
A Guide to Modern Econometrics

4th edition

Marno Verbeek

Rotterdam School of Management, Erasmus University, Rotterdam

B 379887

 **WILEY**

A John Wiley & Sons, Ltd., Publication

Contents

Preface	xiii
1 Introduction	1
1.1 About Econometrics	1
1.2 The Structure of this Book	3
1.3 Illustrations and Exercises	4
2 An Introduction to Linear Regression	6
2.1 Ordinary Least Squares as an Algebraic Tool	7
2.1.1 Ordinary Least Squares	7
2.1.2 Simple Linear Regression	9
2.1.3 Example: Individual Wages	10
2.1.4 Matrix Notation	11
2.2 The Linear Regression Model	12
2.3 Small Sample Properties of the OLS Estimator	15
2.3.1 The Gauss–Markov Assumptions	15
2.3.2 Properties of the OLS Estimator	16
2.3.3 Example: Individual Wages (Continued)	20
2.4 Goodness-of-fit	20
2.5 Hypothesis Testing	22
2.5.1 A Simple t -Test	23
2.5.2 Example: Individual Wages (Continued)	25
2.5.3 Testing One Linear Restriction	25
2.5.4 A Joint Test of Significance of Regression Coefficients	26
2.5.5 Example: Individual Wages (Continued)	28
2.5.6 The General Case	29
2.5.7 Size, Power and p -Values	31

2.6	Asymptotic Properties of the OLS Estimator	32
2.6.1	Consistency	32
2.6.2	Asymptotic Normality	34
2.6.3	Small Samples and Asymptotic Theory	36
2.7	Illustration: The Capital Asset Pricing Model	38
2.7.1	The CAPM as a Regression Model	39
2.7.2	Estimating and Testing the CAPM	40
2.7.3	The World's Largest Hedge Fund	42
2.8	Multicollinearity	43
2.8.1	Example: Individual Wages (Continued)	46
2.9	Missing Data, Outliers and Influential Observations	47
2.9.1	Outliers and Influential Observations	47
2.9.2	Robust Estimation Methods	49
2.9.3	Missing Observations	50
2.10	Prediction	52
	Wrap-up	53
	Exercises	54
3	Interpreting and Comparing Regression Models	58
3.1	Interpreting the Linear Model	58
3.2	Selecting the Set of Regressors	62
3.2.1	Misspecifying the Set of Regressors	62
3.2.2	Selecting Regressors	63
3.2.3	Comparing Non-nested Models	67
3.3	Misspecifying the Functional Form	70
3.3.1	Nonlinear Models	70
3.3.2	Testing the Functional Form	71
3.3.3	Testing for a Structural Break	71
3.4	Illustration: Explaining House Prices	72
3.5	Illustration: Predicting Stock Index Returns	76
3.5.1	Model Selection	76
3.5.2	Forecast Evaluation	79
3.6	Illustration: Explaining Individual Wages	81
3.6.1	Linear Models	81
3.6.2	Loglinear Models	84
3.6.3	The Effects of Gender	87
3.6.4	Some Words of Warning	89
	Wrap-up	90
	Exercises	90
4	Heteroskedasticity and Autocorrelation	94
4.1	Consequences for the OLS Estimator	94
4.2	Deriving an Alternative Estimator	96
4.3	Heteroskedasticity	97
4.3.1	Introduction	97
4.3.2	Estimator Properties and Hypothesis Testing	100
4.3.3	When the Variances are Unknown	101

4.3.4	Heteroskedasticity-consistent Standard Errors for OLS	102
4.3.5	Multiplicative Heteroskedasticity	103
4.3.6	Weighted Least Squares with Arbitrary Weights	104
4.4	Testing for Heteroskedasticity	105
4.4.1	Testing for Multiplicative Heteroskedasticity	105
4.4.2	The Breusch–Pagan Test	106
4.4.3	The White Test	106
4.4.4	Which Test?	107
4.5	Illustration: Explaining Labour Demand	107
4.6	Autocorrelation	112
4.6.1	First-order Autocorrelation	113
4.6.2	Unknown ρ	115
4.7	Testing for First-order Autocorrelation	116
4.7.1	Asymptotic Tests	116
4.7.2	The Durbin–Watson Test	117
4.8	Illustration: The Demand for Ice Cream	119
4.9	Alternative Autocorrelation Patterns	122
4.9.1	Higher-order Autocorrelation	122
4.9.2	Moving Average Errors	122
4.10	What to do When you Find Autocorrelation?	123
4.10.1	Misspecification	124
4.10.2	Heteroskedasticity-and-autocorrelation-consistent Standard Errors for OLS	125
4.11	Illustration: Risk Premia in Foreign Exchange Markets	127
4.11.1	Notation	127
4.11.2	Tests for Risk Premia in the 1 Month Market	128
4.11.3	Tests for Risk Premia Using Overlapping Samples	132
	Wrap-up	134
	Exercises	134
5	Endogenous Regressors, Instrumental Variables and GMM	137
5.1	A Review of the Properties of the OLS Estimator	138
5.2	Cases Where the OLS Estimator Cannot be Saved	141
5.2.1	Autocorrelation with a Lagged Dependent Variable	141
5.2.2	Measurement Error in an Explanatory Variable	142
5.2.3	Endogeneity and Omitted Variable Bias	144
5.2.4	Simultaneity and Reverse Causality	146
5.3	The Instrumental Variables Estimator	148
5.3.1	Estimation with a Single Endogenous Regressor and a Single Instrument	148
5.3.2	Back to the Keynesian model	152
5.3.3	Back to the Measurement Error Problem	153
5.3.4	Multiple Endogenous Regressors	153

