

Contents

<i>List of figures</i>	<i>page</i> xii
<i>List of tables</i>	xiv
<i>List of boxes</i>	xvi
<i>List of screenshots</i>	xvii
<i>Preface to the second edition</i>	xix
<i>Acknowledgements</i>	xxiv
1 Introduction	1
1.1 What is econometrics?	1
1.2 Is financial econometrics different from ‘economic econometrics’?	2
1.3 Types of data	3
1.4 Returns in financial modelling	7
1.5 Steps involved in formulating an econometric model	9
1.6 Points to consider when reading articles in empirical finance	10
1.7 Econometric packages for modelling financial data	11
1.8 Outline of the remainder of this book	22
1.9 Further reading	25
Appendix: Econometric software package suppliers	26
2 A brief overview of the classical linear regression model	27
2.1 What is a regression model?	27
2.2 Regression versus correlation	28
2.3 Simple regression	28
2.4 Some further terminology	37
2.5 Simple linear regression in EViews – estimation of an optimal hedge ratio	40
2.6 The assumptions underlying the classical linear regression model	43
2.7 Properties of the OLS estimator	44
2.8 Precision and standard errors	46
2.9 An introduction to statistical inference	51

2.10	A special type of hypothesis test: the t -ratio	65
2.11	An example of the use of a simple t -test to test a theory in finance: can US mutual funds beat the market?	67
2.12	Can UK unit trust managers beat the market?	69
2.13	The overreaction hypothesis and the UK stock market	71
2.14	The exact significance level	74
2.15	Hypothesis testing in EViews – example 1: hedging revisited	75
2.16	Estimation and hypothesis testing in EViews – example 2: the CAPM	77
	Appendix: Mathematical derivations of CLRM results	81
3	Further development and analysis of the classical linear regression model	88
3.1	Generalising the simple model to multiple linear regression	88
3.2	The constant term	89
3.3	How are the parameters (the elements of the β vector) calculated in the generalised case?	91
3.4	Testing multiple hypotheses: the F -test	93
3.5	Sample EViews output for multiple hypothesis tests	99
3.6	Multiple regression in EViews using an APT-style model	99
3.7	Data mining and the true size of the test	105
3.8	Goodness of fit statistics	106
3.9	Hedonic pricing models	112
3.10	Tests of non-nested hypotheses	115
	Appendix 3.1: Mathematical derivations of CLRM results	117
	Appendix 3.2: A brief introduction to factor models and principal components analysis	120
4	Classical linear regression model assumptions and diagnostic tests	129
4.1	Introduction	129
4.2	Statistical distributions for diagnostic tests	130
4.3	Assumption 1: $E(u_t) = 0$	131
4.4	Assumption 2: $\text{var}(u_t) = \sigma^2 < \infty$	132
4.5	Assumption 3: $\text{cov}(u_i, u_j) = 0$ for $i \neq j$	139
4.6	Assumption 4: the x_t are non-stochastic	160
4.7	Assumption 5: the disturbances are normally distributed	161
4.8	Multicollinearity	170
4.9	Adopting the wrong functional form	174
4.10	Omission of an important variable	178
4.11	Inclusion of an irrelevant variable	179

4.12	Parameter stability tests	180
4.13	A strategy for constructing econometric models and a discussion of model-building philosophies	191
4.14	Determinants of sovereign credit ratings	194
5	Univariate time series modelling and forecasting	206
5.1	Introduction	206
5.2	Some notation and concepts	207
5.3	Moving average processes	211
5.4	Autoregressive processes	215
5.5	The partial autocorrelation function	222
5.6	ARMA processes	223
5.7	Building ARMA models: the Box–Jenkins approach	230
5.8	Constructing ARMA models in EViews	234
5.9	Examples of time series modelling in finance	239
5.10	Exponential smoothing	241
5.11	Forecasting in econometrics	243
5.12	Forecasting using ARMA models in EViews	256
5.13	Estimating exponential smoothing models using EViews	258
6	Multivariate models	265
6.1	Motivations	265
6.2	Simultaneous equations bias	268
6.3	So how can simultaneous equations models be validly estimated?	269
6.4	Can the original coefficients be retrieved from the π s?	269
6.5	Simultaneous equations in finance	272
6.6	A definition of exogeneity	273
6.7	Triangular systems	275
6.8	Estimation procedures for simultaneous equations systems	276
6.9	An application of a simultaneous equations approach to modelling bid–ask spreads and trading activity	279
6.10	Simultaneous equations modelling using EViews	285
6.11	Vector autoregressive models	290
6.12	Does the VAR include contemporaneous terms?	295
6.13	Block significance and causality tests	297
6.14	VARs with exogenous variables	298
6.15	Impulse responses and variance decompositions	298
6.16	VAR model example: the interaction between property returns and the macroeconomy	302
6.17	VAR estimation in EViews	308

