Preface

to

QUANTITATIVE ECONOMICS & ECONOMETRICS

Brennan

and

Carroll

Fourth

Edition

Preface to Quantitative Economics & Econometrics

Michael J. Brennan

Professor of Economics
Wesleyan University

Thomas M. Carroll

Professor of Economics
University of Nevada, Las Vegas

Fourth Edition

H23

Published by

SOUTH-WESTERN PUBLISHING CO.

CINCINNATI   WEST CHICAGO, IL   DALLAS   LIVERMORE, CA

ISBN: 0-538-08230-5

Library of Congress Catalog Card Number: 86-60209

2 3 4 5 6 7 D 2 1 0 9 8 7

Printed in the United States of America

Preface to Fourth Edition

It has been twenty-four years since the first edition of Michael Brennan's *Preface to Econometrics* and thirteen years since the third edition. In that time there has been considerable advancement in the techniques and technology of econometrics. Professor Brennan's book has continued to be a mainstay in the instruction of quantitative economics and econometrics because he addresses himself directly to the student, communicating clearly the concepts of quantitative economics and econometrics.

In collaborating with Professor Brennan on a fourth edition of his text, I have tried to remain true to those qualities that made earlier editions of this text classics. Where there have been changes, the goal has been to enhance the original Brennan concept. First, we have changed the name to more clearly reflect the content of the text. Unlike many texts that divorce the principles of mathematical economics from those of econometrics, we have emphasized that both quantitative economics and econometrics are critical to economic research. Second, we have blended the economic examples more directly into the text, showing in every chapter real-world applications of the mathematical and statistical concepts. Third, we have extended the discussions of matrix algebra (Chapter 5) and statistical estimation (Part III) to incorporate major advancements in computer software. No longer do we treat matrix inversion and multiple regression as the tedious processes they were even a decade ago.

I wish to thank Michael Brennan for the confidence he showed by agreeing to this collaboration. Second, I would like to thank my former colleagues at Memphis State University, particularly Professor Karen Pickerill, for constructive comments over a period of eighteen months. Finally, I thank my undergraduate and graduate students at Memphis State, who read and criticized early drafts of this text as the book was classroom tested.

Thomas M. Carroll
University of Nevada, Las Vegas

Introduction

This text addresses two topics: *quantitative economics* and *econometrics*. Quantitative economics deals with the use of mathematical concepts and techniques in the analysis of economic phenomena. Econometrics involves the use of statistical methods to estimate the equations that describe economic events. Quantitative economics and econometrics are two closely related, but distinct aspects of economic science. They are related in the sense that the econometrician uses mathematical methods to generate economic models to be subjected to empirical testing. Furthermore, the procedures of econometrics provide estimates of the unknowns in mathematical models. It would not be exaggerating to say that mathematical economics is the foundation of econometrics, and that econometrics adds realism to abstract mathematical models. Hence, it is natural for a text to treat both quantitative economics and econometrics. But it is also reasonable to address these topics separately.

Despite the intimate relation between mathematical economics and econometrics, they both lead lives of their own. Many theoretical economics texts and articles are written in general mathematical terms without concern for the precise values of economic variables. Mathematics, by itself, lends precision to economic concepts that are more difficult to grasp with nonmathematical language. For instance, a mathematical economist may use a particular symbol to stand for the price and another to stand for the marginal cost of production. The principles of calculus can be employed to show that profit maximization by a competitive firm requires that marginal cost equal price.

While mathematical economics can stand without econometrics, it is less clear that econometrics can be used or understood without a basic understanding of mathematical concepts. It is true that modern electronic computers have taken much of the drudgery out of mathematical computation. But computers, even loaded with sophisticated econometric software, only do what they are told. We must have a pretty good idea of what we want the computer to do, or we will have a difficult time interpreting the computer output. That is why this book devotes half of its contents to mathematical economics. Chapters 1 through 11 develop the major mathematical tools necessary for understanding econometrics. These chapters can stand alone in a course in quantitative or mathematical economics, or they can serve as background in a course in econometrics.

Econometrics might simply be defined as a bridge between economic theory and reality. Obviously some elaboration is in order. Life is rich in experience, vast in complexity, and intimidating in importance. Theory is simple, precise, and elegant. Theory cuts away irrelevant detail and organizes what remains into a model of an ideal system. Some people are put off by the lack of realism of theory; they fail to appreciate that it is precisely its abstraction that gives theory its ability to predict. Ironically, some theorists lose patience with the sloppiness of the real world; they fail to appreciate the fact that the ultimate test of a theory is its ability to predict and/or explain actual events.