5.4	Illustration: Estimating the Returns to Schooling	154
5.5	The Generalized Instrumental Variables Estimator	158
5.5.1	Multiple Endogenous Regressors with an Arbitrary Number of Instruments	159
5.5.2	Two-stage Least Squares and the Keynesian Model Again	162
5.5.3	Specification Tests	163
5.5.4	Weak Instruments	164
5.6	The Generalized Method of Moments	166
5.6.1	Example	166
5.6.2	The Generalized Method of Moments	167
5.6.3	Some Simple Examples	170
5.6.4	Weak Identification	171
5.7	Illustration: Estimating Intertemporal Asset Pricing Models	171
	Wrap-up	175
	Exercises	176
6	Maximum Likelihood Estimation and Specification Tests	179
6.1	An Introduction to Maximum Likelihood	180
6.1.1	Some Examples	180
6.1.2	General Properties	183
6.1.3	An Example (Continued)	186
6.1.4	The Normal Linear Regression Model	187
6.2	Specification Tests	189
6.2.1	Three Test Principles	189
6.2.2	Lagrange Multiplier Tests	191
6.2.3	An Example (Continued)	194
6.3	Tests in the Normal Linear Regression Model	195
6.3.1	Testing for Omitted Variables	196
6.3.2	Testing for Heteroskedasticity	197
6.3.3	Testing for Autocorrelation	198
6.4	Quasi-maximum Likelihood and Moment Conditions Tests	199
6.4.1	Quasi-maximum Likelihood	199
6.4.2	Conditional Moment Tests	201
6.4.3	Testing for Normality	202
	Wrap-up	203
	Exercises	203
7	Models with Limited Dependent Variables	206
7.1	Binary Choice Models	207
7.1.1	Using Linear Regression?	207
7.1.2	Introducing Binary Choice Models	207
7.1.3	An Underlying Latent Model	210
7.1.4	Estimation	211

7.1.5	Goodness-of-fit	212
7.1.6	Illustration: The Impact of Unemployment Benefits on Reciprocity	215
7.1.7	Specification Tests in Binary Choice Models	217
7.1.8	Relaxing Some Assumptions in Binary Choice Models	219
7.2	Multiresponse Models	220
7.2.1	Ordered Response Models	221
7.2.2	About Normalization	222
7.2.3	Illustration: Explaining Firms' Credit Ratings	223
7.2.4	Illustration: Willingness to Pay for Natural Areas	225
7.2.5	Multinomial Models	228
7.3	Models for Count Data	231
7.3.1	The Poisson and Negative Binomial Models	231
7.3.2	Illustration: Patents and R&D Expenditures	235
7.4	Tobit Models	238
7.4.1	The Standard Tobit Model	238
7.4.2	Estimation	241
7.4.3	Illustration: Expenditures on Alcohol and Tobacco (Part 1)	242
7.4.4	Specification Tests in the Tobit Model	245
7.5	Extensions of Tobit Models	247
7.5.1	The Tobit II Model	248
7.5.2	Estimation	250
7.5.3	Further Extensions	253
7.5.4	Illustration: Expenditures on Alcohol and Tobacco (Part 2)	253
7.6	Sample Selection Bias	257
7.6.1	The Nature of the Selection Problem	257
7.6.2	Semi-parametric Estimation of the Sample-Selection Model	260
7.7	Estimating Treatment Effects	260
7.7.1	Regression-based Estimators	262
7.7.2	Alternative Approaches	266
7.8	Duration Models	268
7.8.1	Hazard Rates and Survival Functions	268
7.8.2	Samples and Model Estimation	270
7.8.3	Illustration: Duration of Bank Relationships	273
	Wrap-up	274
	Exercises	274
8	Univariate Time Series Models	278
8.1	Introduction	279
8.1.1	Some Examples	279
8.1.2	Stationarity and the Autocorrelation Function	281
8.2	General ARMA Processes	284

