Introductory Econometrics for Finance

This bestselling and thoroughly classroom-tested textbook is a complete resource for finance students. A comprehensive and illustrated discussion of the most common empirical approaches in finance prepares students for using econometrics in practice, while detailed case studies help them understand how the techniques are used in relevant financial contexts. Learning outcomes, key concepts and end-of-chapter review questions (with full solutions online) highlight the main chapter takeaways and allow students to self-assess their understanding. Building on the successful data- and problem-driven approach of previous editions, this fourth edition has been updated with new examples, additional introductory material on mathematics and dealing with data, as well as more advanced material on extreme value theory, the generalised method of moments and state space models. A dedicated website, with numerous student and instructor resources including videos and a set of companion manuals for various statistical software – all available free of charge – completes the learning package.

Chris Brooks is Professor of Finance at the ICMA Centre, Henley Business School, University of Reading, UK where he also obtained his PhD. Chris has diverse research interests and has published over a hundred articles in leading academic and practitioner journals, and six books. He is Associate Editor of several journals, including the *Journal of Business Finance and Accounting* and the *British Accounting Review*. He acts as consultant and advisor for various banks, corporations and regulatory and professional bodies in the fields of finance, real estate and econometrics.

Introductory Econometrics for Finance

FOURTH EDITION

CHRIS BROOKS
The ICMA Centre, Henley Business School, University of Reading

CAMBRIDGE UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108422536
DOI: 10.1017/9781108524872

First published 2002
Second edition published 2008
Third edition published 2014
Fourth edition published 2019

Printed and bound in Great Britain by Clays Ltd, Elcograf S.p.A.

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Names: Brooks, Chris, 1971– author.
Title: Introductory econometrics for finance / Chris Brooks, The ICMA Centre, Henley Business School, University of Reading.
Description: Fourth edition. | Cambridge, United Kingdom ; New York, NY : Cambridge University Press, 2019. | Includes bibliographical references and index.
Identifiers: LCCN 2018061692 | ISBN 9781108422536 (hardback : alk. paper) | ISBN 9781108436823 (pbk. : alk. paper)
Subjects: LCSH: Finance–Econometric models. | Econometrics.
Classification: LCC HG173 .B76 2019 | DDC 332.01/5195–dc23
LC record available at https://lccn.loc.gov/2018061692

ISBN 978-1-108-42253-6 Hardback
ISBN 978-1-108-43682-3 Paperback

Additional resources for this publication at www.cambridge.org/brooks4

Contents in Brief

List of Figures *page* xiii
List of Tables xvi
List of Boxes xix
List of Screenshots xxi
Preface to the Fourth Edition xxiii
Acknowledgements xxvii
Outline of the Remainder of this Book xxviii

Chapter 1 Introduction and Mathematical Foundations 1
Chapter 2 Statistical Foundations and Dealing with Data 41
Chapter 3 A Brief Overview of the Classical Linear Regression Model 94
Chapter 4 Further Development and Analysis of the Classical Linear Regression Model 146
Chapter 5 Classical Linear Regression Model Assumptions and Diagnostic Tests 182
Chapter 6 Univariate Time-Series Modelling and Forecasting 246
Chapter 7 Multivariate Models 293
Chapter 8 Modelling Long-Run Relationships in Finance 334
Chapter 9 Modelling Volatility and Correlation 384
Chapter 10 Switching and State Space Models 447
Chapter 11 Panel Data 490
Chapter 12 Limited Dependent Variable Models 516
Chapter 13 Simulation Methods 546
Chapter 14 Additional Econometric Techniques for Financial Research 571
Chapter 15 Conducting Empirical Research or Doing a Project or Dissertation in Finance 617

Appendix 1 Sources of Data Used in This Book and the Accompanying Software Manuals 632
Appendix 2 Tables of Statistical Distributions 633

Glossary 646
References 672
Index 688

Detailed Contents

List of Figures *page* xiii
List of Tables xvi
List of Boxes xix
List of Screenshots xxi
Preface to the Fourth Edition xxiii
Acknowledgements xxvii
Outline of the Remainder of this Book xxviii

Chapter 1 Introduction and Mathematical Foundations 1

1.1 What is Econometrics? 2
1.2 Is Financial Econometrics Different? 3
1.3 Steps Involved in Formulating an Econometric Model 4
1.4 Points to Consider When Reading Articles 6
1.5 Functions 7
1.6 Differential Calculus 19
1.7 Matrices 28

Chapter 2 Statistical Foundations and Dealing with Data 41

2.1 Probability and Probability Distributions 41
2.2 A Note on Bayesian versus Classical Statistics 47
2.3 Descriptive Statistics 48
2.4 Types of Data and Data Aggregation 63
2.5 Arithmetic and Geometric Series 67
2.6 Future Values and Present Values 68
2.7 Returns in Financial Modelling 77
2.8 Portfolio Theory Using Matrix Algebra 82

Chapter 3 A Brief Overview of the Classical Linear Regression Model 94

3.1 What is a Regression Model? 94
3.2 Regression versus Correlation 95
3.3 Simple Regression 95
3.4 Some Further Terminology 103

3.5 The Assumptions Underlying the Model 106
3.6 Properties of the OLS Estimator 107
3.7 Precision and Standard Errors 110
3.8 An Introduction to Statistical Inference 115
3.9 A Special Type of Hypothesis Test 129
3.10 An Example of a Simple *t*-test of a Theory 131
3.11 Can UK Unit Trust Managers Beat the Market? 133
3.12 The Overreaction Hypothesis 134
3.13 The Exact Significance Level 138
Appendix 3.1 Mathematical Derivations of CLRM Results 139

Chapter 4 Further Development and Analysis of the Classical Linear Regression Model 146

