

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.

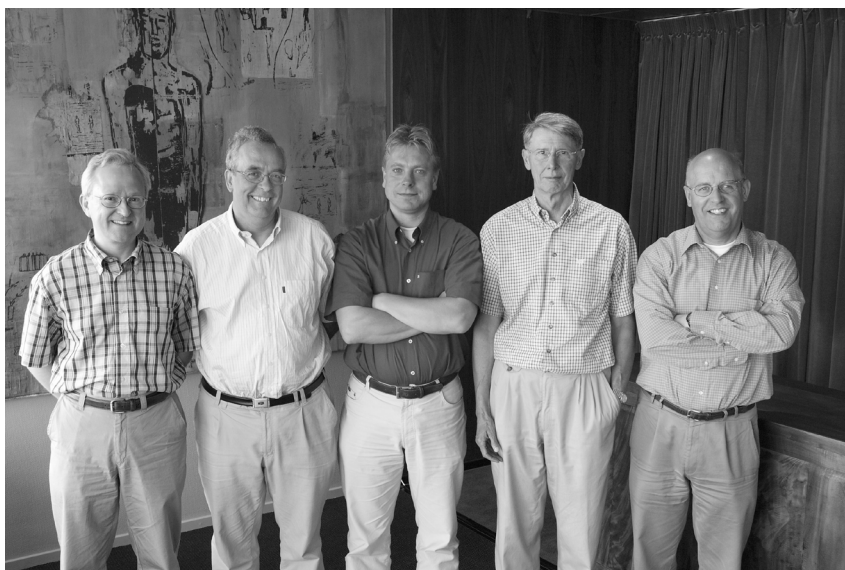
Several people helped us in the process of writing the book and the solutions manual. First of all we should mention our colleague Zsolt Sandor and our (current and former) Ph.D.

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

<i>Detailed Contents</i>	ix
<i>List of Exhibits</i>	xvi
<i>Abbreviations</i>	xix
<i>Guide to the Book</i>	xxi
Introduction	1
1 Review of Statistics	11
2 Simple Regression	75
3 Multiple Regression	117
4 Non-Linear Methods	187
5 Diagnostic Tests and Model Adjustments	273
6 Qualitative and Limited Dependent Variables	437
7 Time Series and Dynamic Models	531
Appendix A. Matrix Methods	723
Appendix B. Data Sets	747
<i>Index</i>	773

Detailed Contents

<i>List of Exhibits</i>	xvi
<i>Abbreviations</i>	xix
<i>Guide to the Book</i>	xxi

Introduction **1**

Econometrics	1
Purpose of the book	2
Characteristic features of the book	3
Target audience and required background knowledge	4
Brief contents of the book	4
Study advice	5
Teaching suggestions	6
Some possible course structures	8

1 Review of Statistics **11**

1.1 Descriptive statistics	12
1.1.1 Data graphs	12
1.1.2 Sample statistics	16
1.2 Random variables	20
1.2.1 Single random variables	20
1.2.2 Joint random variables	23
1.2.3 Probability distributions	29
1.2.4 Normal random samples	35
1.3 Parameter estimation	38
1.3.1 Estimation methods	38
1.3.2 Statistical properties	42
1.3.3 Asymptotic properties	47
1.4 Tests of hypotheses	55
1.4.1 Size and power	55
1.4.2 Tests for mean and variance	59
1.4.3 Interval estimates and the bootstrap	63

Summary, further reading, and keywords	68
Exercises	71

2 Simple Regression **75**

2.1 Least squares	76
2.1.1 Scatter diagrams	76
2.1.2 Least squares	79
2.1.3 Residuals and R^2	82
2.1.4 Illustration: Bank Wages	84
2.2 Accuracy of least squares	87
2.2.1 Data generating processes	87
2.2.2 Examples of regression models	91
2.2.3 Seven assumptions	92
2.2.4 Statistical properties	94
2.2.5 Efficiency	97
2.3 Significance tests	99
2.3.1 The t -test	99
2.3.2 Examples	101
2.3.3 Use under less strict conditions	103
2.4 Prediction	105
2.4.1 Point predictions and prediction intervals	105
2.4.2 Examples	107
Summary, further reading, and keywords	111
Exercises	113

