
A Guide to Modern Econometrics

Fifth Edition

Marno Verbeek

Rotterdam School of Management, Erasmus University, Rotterdam

Wiley Custom

Contents

Preface	xi
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	11
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	23
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	30
2.5.8 Reporting Regression Results	32

2.6	Asymptotic Properties of the OLS Estimator	33
2.6.1	Consistency	33
2.6.2	Asymptotic Normality	35
2.6.3	Small Samples and Asymptotic Theory	37
2.7	Illustration: The Capital Asset Pricing Model	39
2.7.1	The CAPM as a Regression Model	40
2.7.2	Estimating and Testing the CAPM	41
2.7.3	The World's Largest Hedge Fund	43
2.8	Multicollinearity	44
2.8.1	Example: Individual Wages (Continued)	47
2.9	Missing Data, Outliers and Influential Observations	48
2.9.1	Outliers and Influential Observations	48
2.9.2	Robust Estimation Methods	50
2.9.3	Missing Observations	51
2.10	Prediction	53
	Wrap-up	54
	Exercises	55
3	Interpreting and Comparing Regression Models	60
3.1	Interpreting the Linear Model	60
3.2	Selecting the Set of Regressors	65
3.2.1	Misspecifying the Set of Regressors	65
3.2.2	Selecting Regressors	66
3.2.3	Comparing Non-nested Models	71
3.3	Misspecifying the Functional Form	73
3.3.1	Nonlinear Models	73
3.3.2	Testing the Functional Form	74
3.3.3	Testing for a Structural Break	74
3.4	Illustration: Explaining House Prices	76
3.5	Illustration: Predicting Stock Index Returns	79
3.5.1	Model Selection	80
3.5.2	Forecast Evaluation	82
3.6	Illustration: Explaining Individual Wages	85
3.6.1	Linear Models	85
3.6.2	Loglinear Models	88
3.6.3	The Effects of Gender	91
3.6.4	Some Words of Warning	92
	Wrap-up	93
	Exercises	94
4	Heteroskedasticity and Autocorrelation	97
4.1	Consequences for the OLS Estimator	98
4.2	Deriving an Alternative Estimator	99
4.3	Heteroskedasticity	100
4.3.1	Introduction	100
4.3.2	Estimator Properties and Hypothesis Testing	103

4.3.3	When the Variances Are Unknown	104
4.3.4	Heteroskedasticity-consistent Standard Errors for OLS	105
4.3.5	Multiplicative Heteroskedasticity	106
4.3.6	Weighted Least Squares with Arbitrary Weights	107
4.4	Testing for Heteroskedasticity	108
4.4.1	Testing for Multiplicative Heteroskedasticity	108
4.4.2	The Breusch–Pagan Test	109
4.4.3	The White Test	109
4.4.4	Which Test?	110
4.5	Illustration: Explaining Labour Demand	110
4.6	Autocorrelation	114
4.6.1	First-order Autocorrelation	116
4.6.2	Unknown ρ	118
4.7	Testing for First-order Autocorrelation	119
4.7.1	Asymptotic Tests	119
4.7.2	The Durbin–Watson Test	120
4.8	Illustration: The Demand for Ice Cream	121
4.9	Alternative Autocorrelation Patterns	124
4.9.1	Higher-order Autocorrelation	124
4.9.2	Moving Average Errors	125
4.10	What to Do When You Find Autocorrelation?	126
4.10.1	Misspecification	126
4.10.2	Heteroskedasticity-and-autocorrelation-consistent Standard Errors for OLS	128
4.11	Illustration: Risk Premia in Foreign Exchange Markets	129
4.11.1	Notation	129
4.11.2	Tests for Risk Premia in the 1-Month Market	131
4.11.3	Tests for Risk Premia Using Overlapping Samples	134
	Wrap-up	136
	Exercises	136
5	Endogenous Regressors, Instrumental Variables and GMM	139
5.1	A Review of the Properties of the OLS Estimator	140
5.2	Cases Where the OLS Estimator Cannot Be Saved	143
5.2.1	Autocorrelation with a Lagged Dependent Variable	143
5.2.2	Measurement Error in an Explanatory Variable	144
5.2.3	Endogeneity and Omitted Variable Bias	146
5.2.4	Simultaneity and Reverse Causality	148
5.3	The Instrumental Variables Estimator	150
5.3.1	Estimation with a Single Endogenous Regressor and a Single Instrument	150
5.3.2	Back to the Keynesian Model	155
5.3.3	Back to the Measurement Error Problem	156
5.3.4	Multiple Endogenous Regressors	156
5.4	Illustration: Estimating the Returns to Schooling	157