4.1 Generalising the Simple Model 146
4.2 The Constant Term 147
4.3 How are the Parameters Calculated? 149
4.4 Testing Multiple Hypotheses: The *F*-test 150
4.5 Data Mining and the True Size of the Test 157
4.6 Qualitative Variables 158
4.7 Goodness of Fit Statistics 159
4.8 Hedonic Pricing Models 163
4.9 Tests of Non-Nested Hypotheses 167
4.10 Quantile Regression 168
Appendix 4.1 Mathematical Derivations of CLRM Results 173
Appendix 4.2 A Brief Introduction to Factor Models and Principal Components Analysis 175

Chapter 5 Classical Linear Regression Model Assumptions and Diagnostic Tests 182

5.1 Introduction 182
5.2 Statistical Distributions for Diagnostic Tests 183
5.3 Assumption (1): $E(u_t) = 0$ 184
5.4 Assumption (2): $\text{var}(u_t) = \sigma^2 < \infty$ 185
5.5 Assumption (3): $\text{cov}(u_i, u_j) = 0$ for $i \neq j$ 189
5.6 Assumption (4): The x_t are Non-Stochastic 208
5.7 Assumption (5): The Disturbances are Normally Distributed 209
5.8 Multicollinearity 213
5.9 Adopting the Wrong Functional Form 217
5.10 Omission of an Important Variable 221
5.11 Inclusion of an Irrelevant Variable 221
5.12 Parameter Stability Tests 222
5.13 Measurement Errors 230

5.14 A Strategy for Constructing Econometric Models 231
5.15 Determinants of Sovereign Credit Ratings 234

Chapter 6 Univariate Time-Series Modelling and Forecasting 246

6.1 Introduction 246
6.2 Some Notation and Concepts 247
6.3 Moving Average Processes 251
6.4 Autoregressive Processes 254
6.5 The Partial Autocorrelation Function 262
6.6 ARMA Processes 263
6.7 Building ARMA Models: The Box–Jenkins Approach 269
6.8 Examples of Time-Series Modelling in Finance 272
6.9 Exponential Smoothing 274
6.10 Forecasting in Econometrics 277

Chapter 7 Multivariate Models 293

7.1 Motivations 293
7.2 Simultaneous Equations Bias 295
7.3 So how can Simultaneous Equations Models be Validly Estimated? 297
7.4 Can the Original Coefficients be Retrieved from the πs? 297
7.5 Simultaneous Equations in Finance 299
7.6 A Definition of Exogeneity 300
7.7 Triangular Systems 303
7.8 Estimation Procedures for Simultaneous Equations Systems 304
7.9 An Application of a Simultaneous Equations Approach 307
7.10 Vector Autoregressive Models 312
7.11 Does the VAR Include Contemporaneous Terms? 318
7.12 Block Significance and Causality Tests 319
7.13 VARs with Exogenous Variables 322
7.14 Impulse Responses and Variance Decompositions 322
7.15 VAR Model Example: The Interaction Between Property Returns and the Macroeconomy 325
7.16 A Couple of Final Points on VARs 331

Chapter 8 Modelling Long-Run Relationships in Finance 334

8.1 Stationarity and Unit Root Testing 334
8.2 Tests for Unit Roots in the Presence of Structural Breaks 346
8.3 Cointegration 351
8.4 Equilibrium Correction or Error Correction Models 353
8.5 Testing for Cointegration in Regression: A Residuals-Based Approach 355

8.6 Methods of Parameter Estimation in Cointegrated Systems 356
8.7 Lead–Lag and Long-Term Relationships Between Spot and Futures Markets 358
8.8 Testing for and Estimating Cointegration in Systems 365
8.9 Purchasing Power Parity 370
8.10 Cointegration Between International Bond Markets 371
8.11 Testing the Expectations Hypothesis of the Term Structure of Interest Rates 377

Chapter 9 Modelling Volatility and Correlation 384

9.1 Motivations: An Excursion into Non-Linearity Land 384
9.2 Models for Volatility 389
9.3 Historical Volatility 389
9.4 Implied Volatility Models 390
9.5 Exponentially Weighted Moving Average Models 390
9.6 Autoregressive Volatility Models 391
9.7 Autoregressive Conditionally Heteroscedastic (ARCH) Models 392
9.8 Generalised ARCH (GARCH) Models 396
9.9 Estimation of ARCH/GARCH Models 399
9.10 Extensions to the Basic GARCH Model 404
9.11 Asymmetric GARCH Models 404
9.12 The GJR model 405
9.13 The EGARCH Model 405
9.14 Tests for Asymmetries in Volatility 406
9.15 GARCH-in-Mean 408
9.16 Uses of GARCH-Type Models 408
9.17 Testing Non-Linear Restrictions 411
9.18 Volatility Forecasting: Some Examples and Results 414
9.19 Stochastic Volatility Models Revisited 419
9.20 Forecasting Covariances and Correlations 423
9.21 Covariance Modelling and Forecasting in Finance 424
9.22 Simple Covariance Models 426
9.23 Multivariate GARCH Models 427
9.24 Direct Correlation Models 431
9.25 Extensions to the Basic Multivariate GARCH Model 433
9.26 A Multivariate GARCH Model for the CAPM 434
9.27 Estimating a Time-Varying Hedge Ratio 435
9.28 Multivariate Stochastic Volatility Models 439
Appendix 9.1 Parameter Estimation Using Maximum Likelihood 440

