Econometric Methods with Applications in Business and Economics

Econometric Methods with Applications in Business and Economics

Christiaan Heij
Paul de Boer
Philip Hans Franses
Teun Kloek
Herman K. van Dijk

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP
Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in
Oxford New York
Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto
Oxford is a registered trade mark of Oxford University Press
in the UK and certain other countries
Published in the United States
by Oxford University Press Inc., New York
© Christiaan Heij, Paul de Boer, Philip Hans Franses, Teun Kloek, and Herman K. van Dijk, 2004
The moral rights of the author have been asserted
Database right Oxford University Press (maker)
First published 2004
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,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above
You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer
British Library Cataloguing in Publication Data
Data available
Library of Congress Cataloging in Publication Data
Data available
ISBN 0–19–926801–0
1 3 5 7 9 10 8 6 4 2
Typeset by Kolam Information Services Pvt. Ltd, Pondicherry, India
Printed in Great Britain
on acid-free paper by Antony Rowe Ltd, Chippenham, Wiltshire

Preface

Econometric models and methods are applied in the daily practice of virtually all disciplines in business and economics like finance, marketing, microeconomics, and macroeconomics. This book is meant for anyone interested in obtaining a solid understanding and active working knowledge of this field. The book provides the reader both with the required insight in econometric methods and with the practical training needed for successful applications. The guiding principle of the book is to stimulate the reader to work actively on examples and exercises, so that econometrics is learnt the way it works in practice — that is, practical methods for solving questions in business and economics, based on a solid understanding of the underlying methods. In this way the reader gets trained to make the proper decisions in econometric modelling.

This book has grown out of half a century of experience in teaching undergraduate econometrics at the Econometric Institute in Rotterdam. With the support of Jan Tinbergen, Henri Theil founded the institute in 1956 and he developed Econometrics into a full-blown academic programme. Originally, econometrics was mostly concerned with national and international macroeconomic policy; the required computing power to estimate econometric models was expensive and scarcely available, so that econometrics was almost exclusively applied in public (statistical) agencies. Much has changed, and nowadays econometrics finds widespread application in a rich variety of fields. The two major causes of this increased role of econometrics are the information explosion in business and economics (with large data sets — for instance, in finance and marketing) and the enormous growth in cheap computing power and user-friendly software for a wide range of econometric methods.

This development is reflected in the book, as it presents econometric methods as a collection of very useful tools to address issues in a wide range of application areas. First of all, students should learn the essentials of econometrics in a rigorous way, as this forms the indispensable basis for all valid practical work. These essentials are treated in Chapters 1–5, after which two major application areas are discussed in Chapter 6 (on individual choice data with applications in marketing and microeconomics) and Chapter 7 (on time series data with applications in finance and international economics). The Introduction provides more information on the motivation and contents of the book, together with advice for students and instructors, and the Guide to the Book explains the structure and use of the book.

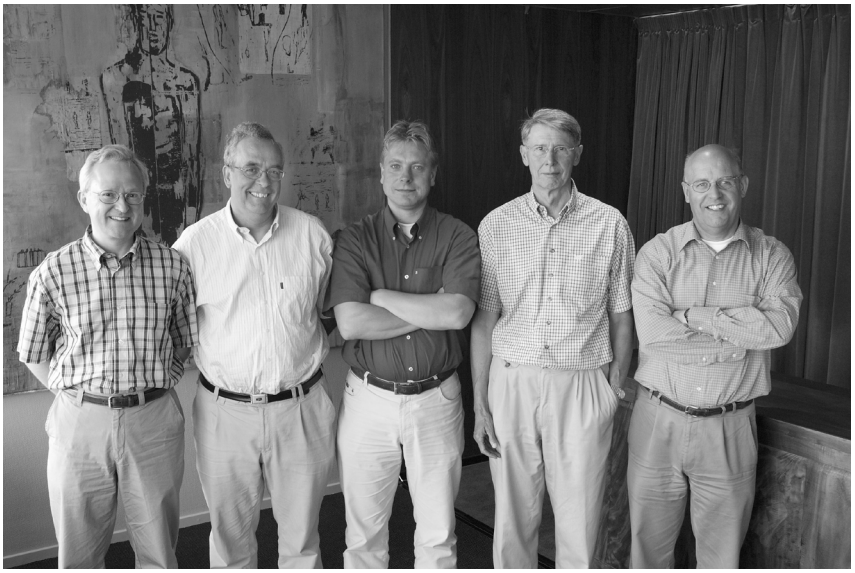
We thank our students, who always stimulate our enthusiasm to teach and who make us feel proud by their achievements in their later careers in econometrics, economics, and business management. We also thank both current and former members of the Econometric Institute in Rotterdam who have inspired our econometric work.

