bob Sego, Bruce Stimpson, and Chris Tregillis.

A. H. Studenmund

Contents

1

An Overview of Regression Analysis

1.1 What Is Econometrics?

"Econometrics is too mathematical; it's the reason my best friend isn't majoring in economics."

"There are two things you don't want to see in the making—sausage and econometric research."[1]

"Econometrics may be defined as the quantitative analysis of actual economic phenomena."[2]

"It's my experience that 'economy-tricks' is usually nothing more than a justification of what the author believed before the research was begun."

1. Attributed to Edward E. Leamer.

2. Paul A. Samuelson, T. C. Koopmans, and J. R. Stone, "Report of the Evaluative Committee for *Econometrica*," *Econometrica,* 1954, p. 141.

Believe it or not, these are all actual quotations. Obviously, econometrics means different things to different people. To beginning students, it may seem as if econometrics is an overly complex obstacle to an otherwise useful education. To skeptical observers, econometric results should be trusted only when the steps that produced those results are completely known. To professionals in the field, econometrics is a fascinating set of techniques that allows the measurement and analysis of economic phenomena and the prediction of future economic trends.

You're probably thinking that such diverse points of view sound like the statements of blind people trying to describe an elephant based on what they happen to be touching, and you're partially right. Econometrics has both a formal definition and a larger context. While you can easily memorize the formal definition, you'll get the complete picture only by understanding the many uses of and alternative approaches to econometrics.

That said, we need a formal definition. **Econometrics,** literally "economic measurement," is the quantitative measurement and analysis of actual economic and business phenomena. It attempts to quantify economic reality and bridge the gap between the worlds of economic theory and actual business activity. To many students, these worlds may seem far apart. On the one hand, economists theorize equilibrium prices based on carefully conceived marginal costs and marginal revenues; on the other, many firms seem to operate as though they have never heard of such concepts. Econometrics allows us to examine data from real-world firms and to quantify the actions of these firms and other factors, such as the actions of consumers and governments. Such measurements have a number of different uses, and an examination of these uses is the first step to understanding econometrics.

1.1.1 Uses of Econometrics

Econometrics has three major uses:

1. the description of economic reality
2. the testing of hypotheses about economic theory
3. the forecasting of future economic activity.

The simplest use of econometrics is **description.** We can use econometrics to quantify economic activity; econometrics allows us to put numbers in equations that previously contained only abstract symbols. For example, consumer demand for a particular commodity often can be thought of as a relationship between the quantity demanded (C) and the commodity's price (P), the price of a substitute good (P_s), and disposable income (Y_d). For most goods, the relationship between consumption and disposable income is expected to be positive, because an increase in disposable income will be associated with an increase in the consumption of the good. Econometrics actually allows us to estimate that relationship based upon past consumption, income, and prices. In other words, a general and purely theoretical functional relationship like:

$$C = f(P, P_s, Y_d) \tag{1.1}$$

can become explicit:

$$C = -60.5 - 0.45P + 0.12P_s + 12.2Y_d \tag{1.2}$$

This technique gives a much more specific and descriptive picture of the function.[3] Let's compare Equations 1.1 and 1.2. Instead of expecting consumption merely to "increase" if there is an increase in disposable income, Equation 1.2 allows us to expect an increase of a specific amount (12.2 units for each unit of increased disposable income). The number 12.2 is called an estimated regression coefficient, and it is the ability to estimate these coefficients that makes econometrics valuable.

The second and perhaps the most common use of econometrics is **hypothesis testing,** the testing of alternative theories with quantitative evidence. Much of economics involves building theoretical models and testing them against evidence, and hypothesis testing is vital to that scientific approach. For example, you could test the hypothesis that the product in Equation 1.1 is what economists call a normal good (one for which the quantity demanded increases when disposable income increases). You could do this by applying various statistical tests to the estimated coefficient (12.2) of disposable income (Y_d) in Equation 1.2. At first glance, the evidence would seem to support this hypothesis because the coefficient's sign is positive, but the "statistical significance" of that estimate would have to be investigated before such a conclusion could be justified. Even though the estimated coefficient is positive, as expected, it may not be sufficiently different from zero to imply that the coefficient is indeed positive instead of zero. Unfortunately, statistical tests of such hypotheses are not always easy, and there are times when two researchers can look at the same set of data

3. The results in Equation 1.2 are from a model of the demand for chicken that we will examine in more detail in Section 6.1.

and come to different conclusions. Even given this possibility, the use of econometrics in testing hypotheses is probably its most important function.