Chapter 10 Switching and State Space Models 447

10.1 Motivations 447
10.2 Seasonalities in Financial Markets 449
10.3 Modelling Seasonality in Financial Data 450
10.4 Estimating Simple Piecewise Linear Functions 458
10.5 Markov Switching Models 459
10.6 A Markov Switching Model for the Real Exchange Rate 462
10.7 A Markov Switching Model for the Gilt–Equity Yield Ratio 464
10.8 Threshold Autoregressive Models 468
10.9 Estimation of Threshold Autoregressive Models 470
10.10 Specification Tests 471
10.11 A SETAR Model for the French franc–German mark Exchange Rate 472
10.12 Threshold Models for FTSE Spot and Futures 474
10.13 Regime Switching Models and Forecasting 477
10.14 State Space Models and the Kalman Filter 477

Chapter 11 Panel Data 490

11.1 Introduction: What Are Panel Techniques? 490
11.2 What Panel Techniques Are Available? 491
11.3 The Fixed Effects Model 493
11.4 Time-Fixed Effects Models 495
11.5 Investigating Banking Competition 496
11.6 The Random Effects Model 500
11.7 Panel Data Application to Credit Stability of Banks 501
11.8 Panel Unit Root and Cointegration Tests 505
11.9 Further Feading 514

Chapter 12 Limited Dependent Variable Models 516

12.1 Introduction and Motivation 516
12.2 The Linear Probability Model 517
12.3 The Logit Model 519
12.4 Using a Logit to Test the Pecking Order Hypothesis 520
12.5 The Probit Model 522
12.6 Choosing Between the Logit and Probit Models 522
12.7 Estimation of Limited Dependent Variable Models 522
12.8 Goodness of Fit Measures for Linear Dependent Variable Models 523
12.9 Multinomial Linear Dependent Variables 526
12.10 The Pecking Order Hypothesis Revisited 529
12.11 Ordered Response Linear Dependent Variables Models 530
12.12 Are Unsolicited Credit Ratings Biased Downwards? An Ordered Probit Analysis 532

12.13 Censored and Truncated Dependent Variables 537
Appendix 12.1 The Maximum Likelihood Estimator for Logit and Probit Models 544

Chapter 13 Simulation Methods 546

13.1 Motivations 546
13.2 Monte Carlo Simulations 547
13.3 Variance Reduction Techniques 548
13.4 Bootstrapping 552
13.5 Random Number Generation 556
13.6 Disadvantages of the Simulation Approach 557
13.7 An Example of Monte Carlo Simulation 558
13.8 An Example of how to Simulate the Price of a Financial Option 558
13.9 An Example of Bootstrapping to Calculate Capital Risk Requirements 561

Chapter 14 Additional Econometric Techniques for Financial Research 571

14.1 Event Studies 571
14.2 Tests of the CAPM and the Fama–French Methodology 586
14.3 Extreme Value Theory 592
14.4 The Generalised Method of Moments 607

Chapter 15 Conducting Empirical Research or Doing a Project or Dissertation in Finance 617

15.1 What is an Empirical Research Project? 617
15.2 Selecting the Topic 618
15.3 Sponsored or Independent Research? 620
15.4 The Research Proposal 622
15.5 Working Papers and Literature on the Internet 623
15.6 Getting the Data 623
15.7 Choice of Computer Software 625
15.8 Methodology 626
15.9 How Might the Finished Project Look? 626
15.10 Presentational Issues 631

Appendix 1 Sources of Data Used in This Book and the Accompanying Software Manuals 632

Appendix 2 Tables of Statistical Distributions 633

Glossary 646
References 672
Index 688

Figures

1.1	Steps involved in formulating an econometric model	*page* 5
1.2	A plot of hours studied (x) against grade-point average (y)	9
1.3	Examples of different straight line graphs	9
1.4	Example of a general polynomial function	10
1.5	Examples of quadratic functions	12
1.6	A plot of an exponential function	14
1.7	A plot of a logarithmic function	16
1.8	The tangents to a curve	20
1.9	$y = f(x)$, its first derivative and its second derivative around the point $x = -6$	24
2.1	The probability distribution function for the sum of two dice	42
2.2	The pdf for a normal distribution	44
2.3	The cdf for a normal distribution	45
2.4	A normal versus a skewed distribution	56
2.5	A normal versus a leptokurtic distribution	57
2.6	A time-series plot and scatter plot of the performance of two fund managers	60
3.1	Scatter plot of two variables, y and x	96
3.2	Scatter plot of two variables with a line of best fit chosen by eye	97
3.3	Method of OLS fitting a line to the data by minimising the sum of squared residuals	98
3.4	Plot of a single observation, together with the line of best fit, the residual and the fitted value	99
3.5	How RSS varies with different values of β	101
3.6	Scatter plot of excess returns on fund XXX versus excess returns on the market portfolio	102
3.7	No observations close to the y-axis	104
3.8	The bias versus variance trade-off when selecting between estimators	109
3.9	Effect on the standard errors of the coefficient estimates when $(x_t - \bar{x})$ are narrowly dispersed	112
3.10	Effect on the standard errors of the coefficient estimates when $(x_t - \bar{x})$ are widely dispersed	113
3.11	Effect on the standard errors of x_t^2 large	113