7 Modelling long-run relationships in finance	318
7.1 Stationarity and unit root testing	318
7.2 Testing for unit roots in EViews	331
7.3 Cointegration	335
7.4 Equilibrium correction or error correction models	337
7.5 Testing for cointegration in regression: a residuals-based approach	339
7.6 Methods of parameter estimation in cointegrated systems	341
7.7 Lead-lag and long-term relationships between spot and futures markets	343
7.8 Testing for and estimating cointegrating systems using the Johansen technique based on VARs	350
7.9 Purchasing power parity	355
7.10 Cointegration between international bond markets	357
7.11 Testing the expectations hypothesis of the term structure of interest rates	362
7.12 Testing for cointegration and modelling cointegrated systems using EViews	365
8 Modelling volatility and correlation	379
8.1 Motivations: an excursion into non-linearity land	379
8.2 Models for volatility	383
8.3 Historical volatility	383
8.4 Implied volatility models	384
8.5 Exponentially weighted moving average models	384
8.6 Autoregressive volatility models	385
8.7 Autoregressive conditionally heteroscedastic (ARCH) models	386
8.8 Generalised ARCH (GARCH) models	392
8.9 Estimation of ARCH/GARCH models	394
8.10 Extensions to the basic GARCH model	404
8.11 Asymmetric GARCH models	404
8.12 The GJR model	405
8.13 The EGARCH model	406
8.14 GJR and EGARCH in EViews	406
8.15 Tests for asymmetries in volatility	408
8.16 GARCH-in-mean	409
8.17 Uses of GARCH-type models including volatility forecasting	411
8.18 Testing non-linear restrictions or testing hypotheses about non-linear models	417
8.19 Volatility forecasting: some examples and results from the literature	420
8.20 Stochastic volatility models revisited	427

8.21	Forecasting covariances and correlations	428
8.22	Covariance modelling and forecasting in finance: some examples	429
8.23	Historical covariance and correlation	431
8.24	Implied covariance models	431
8.25	Exponentially weighted moving average model for covariances	432
8.26	Multivariate GARCH models	432
8.27	A multivariate GARCH model for the CAPM with time-varying covariances	436
8.28	Estimating a time-varying hedge ratio for FTSE stock index returns	437
8.29	Estimating multivariate GARCH models using EViews	441
	Appendix: Parameter estimation using maximum likelihood	444
9	Switching models	451
9.1	Motivations	451
9.2	Seasonalities in financial markets: introduction and literature review	454
9.3	Modelling seasonality in financial data	455
9.4	Estimating simple piecewise linear functions	462
9.5	Markov switching models	464
9.6	A Markov switching model for the real exchange rate	466
9.7	A Markov switching model for the gilt–equity yield ratio	469
9.8	Threshold autoregressive models	473
9.9	Estimation of threshold autoregressive models	474
9.10	Specification tests in the context of Markov switching and threshold autoregressive models: a cautionary note	476
9.11	A SETAR model for the French franc–German mark exchange rate	477
9.12	Threshold models and the dynamics of the FTSE 100 index and index futures markets	480
9.13	A note on regime switching models and forecasting accuracy	484
10	Panel data	487
10.1	Introduction – what are panel techniques and why are they used?	487
10.2	What panel techniques are available?	489
10.3	The fixed effects model	490
10.4	Time-fixed effects models	493
10.5	Investigating banking competition using a fixed effects model	494
10.6	The random effects model	498
10.7	Panel data application to credit stability of banks in Central and Eastern Europe	499
10.8	Panel data with EViews	502
10.9	Further reading	509