Several people helped us in the process of writing the book and the solutions manual. First of all we should mention our colleague Zsolt Sandor and our (current and former) Ph.D.

students Charles Bos, Lennart Hoogerheide, Rutger van Oest, and Björn Vroomen, who all contributed substantially in producing the solutions manual. Further we thank our (current and former) colleagues at the Econometric Institute, Bas Donkers, Rinse Harkema, Johan Kaashoek, Frank Kleibergen, Richard Kleijn, Peter Kooiman, Marius Ooms, and Peter Schotman. We were assisted by our (former) students Arjan van Dijk, Alex Hoogendoorn, and Jesse de Klerk, and we obtained very helpful feedback from our students, in particular from Simone Jansen, Martijn de Jong, Mariëlle Non, Arnoud Pijls, and Gerard Voskuil. Special thanks are for Aletta Henderiks, who never lost her courage in giving us the necessary secretarial support in processing the manuscript. Finally we wish to thank the delegates and staff of Oxford University Press for their assistance, in particular Andrew Schuller, Arthur Attwell, and Hilary Walford.

Christiaan Heij, Paul de Boer, Philip Hans Franses, Teun Kloek, Herman K. van Dijk

Rotterdam, *2004*



From left to right: Christiaan Heij, Paul de Boer, Philip Hans Franses, Teun Kloek, and Herman K. van Dijk

Contents

Detailed Contents

List of Exhibits

Abbreviations

Apart from abbreviations that are common in econometrics, the list also contains the abbreviations (in italics) used to denote the data sets of examples and exercises, but not the abbreviations used to denote the variables in these data sets (see Appendix B for the meaning of the abbreviated variable names).

2SLS	two-stage least squares
3SLS	three-stage least squares
ACF	autocorrelation function
ADF	augmented Dickey–Fuller
ADL	autoregressive distributed lag
AIC	Akaike information criterion
AR	autoregressive
ARCH	autoregressive conditional heteroskedasticity
ARIMA	autoregressive integrated moving average
ARMA	autoregressive moving average
BHHH	method of Berndt, Hall, Hall, and Hausman
BIC	Bayes information criterion
BLUE	best linear unbiased estimator
BWA	bank wages (data set 2)
CAPM	capital asset pricing model
CAR	car production (data set 18)
CDF	cumulative distribution function
CL	conditional logit
COF	coffee sales (data set 4)
CUSUM	cumulative sum
CUSUMSQ	cumulative sum of squares
DGP	data generating process
DJI	Dow-Jones index (data set 15)
DMF	direct marketing for financial product (data set 13)
DUS	duration of strikes (data set 14)
ECM	error correction model
EWMA	exponentially weighted moving average
EXR	exchange rates (data set 21)
FAS	fashion sales (data set 8)
FEX	food expenditure (data set 7)
FGLS	feasible generalized least squares
FWLS	feasible weighted least squares
GARCH	generalized autoregressive conditional heteroskedasticity
GLS	generalized least squares
GMM	generalized method of moments
GNP	gross national product (data set 20)
HAC	heteroskedasticity and autocorrelation consistent
IBR	interest and bond rates (data set 9)
IID	identically and independently distributed
INP	industrial production (data set 10)
IV	instrumental variable
LAD	least absolute deviation
LM	Lagrange multiplier
LOG	natural logarithm
LR	likelihood ratio
MA	moving average
MAE	mean absolute error
MGC	motor gasoline consumption (data set 6)
ML	maximum likelihood
MNL	multinomial logit
MOM	mortality and marriages (data set 16)
MOR	market for oranges (data set 23)
MSE	mean squared error
NEP	nuclear energy production (data set 19)
NID	normally and independently distributed
NLS	non-linear least squares
OLS	ordinary least squares
P	probability (*P*-value)
PACF	partial autocorrelation function
PMI	primary metal industries (data set 5)
QML	quasi-maximum likelihood
RESET	regression specification error test
RMSE	root mean squared error
SACF	sample autocorrelation function
SCDF	sample cumulative distribution function
SEM	simultaneous equation model
SIC	Schwarz information criterion
SMR	stock market returns (data set 3)