3.12	Effect on the standard errors of x_t^2 small	114
3.13	The t-distribution versus the normal	119
3.14	Rejection regions for a two-sided 5% hypothesis test	121
3.15	Rejection region for a one-sided hypothesis test of the form H_0:$\beta = \beta^*$, H_1:$\beta < \beta^*$	121
3.16	Rejection region for a one-sided hypothesis test of the form H_0:$\beta = \beta^*$, H_1:$\beta > \beta^*$	122
3.17	Critical values and rejection regions for a $t_{20;5\%}$	125
3.18	Frequency distribution of t-ratios of mutual fund alphas (gross of transactions costs)	132
3.19	Frequency distribution of t-ratios of mutual fund alphas (net of transactions costs)	132
3.20	Performance of UK unit trusts, 1979–2000	134
4.1	$R^2 = 0$ demonstrated by a flat estimated line, i.e., a zero slope coefficient	161
4.2	$R^2 = 1$ when all data points lie exactly on the estimated line	162
5.1	Effect of no intercept on a regression line	184
5.2	Graphical illustration of heteroscedasticity	185
5.3	Plot of $\hat{u}_t$ against $\hat{u}_{t-1}$, showing positive autocorrelation	192
5.4	Plot of $\hat{u}_t$ over time, showing positive autocorrelation	192
5.5	Plot of $\hat{u}_t$ against $\hat{u}_{t-1}$, showing negative autocorrelation	193
5.6	Plot of $\hat{u}_t$ over time, showing negative autocorrelation	193
5.7	Plot of $\hat{u}_t$ against $\hat{u}_{t-1}$, showing no autocorrelation	194
5.8	Plot of $\hat{u}_t$ over time, showing no autocorrelation	194
5.9	Rejection and non-rejection regions for DW test	197
5.10	Regression residuals from stock return data, showing large outlier for October 1987	211
5.11	Possible effect of an outlier on OLS estimation	212
5.12	Relationship between y and x_2 in a quadratic regression for different values of β_2 and β_3	218
5.13	Plot of a variable showing suggestion for break date	228
6.1	Autocorrelation function for sample MA(2) process	255
6.2	Sample autocorrelation and partial autocorrelation functions for an MA(1) model: $y_t = -0.5u_{t-1} + u_t$	265
6.3	Sample autocorrelation and partial autocorrelation functions for an MA(2) model: $y_t = 0.5u_{t-1} - 0.25u_{t-2} + u_t$	266
6.4	Sample autocorrelation and partial autocorrelation functions for a slowly decaying AR(1) model: $y_t = 0.9y_{t-1} + u_t$	266
6.5	Sample autocorrelation and partial autocorrelation functions for a more rapidly decaying AR(1) model: $y_t = 0.5y_{t-1} + u_t$	267
6.6	Sample autocorrelation and partial autocorrelation functions for a more rapidly decaying AR(1) model with negative coefficient: $y_t = -0.5y_{t-1} + u_t$	267

6.7	Sample autocorrelation and partial autocorrelation functions for a non-stationary model (i.e., a unit coefficient): $y_t = y_{t-1} + u_t$	268
6.8	Sample autocorrelation and partial autocorrelation functions for an ARMA(1, 1) model: $y_t = 0.5y_{t-1} + 0.5u_{t-1} + u_t$	269
6.9	Use of in-sample and out-of-sample periods for analysis	278
7.1	Impulse responses and standard error bands for innovations in unexpected inflation equation errors	330
7.2	Impulse responses and standard error bands for innovations in the dividend yields	330
8.1	Value of R^2 for 1000 sets of regressions of a non-stationary variable on another independent non-stationary variable	335
8.2	Value of t-ratio of slope coefficient for 1,000 sets of regressions of a non-stationary variable on another independent non-stationary variable	336
8.3	Example of a white noise process	340
8.4	Time-series plot of a random walk versus a random walk with drift	340
8.5	Time-series plot of a deterministic trend process	341
8.6	Autoregressive processes with differing values of ϕ (0, 0.8, 1)	341
9.1	Daily S&P returns for August 2003–July 2018	393
9.2	The problem of local optima in maximum likelihood estimation	401
9.3	News impact curves for S&P500 returns using coefficients implied from GARCH and GJR model estimates	407
9.4	Three approaches to hypothesis testing under maximum likelihood	412
9.5	Time-varying hedge ratios derived from symmetric and asymmetric *BEKK* models for FTSE returns	438
10.1	Sample time-series plot illustrating a regime shift	448
10.2	Use of intercept dummy variables for quarterly data	452
10.3	Use of slope dummy variables	455
10.4	Piecewise linear model with threshold x^*	459
10.5	Unconditional distribution of US GEYR together with a normal distribution with the same mean and variance	465
10.6	Value of GEYR and probability that it is in the high GEYR regime for the UK	467
12.1	The fatal flaw of the linear probability model	518
12.2	The logit model	519
12.3	Modelling charitable donations as a function of income	537
14.1	Pdfs for the Weibull, Gumbel and Frechét distributions	595

Tables

1.1	Sample data on hours of study and grades	*page* 8
2.1	Annual performance of two funds	59
2.2	Impact of different compounding frequencies on the effective interest rate and terminal value of an investment	71
2.3	How to construct a series in real terms from a nominal one	81
3.1	Sample data on fund XXX to motivate OLS estimation	102
3.2	Critical values from the standard normal versus t-distribution	119
3.3	Classifying hypothesis testing errors and correct conclusions	128
3.4	Summary statistics for the estimated regression results for equation (3.34)	131
3.5	Summary statistics for unit trust returns, January 1979–May 2000	133
3.6	CAPM regression results for unit trust returns, January 1979–May 2000	134
3.7	Is there an overreaction effect in the UK stock market?	137
4.1	Hedonic model of rental values in Quebec City, 1990. Dependent variable: Canadian dollars per month	165
4.2	OLS and quantile regression results for the Magellan fund	172
4A.1	Principal component ordered eigenvalues for Dutch interest rates, 1962–70	178
4A.2	Factor loadings of the first and second principal components for Dutch interest rates, 1962–70	179
5.1	Constructing a series of lagged values and first differences	191
5.2	Determinants and impacts of sovereign credit ratings	237
5.3	Do ratings add to public information?	239
5.4	What determines reactions to ratings announcements?	241
6.1	Uncovered interest parity test results	275
6.2	Forecast error aggregation	284
7.1	Call bid–ask spread and trading volume regression	310
7.2	Put bid–ask spread and trading volume regression	311
7.3	Granger causality tests and implied restrictions on VAR models	321
7.4	Marginal significance levels associated with joint F-tests	328
7.5	Variance decompositions for the property sector index residuals	329
8.1	Critical values for DF tests (Fuller, 1976, p. 373)	344
8.2	Recursive unit root tests for interest rates allowing for structural breaks	350
8.3	DF tests on log-prices and returns for high frequency FTSE data	360