The problem of testing theory—any theory—rests in reconciling observed events and theoretical explanations. In the physical sciences, theories are tested in controlled experiments. The chemist or the physicist literally creates an imaginary world in the laboratory so that "causes" and "effects" can be clearly isolated and identified. In the behavioral sciences the researcher contrasts an experimental group with a control group. Since factors influencing human (and even animal) behavior are too complex to control directly, researchers rely on the rules of probability to isolate the influence of a stipulated stimulus (e.g., stress) on the subject population. In economics, we use published data on past events to estimate the mathematical structure of theories, comparing the statistical relationship between "causes" and "effects" with those predicted by theory. The greater the probability that the observations could have been generated by chance, the less credence is given to the theory.

Econometrics essentially follows a four-step process, commonly known as the *scientific method*: (1) the assembly of facts and similar information, (2) the formation of a hypothesis or hypotheses about the behavior of economic variables or the causes of events, (3) the derivation of testable assertions or predictions that are logically derived from the hypotheses, and (4) the testing of predictions by the reference to observed facts. We might refer to step (1) as *experience*, step (2) as *theorizing*, step (3) as *mathematical reasoning* and step (4) as *statistical inference*.

Each step can be self-contained, practiced and studied in isolation from the others. Many practitioners of the business arts are sometimes bored by theory, terrified of mathematics, and confused by statistics. Yet it is amazing how frequently they rely on "rules of thumb" that are little else but *ad hoc* theories. They also survive by "seat of the pants" calculations and playing their hunches. When we think about it, most businesspersons are actually amateur econometricians!

Similarly, the economics profession, and indeed the world, is full of "pure" theorists, thinkers who so admire the elegance of a theoretical system that they disdain the real world, or perhaps more dangerously, mistake the real world for their ideal type. Having a nearly religious commitment to the principles of their theories, they see no need to test their beliefs. If "the real world" fails to conform to the predictions of their theories, so much the worse for the real world! Alas, when reality is seen as the servant of theory, rather than the other way around, hostile camps of conflicting theories coalesce around conflicting interest groups. The reputations and careers that are destroyed in the name of preserving the purity of theory are legion. If different theoretical propositions are but put to the test, much of the disagreement could be resolved.

Mathematics is a language that facilitates the communication of economic concepts. Most economic concepts are quantitative; prices, income, saving, amounts of commodities produced and consumed, and many other economic magnitudes are, at least in principle, measurable. The use of mathematical symbols is perfectly natural in economics. There is no fundamental difference between mathematical economic theory and economic theory that does not use mathematics. Although the same conclusions can be reached whether the theory is stated verbally or mathematically, there are definite advantages to mathematical formulation: (1) it introduces rigor into the definitions and relationships, (2) it makes the assumptions explicit at each stage of the reasoning process and thereby avoids hidden assumptions not easily discovered, (3) it brings out clearly the limitations of the theory, and (4) it identifies implications that might be overlooked in purely verbal presentations.

Econometrics differs somewhat from mathematical economics, however. It differs in that its mathematical formulations are designed with a view to statistical measurement and testing. Statistics is also an important aspect of econometrics, yet we must distinguish econometrics from what we may call statistical economics. The latter is a form of quantitative economics that avoids economic theory and claims to provide a statistical summary of the economic data themselves. The recording and charting of the gross national product of the total United States is an example of statistical economics. Another example is the tabulation of the consumer price index. But the mere accumulation and ordering of data seldom provide solutions to important economic questions. Some theory is required to interpret the behavior of items in the data; explanations do not come ready-made. Since some theory is unavoidable, it is usually best to state the theory explicitly. Econometrics, unlike statistical economics, synthesizes theoretical ideas, mathematical form, and quantitative evidence.

Econometrics, when it is done well, blends the elements of experience, theory, mathematics, and statistical inference to provide a unified picture of economics. A practicing econometrician is a generalist, a practical theorist with a working understanding of both mathematics and statistics. This book is designed to be the first step in the training of an econometrician. No doubt you already possess, in varying degrees, the various talents and aptitudes necessary to master the subject matter of this text. You certainly have had some experiences. You have also taken previous courses in mathematics, statistics, and economics. Yet if your experience is like that of most students (including that of the authors of this book), these four facets of econometrics have been learned in isolation from each other. So this book is designed to put the pieces together.

Part I covers the essentials of mathematical economics, with the goal of understanding how to translate verbal economic arguments into their mathematical equivalents. Many economic examples are used to develop these topics. You may discover that your knowledge of economics will help you understand mathematical concepts that heretofore seemed like so many abstractions. Alternatively your deepening understanding of mathematics will reveal insights into economics you previously missed with a "verbal" orientation to the subject.