SPACF	sample partial autocorrelation function
SSE	explained sum of squares
SSR	sum of squared residuals
SST	total sum of squares
STAR	smooth transition autoregressive
STP	standard and poor index (data set 22)
STU	student learning (data set 1)
SUR	seemingly unrelated regression
TAR	threshold autoregressive
TBR	Treasury Bill rates (data set 17)
TMSP	total mean squared prediction error
TOP	salaries of top managers (data set 11)
USP	US presidential elections (data set 12)
VAR	vector autoregressive
VECM	vector error correction model
W	Wald
WLS	weighted least squares

Guide to the Book

This guide describes the organization and use of the book. We refer to the Introduction for the purpose of the book, for a synopsis of the contents of the book, for study advice, and for suggestions for instructors as to how the book can be used in different courses.

Learning econometrics: Why, what, and how

The learning student is confronted with three basic questions: Why should I study this? What knowledge do I need? How can I apply this knowledge in practice? Therefore the topics of the book are presented in the following manner:

- explanation by motivating examples;
- discussion of appropriate econometric models and methods;
- illustrative applications in practical examples;
- training by empirical exercises (using an econometric software package);
- optional deeper understanding (theory text parts and theory and simulation exercises).

The book can be used for applied courses that focus on the 'how' of econometrics and also for more advanced courses that treat both the 'how' and the 'what' of econometrics. The user is free to choose the desired balance between econometric applications and econometric theory.

- In applied courses, the theory parts (clearly marked in the text) and the theory and simulation exercises can be skipped without any harm. Even without these parts, the text still provides a good understanding of the 'what' of econometrics that is required in sound applied work, as there exist no standard 'how-to-do' recipes that can be applied blindly in practice.
- In more advanced courses, students get a deeper understanding of econometrics — in addition to the practical skills of applied courses — by studying also the theory parts and by doing the theory and simulation exercises. This allows them to apply econometrics in new situations that require a creative mind in developing alternative models and methods.

Text structure

The required background material is covered in Chapter 1 (which reviews statistical methods that are fundamental in econometrics) and in Appendix A (which summarizes useful matrix methods, together with computational examples). The core material on econometrics is in Chapters 2–7; Chapters 2–5 treat fundamental econometric methods that are needed for the topics discussed in Chapters 6 and 7. Each chapter has the following structure.

- The chapter starts with a brief statement of the purpose of the chapter, followed by sections and subsections that are divided into manageable parts with clear headings.
- Examples, theory parts, and computational schemes are clearly indicated in the text.
- Summaries are included at many points — especially at the end of all sections in Chapters 5–7.
- The chapter concludes with a brief summary, further reading, and a keyword list that summarizes the treated topics.
- A varied set of exercises is included at the end of each chapter.

To facilitate the use of the book, the required preliminary knowledge is indicated at the start of subsections.

- In Chapters 2–4 we refer to the preliminary knowledge needed from Chapter 1 (on statistics) and Appendix A (on matrix methods). Therefore, it is not necessary to cover all Chapter 1 before starting on the other chapters, as Chapter 1 can be reviewed along the way as one progresses through Chapters 2–4, and the same holds true for the material of Appendix A.
- In Chapters 6 and 7 we indicate which parts of the earlier chapters are needed at each stage. Most of the sections of Chapter 5 can be read independently of each other, and in Chapters 6 and 7 some sections can be skipped depending on the topics of interest for the reader.
- Further details of the text structure are discussed in the Introduction (see the section 'Teaching suggestions' — in particular, Exhibit 0.3).

Examples and data sets

The econometric models and methods are motivated by means of fully worked-out examples using real-world data sets from a variety of applications in business and economics. The examples are clearly marked in the text because they play a crucial role in explaining the application of econometric methods.

The corresponding data sets are available from the web site of the book, and Appendix B explains the type and source of the data and the meaning of the variables in the data files (see p. 748 for a list of all the data sets used in the book). The names of the data sets consist of three parts:

- XM (for examples) and XR (for exercises);
- three digits, indicating the example or exercise number;
- three letters, indicating the data topic.