11 Limited dependent variable models	511
11.1 Introduction and motivation	511
11.2 The linear probability model	512
11.3 The logit model	514
11.4 Using a logit to test the pecking order hypothesis	515
11.5 The probit model	517
11.6 Choosing between the logit and probit models	518
11.7 Estimation of limited dependent variable models	518
11.8 Goodness of fit measures for linear dependent variable models	519
11.9 Multinomial linear dependent variables	521
11.10 The pecking order hypothesis revisited – the choice between financing methods	525
11.11 Ordered response linear dependent variables models	527
11.12 Are unsolicited credit ratings biased downwards? An ordered probit analysis	528
11.13 Censored and truncated dependent variables	533
11.14 Limited dependent variable models in EViews	537
Appendix: The maximum likelihood estimator for logit and probit models	544
12 Simulation methods	546
12.1 Motivations	546
12.2 Monte Carlo simulations	547
12.3 Variance reduction techniques	549
12.4 Bootstrapping	553
12.5 Random number generation	557
12.6 Disadvantages of the simulation approach to econometric or financial problem solving	558
12.7 An example of Monte Carlo simulation in econometrics: deriving a set of critical values for a Dickey-Fuller test	559
12.8 An example of how to simulate the price of a financial option	565
12.9 An example of bootstrapping to calculate capital risk requirements	571
13 Conducting empirical research or doing a project or dissertation in finance	585
13.1 What is an empirical research project and what is it for?	585
13.2 Selecting the topic	586
13.3 Sponsored or independent research?	590
13.4 The research proposal	590
13.5 Working papers and literature on the internet	591
13.6 Getting the data	591

13.7	Choice of computer software	593
13.8	How might the finished project look?	593
13.9	Presentational issues	597
14	Recent and future developments in the modelling of financial time series	598
14.1	Summary of the book	598
14.2	What was not covered in the book	598
14.3	Financial econometrics: the future?	602
14.4	The final word	606
Appendix 1	A review of some fundamental mathematical and statistical concepts	607
A1	Introduction	607
A2	Characteristics of probability distributions	607
A3	Properties of logarithms	608
A4	Differential calculus	609
A5	Matrices	611
A6	The eigenvalues of a matrix	614
Appendix 2	Tables of statistical distributions	616
Appendix 3	Sources of data used in this book	628
	<i>References</i>	629
	<i>Index</i>	641

Figures

<p>1.1 Steps involved in forming an econometric model page 9</p> <p>2.1 Scatter plot of two variables, y and x 29</p> <p>2.2 Scatter plot of two variables with a line of best fit chosen by eye 31</p> <p>2.3 Method of OLS fitting a line to the data by minimising the sum of squared residuals 32</p> <p>2.4 Plot of a single observation, together with the line of best fit, the residual and the fitted value 32</p> <p>2.5 Scatter plot of excess returns on fund XXX versus excess returns on the market portfolio 35</p> <p>2.6 No observations close to the y-axis 36</p> <p>2.7 Effect on the standard errors of the coefficient estimates when $(x_t - \bar{x})$ are narrowly dispersed 48</p> <p>2.8 Effect on the standard errors of the coefficient estimates when $(x_t - \bar{x})$ are widely dispersed 49</p> <p>2.9 Effect on the standard errors of x_t^2 large 49</p> <p>2.10 Effect on the standard errors of x_t^2 small 50</p> <p>2.11 The normal distribution 54</p> <p>2.12 The t-distribution versus the normal 55</p> <p>2.13 Rejection regions for a two-sided 5% hypothesis test 57</p> <p>2.14 Rejection regions for a one-sided hypothesis test of the form $H_0 : \beta = \beta^*$, $H_1 : \beta < \beta^*$ 57</p> <p>2.15 Rejection regions for a one-sided hypothesis test of the form $H_0 : \beta = \beta^*$, $H_1 : \beta > \beta^*$ 57</p> <p>2.16 Critical values and rejection regions for a $t_{20,5\%}$ 61</p>	<p>2.17 Frequency distribution of t-ratios of mutual fund alphas (gross of transactions costs) Source: Jensen (1968). Reprinted with the permission of Blackwell Publishers 68</p> <p>2.18 Frequency distribution of t-ratios of mutual fund alphas (net of transactions costs) Source: Jensen (1968). Reprinted with the permission of Blackwell Publishers 68</p> <p>2.19 Performance of UK unit trusts, 1979–2000 70</p> <p>3.1 $R^2 = 0$ demonstrated by a flat estimated line, i.e. a zero slope coefficient 109</p> <p>3.2 $R^2 = 1$ when all data points lie exactly on the estimated line 109</p> <p>4.1 Effect of no intercept on a regression line 131</p> <p>4.2 Graphical illustration of heteroscedasticity 132</p> <p>4.3 Plot of \hat{u}_t against \hat{u}_{t-1}, showing positive autocorrelation 141</p> <p>4.4 Plot of \hat{u}_t over time, showing positive autocorrelation 142</p> <p>4.5 Plot of \hat{u}_t against \hat{u}_{t-1}, showing negative autocorrelation 142</p> <p>4.6 Plot of \hat{u}_t over time, showing negative autocorrelation 143</p> <p>4.7 Plot of \hat{u}_t against \hat{u}_{t-1}, showing no autocorrelation 143</p> <p>4.8 Plot of \hat{u}_t over time, showing no autocorrelation 144</p> <p>4.9 Rejection and non-rejection regions for DW test 147</p>
--	---