8.4	Estimated potentially cointegrating equation and test for cointegration for high frequency FTSE data	360
8.5	Estimated error correction model for high frequency FTSE data	361
8.6	Comparison of out-of-sample forecasting accuracy	362
8.7	Trading profitability of the error correction model with cost of carry	363
8.8	Cointegration tests of PPP with European data	371
8.9	DF tests for international bond indices	372
8.10	Cointegration tests for pairs of international bond indices	373
8.11	Johansen tests for cointegration between international bond yields	374
8.12	Variance decompositions for VAR of international bond yields	375
8.13	Impulse responses for VAR of international bond yields	376
8.14	Tests of the expectations hypothesis using the US zero coupon yield curve with monthly data	379
9.1	GARCH versus implied volatility	417
9.2	EGARCH versus implied volatility	418
9.3	Out-of-sample predictive power for weekly volatility forecasts	420
9.4	Comparisons of the relative information content of out-of-sample volatility forecasts	421
9.5	Hedging effectiveness: summary statistics for portfolio returns	437
10.1	Values and significances of days of the week coefficients	454
10.2	Day-of-the-week effects with the inclusion of interactive dummy variables with the risk proxy	457
10.3	Estimates of the Markov switching model for real exchange rates	463
10.4	Estimated parameters for the Markov switching models	466
10.5	SETAR model for FRF–DEM	473
10.6	FRF–DEM forecast accuracies	474
10.7	Linear AR(3) model for the basis	476
10.8	A two-threshold SETAR model for the basis	478
10.9	Unit trust performance with time-varying beta estimation	486
11.1	Tests of banking market equilibrium with fixed effects panel models	498
11.2	Tests of competition in banking with fixed effects panel models	499
11.3	Results of random effects panel regression for credit stability of Central and East European banks	504
11.4	Panel unit root test results for economic growth and financial development	512
11.5	Panel cointegration test results for economic growth and financial development	513
12.1	Logit estimation of the probability of external financing	521
12.2	Multinomial logit estimation of the type of external financing	531
12.3	Ordered probit model results for the determinants of credit ratings	534
12.4	Two-step ordered probit model allowing for selectivity bias in the determinants of credit ratings	536
13.1	EGARCH estimates for currency futures returns	563

13.2 Autoregressive volatility estimates for currency futures returns 565
13.3 Minimum capital risk requirements for currency futures as a percentage of the initial value of the position 567
14.1 Fama and MacBeth's results on testing the CAPM 589
14.2 Threshold percentage returns, corresponding empirical quantiles and the number of exceedences 604
14.3 Maximum likelihood estimates of the parameters of the generalised Pareto distribution 605
14.4 Models that predict the actual left tail quantile most accurately 606
14.5 GMM estimates of the effect of stock markets and bank lending on economic growth 614
15.1 Journals in finance and econometrics 621
15.2 Useful internet sites for financial literature 624
15.3 Suggested structure for a typical dissertation or project 627
A2.1 Normal critical values for different values of α 633
A2.2 Critical values of Student's t-distribution for different probability levels, α and degrees of freedom, ν 634
A2.3 Upper 5% critical values for F-distribution 635
A2.4 Upper 1% critical values for F-distribution 636
A2.5 Chi-squared critical values for different values of α and degrees of freedom, υ 637
A2.6 Lower and upper 1% critical values for the Durbin–Watson statistic 639
A2.7 Dickey–Fuller critical values for different significance levels, α 641
A2.8 Critical values for the Engle–Granger cointegration test on regression residuals with no constant in test regression 642
A2.9 Quantiles of the asymptotic distribution of the Johansen cointegration rank test statistics (constant in cointegrating vectors only) 643
A2.10 Quantiles of the asymptotic distribution of the Johansen cointegration rank test statistics (constant, i.e., a drift only in VAR and in cointegrating vector) 644
A2.11 Quantiles of the asymptotic distribution of the Johansen cointegration rank test statistics (constant in cointegrating vector and VAR, trend in cointegrating vector) 645

Boxes

1.1	Examples of the uses of econometrics	*page* 2
1.2	Points to consider when reading a published paper	6
1.3	The roots of a quadratic equation	11
1.4	Manipulating powers and their indices	13
1.5	The laws of logs	15
2.1	The population and the sample	48
2.2	Time-series data	64
2.3	Log returns	78
3.1	Names for y and xs in regression models	95
3.2	Reasons for the inclusion of the disturbance term	97
3.3	Assumptions concerning disturbance terms and their interpretation	107
3.4	Standard error estimators	112
3.5	Conducting a test of significance	120
3.6	Carrying out a hypothesis test using confidence intervals	124
3.7	The test of significance and confidence interval approaches compared	125
3.8	Type I and type II errors	128
3.9	Reasons for stock market overreactions	135
3.10	Ranking stocks and forming portfolios	136
3.11	Portfolio monitoring	136
4.1	The relationship between the regression F-statistic and R^2	166
4.2	Selecting between models	168
5.1	Conducting White's test	187
5.2	'Solutions' for Heteroscedasticity	190
5.3	Conditions for DW to be a valid test	198
5.4	Conducting a Breusch–Godfrey test	199
5.5	The Cochrane–Orcutt procedure	201
5.6	Observations for the dummy variable	211
5.7	Conducting a Chow test	223
6.1	The stationarity condition for an AR(p) model	256
6.2	The invertibility condition for an MA(2) model	263
6.3	Naive forecasting methods	280
7.1	Determining whether an equation is identified	298