Part II presents a transition from quantitative economics to econometrics, reviewing how risk and uncertainty play a crucial role in the estimation of economic models. Part III deals with the problem of statistical inference as it relates to the testing of economic theories. These chapters are rich in economic applications, including many tests of controversial theories. It is likely that you will not agree with all the conclusions we make. But that is the point of econometrics. Disagreements should be resolved by recourse to the data. Accordingly, the data for most of the examples in the text are provided on a microcomputer disk that accompanies your professor's instructor's manual. You should find that by retracing our steps, and trying a few twists of your own, many disagreements can be resolved.

Many students find that the easiest way through a course is to psyche out the professor, memorize the relevant concepts, and forget the material after the final exam. But that is not the best approach to econometrics. Econometrics is "hands-on" economics, designed to subject our most closely held convictions to objective tests. Econometrics must be understood in order to be applied, and it must be applied in order to be understood. This book is a preface to econometrics, a first step to understanding the interaction between economic theory and the real world.

Contents

Part II From Mathematics to Statistics

Part III Econometrics

1

Variables and Functions

As shown in the preceding introductory sketch of econometrics, there are two aspects to the empirical science of economics: (1) theoretical explanations of events in the economy are phrased in terms of mathematics; and (2) the mathematical equations that describe economic behavior are subjected to statistical analysis. In the early chapters we will focus on the former aspect. We shall see how familiar notions of economic theory can be phrased in mathematical terms. Mathematics is a language for the expression of theoretical propositions. For the most part, you will already be familiar with the economic ideas we will use to introduce some elementary mathematical terms. We shall see how these terms, once understood, provide deeper insights into economic concepts by adding both rigor and clarity.

VARIABLES

Before we can formulate economic theory in mathematical terms, it is necessary to understand the mathematical elements that make up that formulation. The most basic of these is the variable. A *variable* may be defined as a quantity that can assume any value from a given set of numbers. A variable is usually symbolized by a letter, such as x, y, or z, or a terse group of letters used to describe its meaning (e.g., *CON* for consumption, *MC* for marginal cost, or *MR* for marginal revenue). Suppose that the variable under consideration is called x and that x can take on any positive value from zero to one million. We represent the potential values of x as a set, called the *domain* of the variable. Individual values of the domain, such as the numbers 0, 0.5, 1, 2, 15.125, 27, and 1,000, are each elements of that set. We show that 2 is an element of the set of numbers between zero and one million in the following way: $2 \in \{0 < x < 1{,}000{,}000\}$, where "$\epsilon$" (the Greek letter "epsilon") means "is an element of."

The domain of a variable may be either finite or infinite. A *finite* set has a specified number of elements; the number of elements can be counted. The set of integers between 0 and 10 is a finite set; the entire set can be enumerated: $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. It follows that an infinite set cannot be listed explicitly because the number of elements cannot be counted. The set of positive integers is an infinite set. The entire set can only be listed by stipulation of the rule for inclusion: $\{x \text{ is an integer and } x > 0\}$.

You are no doubt familiar with a number of mathematical operations that can be performed on variables. For instance, if we are concerned with the set of positive integers, the operations addition, multiplication, and exponentiation performed between two or more elements in the set of positive integers generates a result (value for x) that is also an element of that domain. In other words, the sum of two or more positive integers is also a positive integer; the product of two or more positive integers is also a positive integer; and a positive integer raised to a power that is also a positive integer produces an answer that is a positive integer. By contrast, the operations subtraction and division, when performed between two or more members of the set of positive integers, do not always yield results that are positive integers. If we take two values of x, say x_1 and x_2, such that both are positive integers and $x_2 > x_1$, then $(x_2 - x_1) \in \{\text{positive integers}\}$ while $(x_1 - x_2) \notin \{\text{positive integers}\}$. (The epsilon with a slash through it is read "is not an element of.") If x_1 is a prime number, there is no positive integer, aside from x_1 itself and the number "1" which, when divided into x_1, will yield another positive integer.

The set of *real numbers* is used frequently in economics. This set is composed of all *rational numbers* (numbers that can be expressed as the *ratio* of two integers) and the set of *irrational numbers* (which cannot be expressed as the ratio of two integers). An attractive feature of the set of *real numbers* is that the five common mathematical operations—addition, subtraction, multiplication, division, and exponentiation—are all defined over the domain of real numbers. If we add any two real numbers, our result is a real number; if we multiply any two real numbers, our result is a real number. The set of real numbers includes negative as well as positive integers. Hence, if we subtract a larger integer from a smaller integer, our result is a negative integer, which is, of course a real number. When one integer is divided by another, the result is a rational number, although not necessarily an integer. All rational numbers are included in the set of real numbers.