The third and most difficult use of econometrics is to **forecast** or predict what is likely to happen next quarter, next year, or further into the future. For example, economists use econometric models to make forecasts of variables like sales, profits, Gross National Product (GNP), and the inflation rate. The accuracy of such forecasts depends in large measure on the degree to which the past is a good guide to the future. Business leaders and politicians tend to be especially interested in this use of econometrics because they need to make decisions about the future, and the penalty for being wrong (bankruptcy for the entrepreneur and political defeat for the candidate) is high. To the extent that econometrics can shed light on the impact of their policies, business and government leaders will be better equipped to make decisions. For example, if the president of a company that sold the product modeled in Equation 1.1 wanted to decide whether to increase prices, forecasts of sales with and without the price increase could be calculated and compared to help make such a decision. In this way, econometrics can be used not only for forecasting but also for policy analysis.

1.1.2 Alternative Econometric Approaches

There are many different approaches to econometrics. For example, the fields of biology, psychology, and physics all face quantitative questions similar to those faced in economics and business. However, these fields tend to use somewhat different techniques for analysis because the problems they face aren't the same. Different approaches also make sense within the field of economics. The kind of econometric tools used to quantify a particular function depends in part on the uses to which that equation will be put. A model built solely for descriptive purposes might be different from a forecasting model, for example.

To get a better picture of these approaches, let's look at the steps necessary for any kind of quantitative research:

1. specifying the models or relationships to be studied
2. collecting the data needed to quantify the models
3. quantifying the models with the data.

Steps 1 and 2 are similar in all quantitative work, but the techniques used in step 3, the quantification of models, differ widely between and within disciplines. Choosing among techniques for the quantification of a model given a particular set of data is often referred to as the "art" of econometrics. There are many different alternative approaches to quantifying the same equation, and each approach may give somewhat different results. The choice of approach is left to the individual econometrician (the researcher using econometrics), but each researcher should be able to justify that choice.

This book will focus primarily on one particular econometric approach: single-equation linear *regression analysis*. The majority of this book will thus discuss only regression analysis, but it is important for every econometrician to remember that regression is only one of many approaches to econometric quantification.

The importance of critical evaluation cannot be stressed enough; a good econometrician is one who can diagnose faults in a particular approach and figure out how to repair them. The limitations of the regression analysis approach must be fully perceived and appreciated by anyone attempting to use regression analysis or its findings. The possibility of missing or inaccurate data, incorrectly formulated relationships, poorly chosen estimating techniques, or improper statistical testing procedures implies that the results from regression analyses need to be viewed with some caution.

1.2 What Is Regression Analysis?

Econometricians use regression analysis to make quantitative estimates of economic relationships that previously have been completely theoretical in nature. After all, anybody can claim that the quantity of compact discs demanded will increase if the price of those discs decreases (holding everything else constant), but not many people can actually put numbers into an equation and estimate *by how many* compact discs the quantity demanded will increase for each dollar that price decreases. To predict the *direction* of the change, you need a knowledge of economic theory and the general characteristics of the product in question. To predict the *amount* of the change, though, you need a sample of data, and you need a way to estimate the relationship. The most frequently used method to estimate such a relationship in econometrics is regression analysis.

1.2.1 Dependent Variables, Independent Variables, and Causality

Regression analysis is a statistical technique that attempts to "explain" movements in one variable, the **dependent variable,** as a function of movements in a set of other variables, called the **independent** (*or* **explanatory**) **variables,** through the quantification of a single equation. For example, in Equation 1.1:

$$C = f(P, P_s, Y_d) \tag{1.1}$$

C is the dependent variable and P, P_s, and Y_d are the independent variables. Regression analysis is a natural tool for economists because most economic propositions can be stated in such single-equation[4] functional forms. For example, the quantity demanded (dependent variable) is a function of price, income, and the prices of substitutes (independent variables).

Much of economics and business is concerned with cause-and-effect propositions: If the price of a good increases by one unit, then the quantity demanded decreases on average by a certain amount, depending on the price elasticity of demand (defined as the percentage change in the quantity demanded that is caused by a one percent change in price). Similarly, if the quantity of capital employed increases by one unit, then output increases by a certain amount, called the marginal productivity of capital. Propositions such as these pose an if-then, or causal, relationship that logically postulates a dependent variable having movements that are causally determined by movements in a number of specified independent variables.

Don't be deceived by the words dependent and independent, however. While many economic relationships are causal by their very nature, a regression result, no matter how statistically significant, cannot prove causality. All regression analysis can do is test whether a significant quantitative relationship exists. Judgments as to causality must also include a healthy dose of economic theory and common sense. For example, the fact that the bell on the door of a flower shop rings just before a customer enters and purchases some flowers by no means implies that the ringing of the bell causes the purchase! If events A and B are related statistically, it may be that A causes B, that B causes A, that some omitted factor causes both, or that a chance correlation exists between the two.

The cause and effect relationship is often so subtle that it fools

4. Often there are several related propositions that, when taken as a group, suggest a *system* of regression equations. An example is a two-equation model of supply and demand. Usually, these two equations must be considered simultaneously instead of separately. The estimation of such simultaneous systems will be discussed in Chapter 14.