3 Multiple Regression **117**

3.1 Least squares in matrix form	118
3.1.1 Introduction	118
3.1.2 Least squares	120
3.1.3 Geometric interpretation	123
3.1.4 Statistical properties	125
3.1.5 Estimating the disturbance variance	127
3.1.6 Coefficient of determination	129
3.1.7 Illustration: Bank Wages	131
3.2 Adding or deleting variables	134
3.2.1 Restricted and unrestricted models	135
3.2.2 Interpretation of regression coefficients	139
3.2.3 Omitting variables	142
3.2.4 Consequences of redundant variables	143
3.2.5 Partial regression	145

3.3 The accuracy of estimates	152
3.3.1 The t -test	152
3.3.2 Illustration: Bank Wages	154
3.3.3 Multicollinearity	156
3.3.4 Illustration: Bank Wages	159
3.4 The F -test	161
3.4.1 The F -test in different forms	161
3.4.2 Illustration: Bank Wages	166
3.4.3 Chow forecast test	169
3.4.4 Illustration: Bank Wages	174
Summary, further reading, and keywords	178
Exercises	180
<hr/>	
4 Non-Linear Methods	187
<hr/>	
4.1 Asymptotic analysis	188
4.1.1 Introduction	188
4.1.2 Stochastic regressors	191
4.1.3 Consistency	193
4.1.4 Asymptotic normality	196
4.1.5 Simulation examples	198
4.2 Non-linear regression	202
4.2.1 Motivation	202
4.2.2 Non-linear least squares	205
4.2.3 Non-linear optimization	209
4.2.4 The Lagrange Multiplier test	212
4.2.5 Illustration: Coffee Sales	218
4.3 Maximum likelihood	222
4.3.1 Motivation	222
4.3.2 Maximum likelihood estimation	224
4.3.3 Asymptotic properties	228
4.3.4 The Likelihood Ratio test	230
4.3.5 The Wald test	232
4.3.6 The Lagrange Multiplier test	235
4.3.7 LM -test in the linear model	238
4.3.8 Remarks on tests	240
4.3.9 Two examples	243
4.4 Generalized method of moments	250
4.4.1 Motivation	250
4.4.2 GMM estimation	252
4.4.3 GMM standard errors	255
4.4.4 Quasi-maximum likelihood	259

4.4.5 GMM in simple regression	260
4.4.6 Illustration: Stock Market Returns	262
Summary, further reading, and keywords	266
Exercises	268

5 Diagnostic Tests and Model Adjustments **273**

5.1 Introduction	274
5.2 Functional form and explanatory variables	277
5.2.1 The number of explanatory variables	277
5.2.2 Non-linear functional forms	285
5.2.3 Non-parametric estimation	289
5.2.4 Data transformations	296
5.2.5 Summary	302
5.3 Varying parameters	303
5.3.1 The use of dummy variables	303
5.3.2 Recursive least squares	310
5.3.3 Tests for varying parameters	313
5.3.4 Summary	318
5.4 Heteroskedasticity	320
5.4.1 Introduction	320
5.4.2 Properties of OLS and White standard errors	324
5.4.3 Weighted least squares	327
5.4.4 Estimation by maximum likelihood and feasible WLS	334
5.4.5 Tests for homoskedasticity	343
5.4.6 Summary	352
5.5 Serial correlation	354
5.5.1 Introduction	354
5.5.2 Properties of OLS	358
5.5.3 Tests for serial correlation	361
5.5.4 Model adjustments	368
5.5.5 Summary	376
5.6 Disturbance distribution	378
5.6.1 Introduction	378
5.6.2 Regression diagnostics	379
5.6.3 Test for normality	386
5.6.4 Robust estimation	388
5.6.5 Summary	394
5.7 Endogenous regressors and instrumental variables	396
5.7.1 Instrumental variables and two-stage least squares	396
5.7.2 Statistical properties of IV estimators	404
5.7.3 Tests for exogeneity and validity of instruments	409
5.7.4 Summary	418