4.10	A normal versus a skewed distribution	162	7.2	Value of t -ratio of slope coefficient for 1,000 sets of regressions of a non-stationary variable on another independent non-stationary variable	320
4.11	A leptokurtic versus a normal distribution	162	7.3	Example of a white noise process	324
4.12	Regression residuals from stock return data, showing large outlier for October 1987	165	7.4	Time series plot of a random walk versus a random walk with drift	324
4.13	Possible effect of an outlier on OLS estimation	166	7.5	Time series plot of a deterministic trend process	325
4.14	Plot of a variable showing suggestion for break date	185	7.6	Autoregressive processes with differing values of ϕ (0, 0.8, 1)	325
5.1	Autocorrelation function for sample MA(2) process	215	8.1	Daily S&P returns for January 1990–December 1999	387
5.2	Sample autocorrelation and partial autocorrelation functions for an MA(1) model: $y_t = -0.5u_{t-1} + u_t$	226	8.2	The problem of local optima in maximum likelihood estimation	397
5.3	Sample autocorrelation and partial autocorrelation functions for an MA(2) model: $y_t = 0.5u_{t-1} - 0.25u_{t-2} + u_t$	226	8.3	News impact curves for S&P500 returns using coefficients implied from GARCH and GJR model estimates	410
5.4	Sample autocorrelation and partial autocorrelation functions for a slowly decaying AR(1) model: $y_t = 0.9y_{t-1} + u_t$	227	8.4	Three approaches to hypothesis testing under maximum likelihood	418
5.5	Sample autocorrelation and partial autocorrelation functions for a more rapidly decaying AR(1) model: $y_t = 0.5y_{t-1} + u_t$	227	8.5	Source: Brooks, Henry and Persaud (2002). Time-varying hedge ratios derived from symmetric and asymmetric BEKK models for FTSE returns.	440
5.6	Sample autocorrelation and partial autocorrelation functions for a more rapidly decaying AR(1) model with negative coefficient: $y_t = -0.5y_{t-1} + u_t$	228	9.1	Sample time series plot illustrating a regime shift	452
5.7	Sample autocorrelation and partial autocorrelation functions for a non-stationary model (i.e. a unit coefficient): $y_t = y_{t-1} + u_t$	228	9.2	Use of intercept dummy variables for quarterly data	456
5.8	Sample autocorrelation and partial autocorrelation functions for an ARMA(1, 1) model: $y_t = 0.5y_{t-1} + 0.5u_{t-1} + u_t$	229	9.3	Use of slope dummy variables	459
5.9	Use of an in-sample and an out-of-sample period for analysis	245	9.4	Piecewise linear model with threshold x^*	463
6.1	Impulse responses and standard error bands for innovations in unexpected inflation equation errors	307	9.5	Source: Brooks and Persaud (2001b). Unconditional distribution of US GEYR together with a normal distribution with the same mean and variance	470
6.2	Impulse responses and standard error bands for innovations in the dividend yields	307	9.6	Source: Brooks and Persaud (2001b). Value of GEYR and probability that it is in the High GEYR regime for the UK	471
7.1	Value of R^2 for 1,000 sets of regressions of a non-stationary variable on another independent non-stationary variable	319	11.1	The fatal flaw of the linear probability model	513
			11.2	The logit model	515
			11.3	Modelling charitable donations as a function of income	534
			11.4	Fitted values from the failure probit regression	542