8.2.1	Formulating ARMA Processes	284
8.2.2	Invertibility of Lag Polynomials	287
8.2.3	Common Roots	288
8.3	Stationarity and Unit Roots	289
8.4	Testing for Unit Roots	291
8.4.1	Testing for Unit Roots in a First-order Autoregressive Model	291
8.4.2	Testing for Unit Roots in Higher-order Autoregressive Models	294
8.4.3	Extensions	296
8.4.4	Illustration: Stock Prices and Earnings	297
8.5	Illustration: Long-run Purchasing Power Parity (Part 1)	300
8.6	Estimation of ARMA Models	304
8.6.1	Least Squares	304
8.6.2	Maximum Likelihood	305
8.7	Choosing a Model	306
8.7.1	The Autocorrelation Function	306
8.7.2	The Partial Autocorrelation Function	308
8.7.3	Diagnostic Checking	309
8.7.4	Criteria for Model Selection	310
8.8	Illustration: The Persistence of Inflation	311
8.9	Predicting with ARMA Models	314
8.9.1	The Optimal Predictor	315
8.9.2	Prediction Accuracy	317
8.9.3	Evaluating predictions	319
8.10	Illustration: The Expectations Theory of the Term Structure	320
8.11	Autoregressive Conditional Heteroskedasticity	325
8.11.1	ARCH and GARCH Models	325
8.11.2	Estimation and Prediction	329
8.11.3	Illustration: Volatility in Daily Exchange Rates	331
8.12	What about Multivariate Models?	333
	Wrap-up	333
	Exercises	334
9	Multivariate Time Series Models	338
9.1	Dynamic Models with Stationary Variables	339
9.2	Models with Nonstationary Variables	342
9.2.1	Spurious Regressions	342
9.2.2	Cointegration	343
9.2.3	Cointegration and Error-correction Mechanisms	346
9.3	Illustration: Long-run Purchasing Power Parity (Part 2)	348
9.4	Vector Autoregressive Models	350
9.5	Cointegration: the Multivariate Case	354
9.5.1	Cointegration in a VAR	354
9.5.2	Example: Cointegration in a Bivariate VAR	356
9.5.3	Testing for Cointegration	358

9.5.4	Illustration: Long-run Purchasing Power Parity (Part 3)	360
9.6	Illustration: Money Demand and Inflation	362
	Wrap-up	368
	Exercises	369
10	Models Based on Panel Data	372
10.1	Introduction to Panel Data Modelling	373
10.1.1	Efficiency of Parameter Estimators	374
10.1.2	Identification of Parameters	375
10.2	The Static Linear Model	376
10.2.1	The Fixed Effects Model	377
10.2.2	The First-difference Estimator	379
10.2.3	The Random Effects Model	381
10.2.4	Fixed Effects or Random Effects?	384
10.2.5	Goodness-of-fit	386
10.2.6	Alternative Instrumental Variables Estimators	387
10.2.7	Robust Inference	389
10.2.8	Testing for Heteroskedasticity and Autocorrelation	391
10.2.9	The Fama–MacBeth Approach	392
10.3	Illustration: Explaining Individual Wages	394
10.4	Dynamic Linear Models	396
10.4.1	An Autoregressive Panel Data Model	396
10.4.2	Dynamic Models with Exogenous Variables	401
10.4.3	Too Many Instruments	403
10.5	Illustration: Explaining Capital Structure	405
10.6	Panel Time Series	410
10.6.1	Heterogeneity	411
10.6.2	First Generation Panel Unit Root Tests	412
10.6.3	Second Generation Panel Unit Root Tests	415
10.6.4	Panel Cointegration Tests	416
10.7	Models with Limited Dependent Variables	417
10.7.1	Binary Choice Models	418
10.7.2	The Fixed Effects Logit Model	419
10.7.3	The Random Effects Probit Model	420
10.7.4	Tobit Models	422
10.7.5	Dynamics and the Problem of Initial Conditions	423
10.7.6	Semi-parametric Alternatives	424
10.8	Incomplete Panels and Selection Bias	425
10.8.1	Estimation with Randomly Missing Data	426
10.8.2	Selection Bias and Some Simple Tests	427
10.8.3	Estimation with Nonrandomly Missing Data	429
10.9	Pseudo Panels and Repeated Cross-sections	430
10.9.1	The Fixed Effects Model	431
10.9.2	An Instrumental Variables Interpretation	433

10.9.3 Dynamic Models	434
Wrap-up	435
Exercises	436
A Vectors and Matrices	441
A.1 Terminology	441
A.2 Matrix Manipulations	442
A.3 Properties of Matrices and Vectors	443
A.4 Inverse Matrices	444
A.5 Idempotent Matrices	445
A.6 Eigenvalues and Eigenvectors	445
A.7 Differentiation	446
A.8 Some Least Squares Manipulations	447
B Statistical and Distribution Theory	449
B.1 Discrete Random Variables	449
B.2 Continuous Random Variables	450
B.3 Expectations and Moments	451
B.4 Multivariate Distributions	452
B.5 Conditional Distributions	453
B.6 The Normal Distribution	454
B.7 Related Distributions	457
Bibliography	459
Index	477