5.8 Illustration: Salaries of top managers	419
Summary, further reading, and keywords	424
Exercises	427

6 Qualitative and Limited Dependent Variables 437

6.1 Binary response	438
6.1.1 Model formulation	438
6.1.2 Probit and logit models	443
6.1.3 Estimation and evaluation	447
6.1.4 Diagnostics	452
6.1.5 Model for grouped data	459
6.1.6 Summary	461
6.2 Multinomial data	463
6.2.1 Unordered response	463
6.2.2 Multinomial and conditional logit	466
6.2.3 Ordered response	474
6.2.4 Summary	480
6.3 Limited dependent variables	482
6.3.1 Truncated samples	482
6.3.2 Censored data	490
6.3.3 Models for selection and treatment effects	500
6.3.4 Duration models	511
6.3.5 Summary	521
Summary, further reading, and keywords	523
Exercises	525

7 Time Series and Dynamic Models 531

7.1 Models for stationary time series	532
7.1.1 Introduction	532
7.1.2 Stationary processes	535
7.1.3 Autoregressive models	538
7.1.4 ARMA models	542
7.1.5 Autocorrelations and partial autocorrelations	545
7.1.6 Forecasting	550
7.1.7 Summary	553
7.2 Model estimation and selection	555
7.2.1 The modelling process	555
7.2.2 Parameter estimation	558
7.2.3 Model selection	563

7.2.4 Diagnostic tests	567
7.2.5 Summary	576
7.3 Trends and seasonals	578
7.3.1 Trend models	578
7.3.2 Trend estimation and forecasting	585
7.3.3 Unit root tests	592
7.3.4 Seasonality	604
7.3.5 Summary	611
7.4 Non-linearities and time-varying volatility	612
7.4.1 Outliers	612
7.4.2 Time-varying parameters	616
7.4.3 GARCH models for clustered volatility	620
7.4.4 Estimation and diagnostic tests of GARCH models	626
7.4.5 Summary	636
7.5 Regression models with lags	637
7.5.1 Autoregressive models with distributed lags	637
7.5.2 Estimation, testing, and forecasting	640
7.5.3 Regression of variables with trends	647
7.5.4 Summary	654
7.6 Vector autoregressive models	656
7.6.1 Stationary vector autoregressions	656
7.6.2 Estimation and diagnostic tests of stationary VAR models	661
7.6.3 Trends and cointegration	667
7.6.4 Summary	681
7.7 Other multiple equation models	682
7.7.1 Introduction	682
7.7.2 Seemingly unrelated regression model	684
7.7.3 Panel data	692
7.7.4 Simultaneous equation model	700
7.7.5 Summary	709
Summary, further reading, and keywords	710
Exercises	713

Appendix A. Matrix Methods **723**

A.1 Summations	723
A.2 Vectors and matrices	725
A.3 Matrix addition and multiplication	727
A.4 Transpose, trace, and inverse	729
A.5 Determinant, rank, and eigenvalues	731
A.6 Positive (semi)definite matrices and projections	736
A.7 Optimization of a function of several variables	738