5.5	Alternative Approaches to Estimate Causal Effects	162
5.6	The Generalized Instrumental Variables Estimator	163
5.6.1	Multiple Endogenous Regressors with an Arbitrary Number of Instruments	163
5.6.2	Two-stage Least Squares and the Keynesian Model Again	167
5.6.3	Specification Tests	168
5.6.4	Weak Instruments	169
5.6.5	Implementing and Reporting Instrumental Variables Estimators	170
5.7	Institutions and Economic Development	171
5.8	The Generalized Method of Moments	175
5.8.1	Example	175
5.8.2	The Generalized Method of Moments	177
5.8.3	Some Simple Examples	179
5.8.4	Weak Identification	180
5.9	Illustration: Estimating Intertemporal Asset Pricing Models	181
	Wrap-up	184
	Exercises	185
6	Maximum Likelihood Estimation and Specification Tests	187
6.1	An Introduction to Maximum Likelihood	188
6.1.1	Some Examples	188
6.1.2	General Properties	191
6.1.3	An Example (Continued)	194
6.1.4	The Normal Linear Regression Model	195
6.1.5	The Stochastic Frontier Model	197
6.2	Specification Tests	198
6.2.1	Three Test Principles	198
6.2.2	Lagrange Multiplier Tests	200
6.2.3	An Example (Continued)	203
6.3	Tests in the Normal Linear Regression Model	204
6.3.1	Testing for Omitted Variables	204
6.3.2	Testing for Heteroskedasticity	206
6.3.3	Testing for Autocorrelation	207
6.4	Quasi-maximum Likelihood and Moment Conditions Tests	208
6.4.1	Quasi-maximum Likelihood	208
6.4.2	Conditional Moment Tests	210
6.4.3	Testing for Normality	211
	Wrap-up	212
	Exercises	212
7	Models with Limited Dependent Variables	215
7.1	Binary Choice Models	216
7.1.1	Using Linear Regression?	216
7.1.2	Introducing Binary Choice Models	216
7.1.3	An Underlying Latent Model	219

7.1.4	Estimation	219
7.1.5	Goodness-of-Fit	221
7.1.6	Illustration: The Impact of Unemployment Benefits on Reciprocity	223
7.1.7	Specification Tests in Binary Choice Models	226
7.1.8	Relaxing Some Assumptions in Binary Choice Models	228
7.2	Multiresponse Models	229
7.2.1	Ordered Response Models	230
7.2.2	About Normalization	231
7.2.3	Illustration: Explaining Firms' Credit Ratings	231
7.2.4	Illustration: Willingness to Pay for Natural Areas	234
7.2.5	Multinomial Models	237
7.3	Models for Count Data	240
7.3.1	The Poisson and Negative Binomial Models	240
7.3.2	Illustration: Patents and R&D Expenditures	244
7.4	Tobit Models	246
7.4.1	The Standard Tobit Model	247
7.4.2	Estimation	249
7.4.3	Illustration: Expenditures on Alcohol and Tobacco (Part 1)	250
7.4.4	Specification Tests in the Tobit Model	253
7.5	Extensions of Tobit Models	256
7.5.1	The Tobit II Model	256
7.5.2	Estimation	259
7.5.3	Further Extensions	261
7.5.4	Illustration: Expenditures on Alcohol and Tobacco (Part 2)	262
7.6	Sample Selection Bias	265
7.6.1	The Nature of the Selection Problem	266
7.6.2	Semi-parametric Estimation of the Sample Selection Model	268
7.7	Estimating Treatment Effects	269
7.7.1	Regression-based Estimators	271
7.7.2	Regression Discontinuity Design	274
7.7.3	Weighting and Matching	276
7.8	Duration Models	278
7.8.1	Hazard Rates and Survival Functions	278
7.8.2	Samples and Model Estimation	281
7.8.3	Illustration: Duration of Bank Relationships	283
	Wrap-up	284
	Exercises	285
8	Univariate Time Series Models	288
8.1	Introduction	289
8.1.1	Some Examples	289
8.1.2	Stationarity and the Autocorrelation Function	291