Recall that an integer raised to an integral power generates an integer; but an integer raised to a power that is not an integer will not normally produce an integer. A variable raised to an exponent that is a negative number is equivalent to taking the reciprocal of that variable raised to the absolute value of that power; e.g., $x^{-2} = 1/(x^2)$. A variable raised to a fractional power yields a result that is equivalent to raising the variable to the power indicated by the numerator of the fraction, then taking the root indicated by the denominator of the fraction; e.g., $x^{1/2} = \sqrt{x}$; $x^{2/3} = \sqrt[3]{x^2}$.

It should be clear why economists and other scientists usually specify the set of real numbers as the domain for the variables of their theories; specification of a more restricted domain constricts the freedom of investigators to employ mathematical operations on those variables. Nevertheless, there are some instances when relevance dictates a limit to the set of values a variable can assume. In this universe, mass and energy cannot assume negative numbers. When economists speak of technology, their subject is bound by the rules of known science. Therefore, we cannot speak of negative outputs or negative inputs, although changes of inputs or outputs could be positive, negative, or zero. By similar logic, it usually makes better sense to constrain prices and costs to the set of positive real numbers.

FUNCTIONS

As you can see, many diverse phenomena can be treated as variables in economics: rates of input and output; flows of income, cost, and revenue; stocks of machinery and inventories. But not even the pure mathematician is interested in variables for their own sake. The mathematician attempts to relate one or more variables to others. In economics the theorist tries to determine the connection among relevant concepts (variables). A *relation* is formally defined as *a set of ordered pairs*. This means that the relationship between two variables, say x and y, defines a set of paired values, in which the first element in each pair is a value for x, and the second element of that pair is a corresponding value of y. We can let the letter R denote some *relation* between them. For instance, let $R(x, y)$ stand for the relationship "y is greater than x": $R(x, y) = \{(x, y) \text{ such that } y > x\}$. The pair of numbers (2, 4) belongs to the relation, whereas the pair (4, 2) does not.

A *function* is a special kind of relation with the property that each value of the first element (drawn from a set called the *domain*) of the ordered pair is associated with a unique value of the second element (drawn from a set called the *range*). The relation defined in the previous paragraph is not a function; we can find two pairs, say (2, 4) and (2, 5), which both satisfy the relation. For a specified value of x, there is more than one admissible value of y. By contrast, the relation $y = x^2$ is a function; whatever real number we specify as the value of x, there is only one value of y that is admissible: namely, the result of x multiplied by itself. The variable y is said to be a function of x if, once the value of x has been specified, the value of y is uniquely determined. The variable x, whose value may be arbitrarily assigned, is called the *independent* variable. The variable y, whose numerical value is stipulated by the selection of x, is called the *dependent* variable.

To this point we have defined functions as relations between only two variables. These are the types of functions with which you are most familiar, given the prevalence of two-dimensional diagrams in economics. However, we will learn presently that a dependent variable can also be defined for combinations of two or more independent variables.

The standard notation for defining a functional relation between the dependent variable y and the independent variable x is: $y = f(x)$ (although nearly any letter or symbol could be used in place of "f"). This notation means that the value of y depends on the value of x in some unspecified way.

To appreciate the usefulness of this mathematical symbolism, consider an example from economics. Let p represent the price of a commodity and q the quantity of that commodity that a household will buy, given values for the prices of complements, substitutes, the household's real income, and the tastes of family members. The demand function can be written $q = D(p)$; the quantity demanded is unique for each price, although the relation is as yet unspecified. Estimation of the demand equation could yield explicit mathematical functions, such as $q = 100 - 2p$ or $q = 20p^{-2}$. Likewise, the consumption function might be written as $C = C(Y_d)$, where the dependent variable is aggregate consumption and the independent variable is aggregate disposable income. Statistical techniques could be employed to determine a specific mathematical relation between the two variables. No doubt you can think of many more examples from economics.

We can now discuss the specification of functional form in more detail. As we will see in Part III, before we attempt an empirical estimation of a function, it is necessary to specify a general mathematical form the equation will take. An *equation* is a statement of equality between two mathematical entities. It stipulates the value of one variable, once the value of the other has been specified. The convention is to place the dependent variable on the left side of the equal sign and the independent variable on the right side. This notation is referred to as the *explicit form* of the function. The familiar "intercept-slope" form of the linear equation, e.g., $y = 6 + 3x$, is an example of the explicit form. This equation says that the value of y is determined by taking the value of x, multiplying it by 3, and adding 6. The value of y is unique for each value of x because each time the value of x changes, the value of y changes by three times that amount.