7.2	Conducting a Hausman test for exogeneity	301
7.3	Forecasting with VARs	320
8.1	Stationarity tests	347
8.2	Multiple cointegrating relationships	358
9.1	Testing for 'ARCH effects'	395
9.2	Estimating an ARCH or GARCH model	399
9.3	Using maximum likelihood estimation in practice	402
10.1	How do dummy variables work?	452
10.2	Parameter estimation using the Kalman filter	483
11.1	Fixed or random effects?	502
12.1	Parameter interpretation for probit and logit models	524
12.2	The differences between censored and truncated dependent variables	538
13.1	Conducting a Monte Carlo simulation	548
13.2	Re-sampling the data	554
13.3	Re-sampling from the residuals	555
13.4	Setting up a Monte Carlo simulation	559
13.5	Simulating the price of an option	560
13.6	Generating draws from a GARCH process	560
14.1	The three generalised extreme value distributions	596
15.1	Possible types of research project	619

Screenshots

2.1 Setting up a variance–covariance matrix in Excel *page* 86
2.2 The spreadsheet for constructing the efficient frontier 87
2.3 Completing the Solver window 88
2.4 A plot of the completed efficient frontier 89
2.5 The capital market line and efficient frontier 90

Preface to the Fourth Edition

All of the motivations for the first edition, described below, seem just as important today. Given that the book seems to have gone down well with readers, I have left the style largely unaltered but added a lot of new material. The main motivations for writing the first edition of the book were:

- To write a book that focused on *using and applying* the techniques rather than deriving proofs and learning formulae.
- To write an accessible textbook that required no prior knowledge of econometrics, but which also covered more recently developed approaches usually only found in more advanced texts.
- To use examples and terminology from finance rather than economics since there are many introductory texts in econometrics aimed at students of economics but none for students of finance.
- To populate the book with case studies of the use of econometrics in practice taken from the academic finance literature.
- To include sample instructions, screen dumps and computer output from a popular econometrics package. This enabled readers to see how the techniques can be implemented in practice. In this fourth edition, the EViews instructions have been separated off and are available free of charge on the book's web site along with parallel manuals for other packages including Stata, Python and R.
- To develop a companion web site containing answers to end of chapter questions, a multiple choice question bank with feedback, PowerPoint slides and other supporting materials.

What is New in the Fourth Edition

The fourth edition includes a number of important new features

(1) Students of finance have enormously varying backgrounds, and in particular varying levels of training in elementary mathematics and statistics. In order to make the book more self-contained, the introductory chapter has again been expanded. So the material previously in Chapter 2 has been separated into introductory maths (Chapter 1) and introductory statistics/dealing with data (Chapter 2).

(2) More new material has been added on state space models and their estimation using the Kalman filter in Chapter 10.

(3) A chapter has been added which collects together a number of techniques often used in financial research, including event studies and the Fama MacBeth approach (previously elsewhere in the book) and new sections on using extreme value distribution to model the fat tails in financial series and on estimating models with the generalised method of moments.

(4) The incorporation of EViews directly into the core of the book may have been a distraction for those using other packages. Thus, as stated above, in the new edition the EViews instructions have been separated off and are available free of charge on the book's web site along with parallel manuals for other packages including Stata, Python and R. This package should ensure that the book fits the bill whatever the reader's preferred software.

Motivations for the First Edition

This book had its genesis in two sets of lectures given annually by the author at the ICMA Centre (formerly the ISMA Centre), Henley Business School, University of Reading and arose partly from several years of frustration at the lack of an appropriate textbook. In the past, finance was but a small sub-discipline drawn from economics and accounting, and therefore it was generally safe to assume that students of finance were well grounded in economic principles; econometrics would be taught using economic motivations and examples.

However, finance as a subject has taken on a life of its own in recent years. Drawn in by perceptions of exciting careers in the financial markets, the number of students of finance has grown phenomenally all around the world. At the same time, the diversity of educational backgrounds of students taking finance courses has also expanded. It is not uncommon to find undergraduate students of finance even without advanced high-school qualifications in mathematics or economics. Conversely, many with PhDs in physics or engineering are also attracted to study finance at the Masters level. Unfortunately, authors of textbooks failed to keep pace with the change in the nature of students. In my opinion, the currently available textbooks fall short of the requirements of this market in three main regards, which this book seeks to address

(1) Books fall into two distinct and non-overlapping categories: the introductory and the advanced. Introductory textbooks are at the appropriate level for students with limited backgrounds in mathematics or statistics, but their focus is too narrow. They often spend too long deriving the most basic results, and treatment of important, interesting and relevant topics (such as simulations methods, VAR modelling, etc.) is covered in only the last few pages, if at all. The more advanced textbooks, meanwhile, usually require a quantum leap

in the level of mathematical ability assumed of readers, so that such books cannot be used on courses lasting only one or two semesters, or where students have differing backgrounds. In this book, I have tried to sweep a broad brush over a large number of different econometric techniques that are relevant to the analysis of financial and other data.

