

Preface

Econometrics by Example (EBE) is written primarily for undergraduate students in economics, accounting, finance, marketing, and related disciplines. It is also intended for students in MBA programs and for researchers in business, government, and research organizations.

There are several excellent textbooks in econometrics, written from very elementary to very advanced levels. The writers of these books have their intended audiences. I have contributed to this field with my own books, *Basic Econometrics* (McGraw-Hill, 5th edn, 2009) and *Essentials of Econometrics* (McGraw-Hill, 4th edn, 2009). These books have been well received and have been translated into several languages. EBE is different from my own books and those written by others in that it deals with major topics in econometrics from the point of view of their practical applications. Because of space limitations, textbooks generally discuss econometric theory and illustrate econometric techniques with just a few examples. But space does not permit them to deal with concrete examples in detail.

In EBE, each chapter discusses one or two examples in depth. To give but one illustration of this, Chapter 8 discusses binary dummy dependent variable regression models. This specific example relates to the decision to smoke or not to smoke, taking the value of 1 if a person smokes or the value of 0 if he/she does not smoke. The data consists of a random sample of 119 US males. The explanatory variables considered are age, education, income, and price of cigarettes. There are three approaches to modeling this problem: (1) ordinary least-squares (OLS), which leads to the linear probability model (LPM), (2) the logit model, based on the logistic probability distribution, and (3) the probit model, based on the normal distribution.

Which is a better model? In assessing this, we have to consider the pros and cons of all of these three approaches and evaluate the results based on these three competing models and then decide which one to choose. Most textbooks have a theoretical discussion about this, but do not have the space to discuss all the practical aspects of a given problem.

This book is self-contained in that the basic theory underlying each topic is discussed without complicated mathematics. It has an appendix that discusses the basic concepts of statistics in a user-friendly manner and provides the necessary statistical background to follow the concepts covered therein. In EBE all the examples I analyse look at each problem in depth, starting with model formulation, estimation of the chosen model, testing hypotheses about the phenomenon under study, and post-estimation diagnostics to see how well the model performs. Due attention is paid to commonly encountered problems, such as multicollinearity, heteroscedasticity, autocorrelation, model specification errors, and non-stationarity of economic time series. This step-by-step approach, from model formulation, through estimation and

hypothesis-testing, to post-estimation diagnostics will provide a framework for less experienced students and researchers. It will also help them to understand empirical articles in academic and professional journals.

The specific examples discussed in this book are:

- 1 Determination of hourly wages for a group of US workers
- 2 Cobb–Douglas production function for the USA
- 3 The rate of growth of real GDP, USA, 1960–2007
- 4 The relationship between food expenditure and total expenditure
- 5 Log-linear model of real GDP growth
- 6 Gross private investment and gross private savings, USA, 1959–2007
- 7 Quarterly retail fashion sales
- 8 Married women's hours of work
- 9 Abortion rates in the USA
- 10 US consumption function, 1947–2000
- 11 Deaths from lung cancer and the number of cigarettes smoked
- 12 Model of school choice
- 13 Attitude toward working mothers
- 14 Decision to apply to graduate school
- 15 Patents and R&D expenditure. an application of the Poisson probability distribution
- 16 Dollar/euro exchange rates: are they stationary?
- 17 Closing daily prices of IBM stock: are they a random walk?
- 18 Is the regression of consumption expenditure on disposable personal income spurious?
- 19 Are 3-month and 6-month US Treasury Bills cointegrated?
- 20 ARCH model of dollar/euro exchange rate
- 21 GARCH model of dollar/euro exchange rate
- 22 An ARMA model of IBM daily closing prices
- 23 Vector error correction model (VEC) of 3-month and 6-month Treasury Bill rates
- 24 Testing for Granger causality between consumption expenditure and per capita disposable income
- 25 Charitable donations using panel data
- 26 Duration analysis of recidivism
- 27 Instrumental variable estimation of schooling and socio-economic variables
- 28 The simultaneity between consumption expenditure and income

The book is divided into four parts:

Part I discusses the classical linear regression model, which is the workhorse of econometrics. This model is based on restrictive assumptions. The three chapters cover the linear regression model, functional forms of regression models, and qualitative (dummy) variables regression models.

Part II looks critically at the assumptions of the classical linear regression model and examines the ways these assumptions can be modified and with what effect. Specifically, we discuss the topics of multicollinearity, heteroscedasticity, autocorrelation, and model specification errors.

Part III discusses important topics in cross-section econometrics. These chapters discuss and illustrate several cross-sectional topics that are, in fact, not usually discussed in depth in most undergraduate textbooks. These are logit and probit models, multinomial regression models, ordinal regression models, censored and truncated regression models, and Poisson and negative binomial distribution models dealing with count data.

The reason for discussing these models is that they are increasingly being used in the fields of economics, education, psychology, political science, and marketing, largely due to the availability of extensive cross-sectional data involving thousands of observations and also because user-friendly software programs are now readily available to deal with not only vast quantities of data but also to deal with some of these techniques, which are mathematically involved.

Part IV deals primarily with topics in time series econometrics, such as stationary and nonstationary time series, cointegration and error-correction mechanisms, asset price volatility (the ARCH and GARCH models), and economic forecasting with regression (ARIMA and VAR models).

It also discusses three advanced topics. These are panel data regression models (that is, models that deal with repeated cross-sectional data over time; in particular we discuss the fixed effects and random effects models), survival or duration analysis of phenomena such as the duration of unemployment and survival time of cancer patients, and the method of instrumental variables (IV), which is used to deal with stochastic explanatory variables that may be correlated with the error term, which renders OLS estimators inconsistent.

In sum, as the title suggests, *Econometrics by Example* discusses the major themes in econometrics with detailed worked examples that show how the subject works in practice. With some basic theory and familiarity with econometric software, students will find that “learning by doing” is the best way to learn econometrics. The prerequisites are minimal. An exposure to the two-variable linear regression model, a beginning course in statistics, and facility in algebraic manipulations will be adequate to follow the material in the book. EBE does not use any matrix algebra or advanced calculus.

EBE makes heavy use of the *Stata* and *Eviews* statistical packages. The outputs obtained from these packages are reproduced in the book so the reader can see clearly the results in a compact way. Wherever necessary, graphs are produced to give a visual feel for the phenomenon under study. Most of the chapters include several exercises that the reader may want to attempt to learn more about the various techniques discussed. Although the bulk of the book is free of complicated mathematical derivations, in a few cases some advanced material is put in the appendices.

Companion website

The data used in this textbook are posted on the companion website and notes within each chapter direct the reader to this at the relevant points. Students are encouraged to use these data in several end-of-chapter exercises to practice applying what they

have learned to different scenarios. The instructor may also want to use these data for classroom assignments to develop and estimate alternative econometric models. For the instructor, solutions to these end-of-chapter exercises are posted on the companion website in the password protected lecturer zone. Here, (s)he will also find a collection of PowerPoint slides which correspond to each chapter for use in teaching.