Tables

1.1	Econometric software packages for modelling financial data	<i>page</i> 12	5.1	Uncovered interest parity test results	241
2.1	Sample data on fund XXX to motivate OLS estimation	34	5.2	Forecast error aggregation	252
2.2	Critical values from the standard normal versus <i>t</i> -distribution	55	6.1	Call bid-ask spread and trading volume regression	283
2.3	Classifying hypothesis testing errors and correct conclusions	64	6.2	Put bid-ask spread and trading volume regression	283
2.4	Summary statistics for the estimated regression results for (2.52)	67	6.3	Granger causality tests and implied restrictions on VAR models	297
2.5	Summary statistics for unit trust returns, January 1979–May 2000	69	6.4	Marginal significance levels associated with joint <i>F</i> -tests	305
2.6	CAPM regression results for unit trust returns, January 1979–May 2000	70	6.5	Variance decompositions for the property sector index residuals	306
2.7	Is there an overreaction effect in the UK stock market?	73	7.1	Critical values for DF tests (Fuller, 1976, p. 373)	328
2.8	Part of the EViews regression output revisited	75	7.2	DF tests on log-prices and returns for high frequency FTSE data	344
3.1	Hedonic model of rental values in Quebec City, 1990. Dependent variable: Canadian dollars per month	114	7.3	Estimated potentially cointegrating equation and test for cointegration for high frequency FTSE data	345
3A.1	Principal component ordered eigenvalues for Dutch interest rates, 1962–1970	123	7.4	Estimated error correction model for high frequency FTSE data	346
3A.2	Factor loadings of the first and second principal components for Dutch interest rates, 1962–1970	123	7.5	Comparison of out-of-sample forecasting accuracy	346
4.1	Constructing a series of lagged values and first differences	140	7.6	Trading profitability of the error correction model with cost of carry	348
4.2	Determinants and impacts of sovereign credit ratings	197	7.7	Cointegration tests of PPP with European data	356
4.3	Do ratings add to public information?	199	7.8	DF tests for international bond indices	357
4.4	What determines reactions to ratings announcements?	201	7.9	Cointegration tests for pairs of international bond indices	358
			7.10	Johansen tests for cointegration between international bond yields	359
			7.11	Variance decompositions for VAR of international bond yields	360

7.12	Impulse responses for VAR of international bond yields	361	10.2	Tests of competition in banking with fixed effects panel models	497
7.13	Tests of the expectations hypothesis using the US zero coupon yield curve with monthly data	364	10.3	Results of random effects panel regression for credit stability of Central and East European banks	503
8.1	GARCH versus implied volatility	423	11.1	Logit estimation of the probability of external financing	517
8.2	EGARCH versus implied volatility	423	11.2	Multinomial logit estimation of the type of external financing	527
8.3	Out-of-sample predictive power for weekly volatility forecasts	426	11.3	Ordered probit model results for the determinants of credit ratings	531
8.4	Comparisons of the relative information content of out-of-sample volatility forecasts	426	11.4	Two-step ordered probit model allowing for selectivity bias in the determinants of credit ratings	532
8.5	Hedging effectiveness: summary statistics for portfolio returns	439	11.5	Marginal effects for logit and probit models for probability of MSc failure	543
9.1	Values and significances of days of the week coefficients	458	12.1	EGARCH estimates for currency futures returns	574
9.2	Day-of-the-week effects with the inclusion of interactive dummy variables with the risk proxy	461	12.2	Autoregressive volatility estimates for currency futures returns	575
9.3	Estimates of the Markov switching model for real exchange rates	468	12.3	Minimum capital risk requirements for currency futures as a percentage of the initial value of the position	578
9.4	Estimated parameters for the Markov switching models	470	13.1	Journals in finance and econometrics	589
9.5	SETAR model for FRF-DEM	478	13.2	Useful internet sites for financial literature	592
9.6	FRF-DEM forecast accuracies	479	13.3	Suggested structure for a typical dissertation or project	594
9.7	Linear AR(3) model for the basis	482			
9.8	A two-threshold SETAR model for the basis	483			
10.1	Tests of banking market equilibrium with fixed effects panel models	496			