A.8 Concentration and the Lagrange method	743
Exercise	746

Appendix B. Data Sets **747**

List of Data Sets	748
-------------------	-----

Index **773**

List of Exhibits

0.1	Econometrics as an interdisciplinary field	2	3.1	Scatter diagrams of Bank Wage data	119
0.2	Econometric modelling	3	3.2	Least squares	124
0.3	Book structure	8	3.3	Least squares	124
			3.4	Geometric picture of R^2	130
1.1	Student Learning (Example 1.1)	13	3.5	Bank Wages (Section 3.1.7)	132
1.2	Student Learning (Example 1.2)	14	3.6	Direct and indirect effects	134
1.3	Student Learning (Example 1.3)	15	3.7	Bank Wages (Example 3.1)	138
1.4	Student Learning (Example 1.4)	17	3.8	Bank Wages (Example 3.2)	141
1.5	Student Learning (Example 1.5)	19	3.9	Bias and efficiency	145
1.6	Student Learning (Example 1.6)	28	3.10	Bank Wages (Example 3.3)	149
1.7	Normal distribution	30	3.11	Bank Wages (Example 3.3)	150
1.8	χ^2 -distribution	32	3.12	Bank Wages (Section 3.3.2)	155
1.9	t -distribution	33	3.13	Bank Wages (Section 3.3.4)	160
1.10	F -distribution	34	3.14	P -value	162
1.11	Bias and variance	44	3.15	Geometry of F -test	163
1.12	Normal Random Sample (Example 1.9)	47	3.16	Bank Wages (Section 3.4.2)	167
1.13	Consistency	48	3.17	Prediction	170
1.14	Simulated Normal Random Sample (Example 1.10)	54	3.18	Bank Wages (Section 3.4.4)	175
1.15	Simulated Normal Random Sample (Example 1.11)	58	3.19	Bank Wages (Section 3.4.4)	176
1.16	P -value	61	4.1	Bank Wages (Example 4.1)	190
1.17	Student Learning (Example 1.13)	66	4.2	Inconsistency	195
			4.3	Simulation Example (Section 4.1.5)	198
2.1	Stock Market Returns (Example 2.1)	77	4.4	Simulation Example (Section 4.1.5)	200
2.2	Bank Wages (Example 2.2)	78	4.5	Coffee Sales (Example 4.2)	203
2.3	Coffee Sales (Example 2.3)	79	4.6	Food Expenditure (Example 4.3)	204
2.4	Scatter diagram with fitted line	80	4.7	Newton–Raphson	210
2.5	Bank Wages (Section 2.1.4)	85	4.8	Lagrange multiplier	214
2.6	Bank Wages (Section 2.1.4)	86	4.9	Coffee Sales (Section 4.2.5)	219
2.7	Simulated Regression Data (Example 2.4)	89	4.10	Coffee Sales (Section 4.2.5)	220
2.8	Simulated Regression Data (Example 2.4)	90	4.11	Stock Market Returns (Example 4.4)	224
2.9	Accuracy of least squares	97	4.12	Maximum likelihood	225
2.10	Simulated Regression Data (Example 2.8)	101	4.13	Likelihood Ratio test	231
2.11	Bank Wages (Example 2.9)	103	4.14	Wald test	233
2.12	Quantiles of distributions of t -statistics	103	4.15	Lagrange Multiplier test	236
2.13	Prediction error	106	4.16	Comparison of tests	241
2.14	Simulated Regression Data (Example 2.10)	108	4.17	F - and χ^2 -distributions	242
2.15	Bank Wages (Example 2.11)	109	4.18	Stock Market Returns (Example 4.5)	245
			4.19	Coffee Sales (Example 4.6)	248