For example, the file XM101STU contains the data for Example 1.1 on student learning, and the file XR111STU contains the data for Exercise 1.11 on student learning.

Exercises

Students will enhance their understanding and acquire practical skills by working through the exercises, which are of three types.

- *Theory exercises on derivations and model extensions*. These exercises deepen the theoretical understanding of the 'what' of econometrics. The desired level of the course will determine how many of the theory exercises should be covered.
- *Simulation exercises illustrating statistical properties of econometric models and methods*. These exercises provide more intuitive understanding of some of the central theoretical results.
- *Empirical exercises on applications with business and economic data sets to solve questions of practical interest*. These exercises focus on the 'how' of econometrics, so that the student learns to construct appropriate models from real-world data and to draw sound conclusions from the obtained results. Actively working through these empirical exercises is essential to gaining a proper understanding of econometrics and to getting hands-on experience with applications to solve practical problems. The web site of the book contains the data sets of all empirical exercises, and Appendix B contains information on these data sets.

The choice of appropriate exercises is facilitated by cross-references.

- Each subsection concludes with a list of exercises related to the material of that subsection (where T denotes theory exercises, S simulation exercises, and E empirical exercises).
- Every exercise refers to the parts of the chapter that are needed for doing the exercise.
- An asterisk (∗) denotes advanced (parts of) exercises.

Web site and software

The web site of the book contains all the data sets used in the book, in three formats:

- EViews;
- Excel;
- ASCII.

All the examples and all the empirical and simulation exercises in the book can be done with EViews version 3.1 and higher (Quantitative Micro Software, 1994–8), but other econometric software packages can also be used in most cases. The student version of the EViews package suffices for most of the book, but this version has some limitations — for example, it does not support the programs required for the simulation exercises (see the web site of the book for further details). The exhibits for the empirical examples in the text have been obtained by using EViews version 3.1.

Instructor material

Instructors who adopt the book can receive the Solutions Manual of the book for free.

- The manual contains over 350 pages with fully worked-out text solutions of all exercises, both of the theory questions and of the empirical and simulation questions; this will assist instructors in selecting material for exercise sessions and computer sessions as part of their course.
- The manual contains a CD-ROM with solution files (EViews work files with the solutions of all empirical exercises and EViews programs for all simulation exercises).
- This CD-ROM also contains all the exhibits of the book (in Word format) to facilitate lecture presentations.

The printed solutions manual and CD-ROM can be obtained from Oxford University Press, upon request by adopting instructors. For further information and additional material we refer readers to the Oxford University Press web site of the book.

Remarks on notation

In the text we follow the notational conventions commonly used in econometrics.

- Scalar variables and vectors are denoted by lower-case italic letters (x, y, and so on); however, in Section 7.6 vectors of variables are denoted by upper-case italic letters, such as Y_t, in accordance with most of the literature on this topic.
- Matrices are denoted by upper-case italic letters (X, A, and so on).

- The element in row i and column j of a matrix A is generally denoted by a_{ij}, except for the regressor matrix X, where this element is denoted by x_{ji}, which is observation i of variable j (see Section 3.1.2).
- x_i denotes the vector containing the values of all the explanatory variables x_{ji} for observation i (including the value 1 as first element of x_i if the model contains a constant term).
- Transposition is denoted by a prime (X', x', and so on).
- Unknown parameters are denoted by Greek italic letters (β, ε, σ, and so on).
- Estimated quantities are denoted by Latin italic letters (b, e, s, and so on), or sometimes by imposing a hat ($\hat{\beta}$, $\hat{\varepsilon}$, $\hat{\sigma}$, and so on).
- Expected values are denoted by $E[\cdot]$ — for instance, $E[b]$.
- $\log(x)$ denotes the natural logarithm of x (with base $e = 2.71828\ \ldots$).

In many of the exhibits — for instance, the ones related to empirical examples — we show the output as generated by the software program EViews. The notation in these exhibits may differ from the above conventions.

- Scalar variables are denoted by capital letters (X, Y, instead of x, y, and so on).
- Statistics are denoted by text (R-squared, Std. Dev., instead of R^2, s, and so on).

In most cases this does not lead to any confusion, and otherwise the notation is explained in the text or in the caption of the exhibits.