In contrast to the explicit form of an equation, the *implicit* form depicts a mutual relation between variables. The function $f(x, y) = 3y - 9x - 18 = 0$ is equivalent to the explicit function in the previous paragraph, i.e., $y = 6 + 3x$. Note that by adding "$9x + 18$" to (or subtracting "$-9x - 18$" from) each side of the equation, then dividing through by 3, we obtain our implicit equation from the explicit one. Indeed, this is what we mean when we say we have "solved" an equation: we have translated an implicit functional form into an explicit form. Note that we could have solved the implicit form to make x a function of y: $3y - 9x - 18 = 0$ becomes $x = (1/3)y - 2$ by adding $9x$ to both sides of the equation, dividing through by 9, and then transposing. An implicit linear equation has the property that either variable can be assigned the role of dependent variable. This is not always the case with nonlinear functions. The implicit function $y - x^2 - 9 = 0$ can be solved for y to yield the function $y = x^2 + 9$, since each value for x is associated with only one value of y. However, if we solved this equation for x, we would generate the relation $x = \pm\sqrt{y - 9}$, which is not a function. For instance, if we set $y = 13$, there are two values of x implied, namely $x = 2$ and $x = -2$.

Let us return to our demand function example: $q = D(p)$, where p is the price of the commodity and q is the quantity demanded. We might write this demand function explicitly as:

(1) $q = 100 - 2p$

We could also write it implicitly as:

(2) $q + 2p - 100 = 0$

By expressing the demand function explicitly as in (1), we not only state that q depends on p, but go further and state the exact way in which q depends on p: q is determined by multiplying p by 2 and subtracting that product from 100. For any value of p, q is uniquely determined. You have probably already recognized equation (1) as a specific example of the "law of demand."

However, if we solved the implicit equation for p, we would obtain the function $p = 50 - 0.5q$, which could represent the average revenue function for an imperfectly competitive seller. As we will see in Part III, the specification of which variable is independent and which is dependent has a profound influence on the nature of the model being tested. Many controversies in economics involve disagreements about the assignment of independent and dependent variables. For example, in his classic article, A. W. Phillips[1] hypothesized that wage changes are a function of the unemployment rate. In his attack on the "inflation-unemployment trade-off," Milton Friedman[2] disputed the argument that the unemployment rate is a function of the rate of inflation (which presumably reflects money wage changes). In the debate, the unemployment variable was transformed from the independent variable to the dependent variable.

[1]A. W. Phillips, "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957," *Economica* (November 1958), pp. 283-299.

[2]Milton Friedman, "Nobel Lecture: Inflation and Unemployment," *Journal of Political Economy*, (June 1977), pp. 451-472.

Functions of One Variable

The notation $y = f(x)$ signifies no more than the proposition that we are taking y as some explicit function of x. More specifically, it states that y is a function of one variable only, x. Both y and x enter the equation, but there is only one independent variable. Hence, we say that this is a function of one variable. There are several types of explicit functions of one variable, and we shall discuss some types that are commonly used in economics.

Linear Functions

A *linear function* is one in which only the first power of the independent variable appears in the equation. The reason for the name "linear" will become clear when we discuss graphs of functions. The equation for this function is sometimes called a first-degree equation. The following equations all represent linear functions:

(1) $y = x$
(2) $y = 2 - x$
(3) $y = 4 + 3x$
(4) $y = a + bx$

Equation (4) is a general statement of a linear function. The letters a and b refer to constants whose values are unspecified. This general form would take on the configuration of (1) if $a = 0$ and $b = 1$; it would take form (2) if $a = 2$ and $b = -1$; and so forth. Equation (4) is called the general form of the equation because a and b represent unknowns rather than stipulated constants.

Polynomial Functions

When the single independent variable is raised to powers that are nonnegative integers, we have a *polynomial function*. A linear function is a polynomial function whose greatest exponent equals 1. A quadratic function is a polynomial function wherein the independent variable is raised to the second power. The following set of equations all represent quadratic functions:

(1) $y = x^2$
(2) $y = 10 + 3x^2$
(3) $y = 2 - 3x + 4x^2$
(4) $y = a + bx + cx^2$

Equation (4) is the general quadratic form with a, b, and c as unspecified constants.

Higher powers of the independent variable may also be encountered in a polynomial function. The cubic function includes at most the third power