(2) Many of the currently available textbooks with broad coverage are too theoretical in nature and students can often, after reading such a book, still have no idea of how to tackle real-world problems themselves, even if they have mastered the techniques in theory. This book and the accompanying software manuals should assist students who wish to learn how to estimate models for themselves – for example, if they are required to complete a project or dissertation. Some examples have been developed especially for this book, while many others are drawn from the academic finance literature. In my opinion, this is an essential but rare feature of a textbook that should help to show students how econometrics is really applied. It is also hoped that this approach will encourage some students to delve deeper into the literature, and will give useful pointers and stimulate ideas for research projects. It should, however, be stated at the outset that the purpose of including examples from the academic finance print is not to provide a comprehensive overview of the literature or to discuss all of the relevant work in those areas, but rather to illustrate the techniques. Therefore, the literature reviews may be considered deliberately deficient, with interested readers directed to the suggested readings and the references therein.

(3) With few exceptions, almost all textbooks that are aimed at the introductory level draw their motivations and examples from economics, which may be of limited interest to students of finance or business. To see this, try motivating regression relationships using an example such as the effect of changes in income on consumption and watch your audience, who are primarily interested in business and finance applications, slip away and lose interest in the first ten minutes of your course.

Who Should Read this Book?

The intended audience is undergraduates or Masters/MBA and PhD students who require a broad knowledge of modern econometric techniques commonly employed in the finance literature. It is hoped that the book will also be useful for researchers (both academics and practitioners), who require an introduction to the statistical tools commonly employed in the area of finance. The book can be used for courses covering financial time-series analysis or financial econometrics in undergraduate or post-graduate programmes in finance, financial economics, securities and investments.

Although the applications and motivations for model-building given in the book are drawn from finance, the empirical testing of theories in many other disciplines,

such as management studies, business studies, real estate, economics and so on, may usefully employ econometric analysis. For this group, the book may also prove useful.

Finally, while the present text is designed mainly for students at the undergraduate or Masters level, it could also provide introductory reading in financial modelling for finance doctoral programmes where students have backgrounds which do not include courses in modern econometric techniques.

Pre-Requisites for Good Understanding of This Material

In order to make the book as accessible as possible, no prior knowledge of statistics, econometrics or algebra is required, although those with a prior exposure to calculus, algebra (including matrices) and basic statistics will be able to progress more quickly. The emphasis throughout the book is on a valid application of the techniques to real data and problems in finance.

In the finance and investment area, it is assumed that the reader has knowledge of the fundamentals of corporate finance, financial markets and investment. Therefore, subjects such as portfolio theory, the capital asset pricing model (CAPM) and arbitrage pricing theory (APT), the efficient markets hypothesis, the pricing of derivative securities and the term structure of interest rates, which are frequently referred to throughout the book, are not explained from first principles in this text. There are very many good books available in corporate finance, in investments and in futures and options, including those by Brealey and Myers (2013), Bodie, Kane and Marcus (2014) and Hull (2017) respectively.

Acknowledgements

I am grateful to Gita Persand, Olan Henry, James Chong and Apostolos Katsaris, who assisted with various parts of the software applications for the first edition. I am also grateful to Hilary Feltham for assistance with Chapters 1 and 2. I would also like to thank Simon Burke, James Chong and Con Keating for detailed and constructive comments on various drafts of the first edition, Simon Burke for suggestions on parts of the second edition, Mike Clements, Jo Cox, Eunyoung Mallet, Ogonna Nneji, Ioannis Oikonomou and Chardan Wese Simen for comments on part of the third edition and Marcel Prokopczuk for comments on part of the fourth edition.

I have additionally benefited from the comments, suggestions and questions of the following list of people, many of whom sent useful e-mails pointing out typos or inaccuracies: Zary Aftab, Panos Ballis-Papanastasiou, Mirco Balatti, Peter Burridge, Kyongwook Choi, Rishi Chopra, Araceli Ortega Diaz, Xiaoming Ding, Thomas Eilertsen, Waleid Eldien, Junjong Eo, Merlyn Foo, Andrea Gheno, Christopher Gilbert, Kimon Gomozias, Jan de Gooijer and his colleagues, Cherif Guermat, Abid Hameed, Ibrahim Jamali, Kejia Jia, Arty Khemlani, Margaret Lynch, David McCaffrey, Tehri Jokipii, Emese Lazar, Zhao Liuyan, Dimitri Lvov, Bill McCabe, Junshi Ma, Raffaele Mancuso, David Merchan, Yue Min, Victor Murinde, Kyoung Gook Park, Mikael Petitjean, Marcelo Perlin, Thai Pham, Jean-Sebastien Pourchet, Marcel Prokopczuk, Tao Qingmei, Satya Sahoo, Lisa Schopohl, Guilherme Silva, Jerry Sin, Andre-Tudor Stancu, Silvia Stanescu, Fred Sterbenz, Birgit Strikholm, Yiguo Sun, Li Qui, Panagiotis Varlagas, Jakub Vojtek, Henk von Eije, Jue Wang, Robert Wichmann and Meng-Feng Yen.

The publisher and author have used their best endeavours to ensure that the URLs for external web sites referred to in this book are correct and active at the time of going to press. However, the publisher and author have no responsibility for the web sites and can make no guarantee that a site will remain live or that the content is or will remain appropriate.

Outline of the Remainder of this Book

Chapter 1

This covers the key mathematical techniques that readers will need some familiarity with to be able to get the most out of the remainder of this book. It starts with a discussion of what econometrics is about and how to set up an econometric model, then moves on to present the mathematical material on functions, and powers, exponents and logarithms of numbers. It then proceeds to explain the basics of differentiation and matrix algebra, which is illustrated via the construction of optimal portfolio weights.

Chapter 2

This chapter presents the statistical foundations of econometrics and the beginnings of how to work with financial data. It covers key results in statistics, discusses probability distributions, how to summarise data and different types of data. The chapter then moves on to discuss the calculation of present and future values, compounding and discounting, and how to calculate nominal and real returns in various ways.