4.20	Stock Market Returns (Example 4.7)	251	5.45	Motor Gasoline Consumption (Example 5.31)	404
4.21	Stock Market Returns (Section 4.4.6)	264	5.46	Interest and Bond Rates (Example 5.32)	408
5.1	The empirical cycle in econometric modelling	276	5.47	Interest and Bond Rates (Example 5.33)	415
5.2	Bank Wages (Example 5.1)	283	5.48	Motor Gasoline Consumption (Example 5.34)	417
5.3	Bank Wages (Example 5.1)	284	5.49	Salaries of Top Managers (Example 5.35)	420
5.4	Bank Wages (Example 5.2)	287	6.1	Probability models	440
5.5	Tricube weights	291	6.2	Normal and logistic densities	444
5.6	Simulated Data from a Non-Linear Model (Example 5.3)	294	6.3	Direct Marketing for Financial Product (Example 6.2)	451
5.7	Bank Wages (Example 5.4)	295	6.4	Direct Marketing for Financial Product (Example 6.3)	458
5.8	Bank Wages (Example 5.5)	299	6.5	Bank Wages (Example 6.4)	472
5.9	Bank Wages (Example 5.5)	300	6.6	Ordered response	476
5.10	Fashion Sales (Example 5.6)	306	6.7	Bank Wages (Example 6.5)	478
5.11	Fashion Sales (Example 5.6)	307	6.8	Truncated data	484
5.12	Coffee Sales (Example 5.7)	309	6.9	Direct Marketing for Financial Product (Example 6.6)	489
5.13	Bank Wages (Example 5.8)	312	6.10	Censored data	491
5.14	Bank Wages (Example 5.9)	317	6.11	Direct Marketing for Financial Product (Example 6.7)	498
5.15	Bank Wages (Example 5.9)	318	6.12	Student Learning (Example 6.8)	508
5.16	Bank Wages (Example 5.10)	322	6.13	Student Learning (Example 6.8)	509
5.17	Interest and Bond Rates (Example 5.11)	323	6.14	Duration data	512
5.18	Bank Wages; Interest and Bond Rates (Example 5.12)	326	6.15	Hazard rates	514
5.19	Bank Wages (Example 5.13)	331	6.16	Duration of Strikes (Example 6.9)	517
5.20	Bank Wages (Example 5.13)	332	7.1	Industrial Production (Example 7.1)	533
5.21	Interest and Bond Rates (Example 5.14)	333	7.2	Dow-Jones Index (Example 7.2)	534
5.22	Bank Wages (Example 5.15)	338	7.3	Simulated AR Time Series (Example 7.3)	542
5.23	Interest and Bond Rates (Example 5.16)	341	7.4	Simulated MA and ARMA Time Series (Example 7.4)	545
5.24	Bank Wages (Example 5.17)	347	7.5	Simulated Times Series (Example 7.5)	549
5.25	Bank Wages (Example 5.17)	349	7.6	Simulated Time Series (Example 7.6)	553
5.26	Interest and Bond Rates (Example 5.18)	350	7.7	Steps in modelling	555
5.27	Interest and Bond Rates (Example 5.18)	351	7.8	Industrial Production (Example 7.7)	557
5.28	Interest and Bond Rates (Example 5.19)	355	7.9	Industrial Production (Example 7.8)	563
5.29	Food Expenditure (Example 5.20)	358	7.10	Industrial Production (Example 7.10)	566
5.30	Interest and Bond Rates (Example 5.21)	360	7.11	Industrial Production (Example 7.11)	573
5.31	Interest and Bond Rates (Example 5.22)	366	7.12	Industrial Production (Example 7.11)	575
5.32	Food Expenditure (Example 5.23)	367	7.13	Simulated Series with Trends (Example 7.12)	584
5.33	Interest and Bond Rates (Example 5.24)	371	7.14	Industrial Production (Example 7.13)	591
5.34	Food Expenditure (Example 5.25)	373	7.15	Industrial Production (Example 7.13)	592
5.35	Industrial Production (Example 5.26)	375	7.16	Unit root tests	595
5.36	Industrial Production (Example 5.26)	376	7.17	Industrial Production (Example 7.14)	601
5.37	Outliers and OLS	382	7.18	Dow-Jones Index (Example 7.15)	603
5.38	Stock Market Returns (Example 5.27)	385	7.19	Industrial Production (Example 7.16)	609
5.39	Stock Market Returns (Example 5.27)	386	7.20	Industrial Production (Example 7.16)	610
5.40	Stock Market Returns (Example 5.28)	388	7.21	Industrial Production (Example 7.17)	615
5.41	Simulated Data of Normal and Student $t(2)$ Distribution (Example 5.29)	389			
5.42	Three estimation criteria	391			
5.43	Interest and Bond Rates (Example 5.30)	401			
5.44	Motor Gasoline Consumption (Example 5.31)	403			

7.22	Industrial Production (Example 7.18)	619	7.34	Primary Metal Industries (Example 7.29)	689
7.23	Simulated ARCH and GARCH Time Series (Example 7.19)	624	7.35	Primary Metal Industries (Example 7.29)	690
7.24	Industrial Production (Example 7.20)	630	7.36	Primary Metal Industries (Example 7.30)	698
7.25	Dow-Jones Index (Example 7.21)	632	7.37	Simulated Macroeconomic Consumption and Income (Example 7.31)	703
7.26	Interest and Bond Rates (Example 7.22)	645	7.38	Interest and Bond Rates (Example 7.32)	708
7.27	Mortality and Marriages (Example 7.23)	648			
7.28	Simulated Random Walk Data (Example 7.24)	650	A.1	Simulated Data on Student Learning (Example A.1)	724
7.29	Interest and Bond Rates (Example 7.25)	653	A.2	Simulated Data on Student Learning (Example A.7)	735
7.30	Interest and Bond Rates (Example 7.26)	665	A.3	Simulated Data on Student Learning (Example A.11)	742
7.31	Cointegration tests	672			
7.32	Interest and Bond Rates (Example 7.27)	676			
7.33	Treasury Bill Rates (Example 7.28)	679			

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

xx Abbreviations

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

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 $e = 2.71828 \dots$).

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.