8.2	General ARMA Processes	294
8.2.1	Formulating ARMA Processes	294
8.2.2	Invertibility of Lag Polynomials	297
8.2.3	Common Roots	298
8.3	Stationarity and Unit Roots	299
8.4	Testing for Unit Roots	301
8.4.1	Testing for Unit Roots in a First-order Autoregressive Model	301
8.4.2	Testing for Unit Roots in Higher-Order Autoregressive Models	304
8.4.3	Extensions	306
8.4.4	Illustration: Stock Prices and Earnings	307
8.5	Illustration: Long-run Purchasing Power Parity (Part 1)	309
8.6	Estimation of ARMA Models	313
8.6.1	Least Squares	314
8.6.2	Maximum Likelihood	315
8.7	Choosing a Model	316
8.7.1	The Autocorrelation Function	316
8.7.2	The Partial Autocorrelation Function	318
8.7.3	Diagnostic Checking	319
8.7.4	Criteria for Model Selection	319
8.8	Illustration: The Persistence of Inflation	320
8.9	Forecasting with ARMA Models	324
8.9.1	The Optimal Forecast	324
8.9.2	Forecast Accuracy	327
8.9.3	Evaluating Forecasts	329
8.10	Illustration: The Expectations Theory of the Term Structure	330
8.11	Autoregressive Conditional Heteroskedasticity	335
8.11.1	ARCH and GARCH Models	335
8.11.2	Estimation and Prediction	338
8.11.3	Illustration: Volatility in Daily Exchange Rates	340
8.12	What about Multivariate Models?	342
	Wrap-up	343
	Exercises	344
9	Multivariate Time Series Models	348
9.1	Dynamic Models with Stationary Variables	349
9.2	Models with Nonstationary Variables	352
9.2.1	Spurious Regressions	352
9.2.2	Cointegration	353
9.2.3	Cointegration and Error-correction Mechanisms	356
9.3	Illustration: Long-run Purchasing Power Parity (Part 2)	358
9.4	Vector Autoregressive Models	360
9.5	Cointegration: the Multivariate Case	364
9.5.1	Cointegration in a VAR	364
9.5.2	Example: Cointegration in a Bivariate VAR	366
9.5.3	Testing for Cointegration	367

9.5.4	Illustration: Long-run Purchasing Power Parity (Part 3)	370
9.6	Illustration: Money Demand and Inflation	372
	Wrap-up	378
	Exercises	379
10	Models Based on Panel Data	382
10.1	Introduction to Panel Data Modelling	383
10.1.1	Efficiency of Parameter Estimators	384
10.1.2	Identification of Parameters	385
10.2	The Static Linear Model	386
10.2.1	The Fixed Effects Model	386
10.2.2	The First-difference Estimator	388
10.2.3	The Random Effects Model	390
10.2.4	Fixed Effects or Random Effects?	394
10.2.5	Goodness-of-Fit	395
10.2.6	Alternative Instrumental Variables Estimators	396
10.2.7	Robust Inference	398
10.2.8	Testing for Heteroskedasticity and Autocorrelation	400
10.2.9	The Fama–MacBeth Approach	402
10.3	Illustration: Explaining Individual Wages	403
10.4	Dynamic Linear Models	405
10.4.1	An Autoregressive Panel Data Model	406
10.4.2	Dynamic Models with Exogenous Variables	411
10.4.3	Too Many Instruments	412
10.5	Illustration: Explaining Capital Structure	414
10.6	Panel Time Series	419
10.6.1	Heterogeneity	420
10.6.2	First Generation Panel Unit Root Tests	421
10.6.3	Second Generation Panel Unit Root Tests	424
10.6.4	Panel Cointegration Tests	425
10.7	Models with Limited Dependent Variables	426
10.7.1	Binary Choice Models	427
10.7.2	The Fixed Effects Logit Model	428
10.7.3	The Random Effects Probit Model	429
10.7.4	Tobit Models	431
10.7.5	Dynamics and the Problem of Initial Conditions	431
10.7.6	Semi-parametric Alternatives	433
10.8	Incomplete Panels and Selection Bias	433
10.8.1	Estimation with Randomly Missing Data	434
10.8.2	Selection Bias and Some Simple Tests	436
10.8.3	Estimation with Nonrandomly Missing Data	438
10.9	Pseudo Panels and Repeated Cross-sections	439
10.9.1	The Fixed Effects Model	440
10.9.2	An Instrumental Variables Interpretation	441
10.9.3	Dynamic Models	442
	Wrap-up	444
	Exercises	445

A	Vectors and Matrices	450
A.1	Terminology	450
A.2	Matrix Manipulations	451
A.3	Properties of Matrices and Vectors	452
A.4	Inverse Matrices	453
A.5	Idempotent Matrices	454
A.6	Eigenvalues and Eigenvectors	454
A.7	Differentiation	455
A.8	Some Least Squares Manipulations	456
B	Statistical and Distribution Theory	458
B.1	Discrete Random Variables	458
B.2	Continuous Random Variables	459
B.3	Expectations and Moments	460
B.4	Multivariate Distributions	461
B.5	Conditional Distributions	462
B.6	The Normal Distribution	463
B.7	Related Distributions	466
	Bibliography	468
	Index	488