Chapter 3

This introduces the classical linear regression model (CLRM). The ordinary least squares (OLS) estimator is derived and its interpretation discussed. The conditions for OLS optimality are stated and explained. A hypothesis testing framework is developed and examined in the context of the linear model. Examples employed include Jensen's classic study of mutual fund performance measurement and tests of the 'overreaction hypothesis' in the context of the UK stock market.

Chapter 4

This continues and develops the material of Chapter 3 by generalising the bivariate model to multiple regression – i.e., models with many variables. The framework for testing multiple hypotheses is outlined, and measures of how well the model fits the data are described. Case studies include modelling rental values and an application of principal components analysis (PCA) to interest rates.

Chapter 5

Chapter 5 examines the important but often neglected topic of diagnostic testing. The consequences of violations of the CLRM assumptions are described, along with plausible remedial steps. Model-building philosophies are discussed, with particular reference to the general-to-specific approach. Applications covered in this chapter include the determination of sovereign credit ratings.

Chapter 6

This presents an introduction to time-series models, including their motivation and a description of the characteristics of financial data that they can and cannot capture. The chapter commences with a presentation of the features of some standard models of stochastic (white noise, moving average, autoregressive and mixed ARMA) processes. The chapter continues by showing how the appropriate model can be chosen for a set of actual data, how the model is estimated and how model adequacy checks are performed. The generation of forecasts from such models is discussed, as are the criteria by which these forecasts can be evaluated. Examples include model-building for UK house prices, and tests of the exchange rate covered and uncovered interest parity hypotheses.

Chapter 7

This extends the analysis from univariate to multivariate models. Multivariate models are motivated by way of explanation of the possible existence of bi-directional causality in financial relationships, and the simultaneous equations bias that results if this is ignored. Estimation techniques for simultaneous equations models are outlined. Vector autoregressive (VAR) models, which have become extremely popular in the empirical finance literature, are also covered. The interpretation of VARs is explained by way of joint tests of restrictions, causality tests, impulse responses and variance decompositions. Relevant examples discussed in this chapter are the simultaneous relationship between bid–ask spreads and trading volume in the context of options pricing, and the relationship between property returns and macroeconomic variables.

Chapter 8

The first section of the chapter discusses unit root processes and presents tests for non-stationarity in time-series. The concept of and tests for cointegration, and the formulation of error correction models, are then discussed in the context of both the single equation framework of Engle–Granger, and the multivariate framework of Johansen. Applications studied in Chapter 8 include spot and futures markets, tests for cointegration between international bond markets and tests of the purchasing power parity (PPP) hypothesis and of the expectations hypothesis of the term structure of interest rates.

Chapter 9

This covers the important topic of volatility and correlation modelling and forecasting. This chapter starts by discussing in general terms the issue of non-linearity in financial time series. The class of ARCH (autoregressive conditionally heteroscedastic) models and the motivation for this formulation are then discussed. Other models are also presented, including extensions of the basic model such as GARCH, GARCH-M, EGARCH and GJR formulations. Examples of the huge number of applications are discussed, with particular reference to stock returns. Multivariate GARCH and conditional correlation models are described, and applications to the estimation of conditional betas and time-varying hedge ratios, and to financial risk measurement, are given.

Chapter 10

This begins by discussing how to test for and model regime shifts or switches of behaviour in financial series that can arise from changes in government policy, market trading conditions or microstructure, among other causes. This chapter then introduces the Markov switching approach to dealing with regime shifts. Threshold autoregression is also discussed, along with issues relating to the estimation of such models. Examples include the modelling of exchange rates within a managed floating environment, modelling and forecasting the gilt–equity yield ratio and models of movements of the difference between spot and futures prices. Finally, the second part of the chapter moves on to examine how to specify models with time-varying parameters using the state space form and how to estimate them with the Kalman filter.

Chapter 11

This chapter focuses on how to deal appropriately with longitudinal data – that is, data having both time-series and cross-sectional dimensions. Fixed effect and random effect models are explained and illustrated by way of examples on banking competition in the UK and on credit stability in Central and Eastern Europe. Entity fixed and time-fixed effects models are elucidated and distinguished.

Chapter 12

This chapter describes various models that are appropriate for situations where the dependent variable is not continuous. Readers will learn how to construct, estimate and interpret such models, and to distinguish and select between alternative specifications. Examples used include a test of the pecking order hypothesis in corporate finance and the modelling of unsolicited credit ratings.

Chapter 13

This presents an introduction to the use of simulations in econometrics and finance. Motivations are given for the use of repeated sampling, and a distinction is drawn

between Monte Carlo simulation and bootstrapping. The reader is shown how to set up a simulation, and examples are given in options pricing and financial risk management to demonstrate the usefulness of these techniques.

Chapter 14

This chapter presents a collection of techniques that are particularly useful for conducting research in finance. It begins with detailed illustrations of how to conduct event studies, which are commonly used in corporate finance applications, and how to use the Fama-French factor model approach to asset pricing. The chapter then proceeds to present the families of extreme value models that are used to accurately capture the fat tails of asset return distributions and as the basis for value at risk calculations. Finally, the chapter covers the generalised method of moments (GMM) technique, which has become increasingly popular in recent years for estimating a range of different types of models in finance.

Chapter 15

This offers suggestions related to conducting a project or dissertation in empirical finance. It introduces the sources of financial and economic data available on the internet and elsewhere, and recommends relevant online information and literature on research in financial markets and financial time series. The chapter also suggests ideas for what might constitute a good structure for a dissertation on this subject, how to generate ideas for a suitable topic, what format the report could take, and some common pitfalls.