Boxes

1.1	The value of econometrics	page 2	4.6	Observations for the dummy variable	165
1.2	Time series data	4	4.7	Conducting a Chow test	180
1.3	Log returns	8	5.1	The stationarity condition for an AR(p) model	216
1.4	Points to consider when reading a published paper	11	5.2	The invertibility condition for an MA(2) model	224
1.5	Features of EViews	21	5.3	Naive forecasting methods	247
2.1	Names for y and x s in regression models	28	6.1	Determining whether an equation is identified	270
2.2	Reasons for the inclusion of the disturbance term	30	6.2	Conducting a Hausman test for exogeneity	274
2.3	Assumptions concerning disturbance terms and their interpretation	44	6.3	Forecasting with VARs	299
2.4	Standard error estimators	48	7.1	Stationarity tests	331
2.5	Conducting a test of significance	56	7.2	Multiple cointegrating relationships	340
2.6	Carrying out a hypothesis test using confidence intervals	60	8.1	Testing for ‘ARCH effects’	390
2.7	The test of significance and confidence interval approaches compared	61	8.2	Estimating an ARCH or GARCH model	395
2.8	Type I and type II errors	64	8.3	Using maximum likelihood estimation in practice	398
2.9	Reasons for stock market overreactions	71	9.1	How do dummy variables work?	456
2.10	Ranking stocks and forming portfolios	72	10.1	Fixed or random effects?	500
2.11	Portfolio monitoring	72	11.1	Parameter interpretation for probit and logit models	519
3.1	The relationship between the regression F -statistic and R^2	111	11.2	The differences between censored and truncated dependent variables	535
3.2	Selecting between models	117	12.1	Conducting a Monte Carlo simulation	548
4.1	Conducting White’s test	134	12.2	Re-sampling the data	555
4.2	‘Solutions’ for heteroscedasticity	138	12.3	Re-sampling from the residuals	556
4.3	Conditions for DW to be a valid test	148	12.4	Setting up a Monte Carlo simulation	560
4.4	Conducting a Breusch–Godfrey test	149	12.5	Simulating the price of an Asian option	565
4.5	The Cochrane–Orcutt procedure	151	12.6	Generating draws from a GARCH process	566

Screenshots

1.1	Creating a workfile	<i>page</i> 15	5.4	Estimating exponential smoothing models	259
1.2	Importing Excel data into the workfile	16	6.1	Estimating the inflation equation	288
1.3	The workfile containing loaded data	17	6.2	Estimating the rsandp equation	289
1.4	Summary statistics for a series	19	6.3	VAR inputs screen	310
1.5	A line graph	20	6.4	Constructing the VAR impulse responses	313
2.1	Summary statistics for spot and futures	41	6.5	Combined impulse response graphs	314
2.2	Equation estimation window	42	6.6	Variance decomposition graphs	315
2.3	Estimation results	43	7.1	Options menu for unit root tests	332
2.4	Plot of two series	79	7.2	Actual, Fitted and Residual plot to check for stationarity	366
3.1	Stepwise procedure equation estimation window	103	7.3	Johansen cointegration test	368
3.2	Conducting PCA in EViews	126	7.4	VAR specification for Johansen tests	374
4.1	Regression options window	139	8.1	Estimating a GARCH-type model	400
4.2	Non-normality test results	164	8.2	GARCH model estimation options	401
4.3	Regression residuals, actual values and fitted series	168	8.3	Forecasting from GARCH models	415
4.4	Chow test for parameter stability	188	8.4	Dynamic forecasts of the conditional variance	415
4.5	Plotting recursive coefficient estimates	190	8.5	Static forecasts of the conditional variance	416
4.6	CUSUM test graph	191	8.6	Making a system	441
5.1	Estimating the correlogram	235	10.1	Workfile structure window	505
5.2	Plot and summary statistics for the dynamic forecasts for the percentage changes in house prices using an AR(2)	257	11.1	'Equation Estimation' window for limited dependent variables	539
5.3	Plot and summary statistics for the static forecasts for the percentage changes in house prices using an AR(2)	258	11.2	'Equation Estimation' options for limited dependent variables	541
			12.1	Running an EViews program	561