

Introduction to Econometrics

Fourth Edition

G. S. Maddala
Kajal Lahiri



A John Wiley and Sons, Ltd., Publication

CONTENTS

Foreword	xvii
Preface to the Fourth Edition	xix
Part I Introduction and the Linear Regression Model	1
CHAPTER 1	
What is Econometrics?	3
What is in this Chapter?	3
1.1 What is econometrics?	3
1.2 Economic and econometric models	4
1.3 The aims and methodology of econometrics	6
1.4 What constitutes a test of an economic theory?	8
Summary and an outline of the book	9
References	10
CHAPTER 2	
Statistical Background and Matrix Algebra	11
What is in this Chapter?	11
2.1 Introduction	11
2.2 Probability	12
Addition rules of probability	13
Conditional probability and the multiplication rule	14
Bayes' theorem	14
Summation and product operations	15
2.3 Random variables and probability distributions	17
Joint, marginal, and conditional distributions	17
Illustrative example	18
2.4 The normal probability distribution and related distributions	18
The normal distribution	19
Related distributions	19
2.5 Classical statistical inference	21
2.6 Properties of estimators	22
Unbiasedness	23
Efficiency	23
Consistency	23
Other asymptotic properties	24

2.7	Sampling distributions for samples from a normal population	26
2.8	Interval estimation	26
2.9	Testing of hypotheses	28
2.10	Relationship between confidence interval procedures and tests of hypotheses	31
2.11	Combining independent tests	32
	Summary	33
	Exercises	33
	Appendix: Matrix algebra	40
	Exercises on matrix algebra	56
	References	57

CHAPTER 3

	Simple Regression	59
	What is in this Chapter?	59
3.1	Introduction	59
	Example 1: Simple regression	60
	Example 2: Multiple regression	60
3.2	Specification of the relationships	61
3.3	The method of moments	65
	Illustrative example	66
3.4	The method of least squares	68
	Reverse regression	71
	Illustrative example	72
3.5	Statistical inference in the linear regression model	76
	Illustrative example	78
	Confidence intervals for α , β , and σ^2	78
	Testing of hypotheses	80
	Example of comparing test scores from the GRE and GMAT tests	81
	Regression with no constant term	83
3.6	Analysis of variance for the simple regression model	83
3.7	Prediction with the simple regression model	85
	Prediction of expected values	87
	Illustrative example	87
3.8	Outliers	88
	Some illustrative examples	90
3.9	Alternative functional forms for regression equations	95
	Illustrative example	97
*3.10	Inverse prediction in the least squares regression model ¹	99
*3.11	Stochastic regressors	102
*3.12	The regression fallacy	102
	The bivariate normal distribution	102
	Galton's result and the regression fallacy	104
	A note on the term "regression"	105
	Summary	105

¹ Here and below, the * indicates that this is an optional section.

Exercises	106
Appendix: Proofs	113
References	126
CHAPTER 4	
Multiple Regression	127
What is in this Chapter?	127
4.1 Introduction	127
4.2 A model with two explanatory variables	129
The least squares method	129
Illustrative example	132
4.3 Statistical inference in the multiple regression model	134
Illustrative example	135
Formulas for the general case of k explanatory variables	139
Some illustrative examples	140
4.4 Interpretation of the regression coefficients	143
Illustrative example	146
4.5 Partial correlations and multiple correlation	146
4.6 Relationships among simple, partial, and multiple correlation coefficients	147
Two illustrative examples	148
4.7 Prediction in the multiple regression model	153
Illustrative example	154
4.8 Analysis of variance and tests of hypotheses	155
Nested and nonnested hypotheses	157
Tests for linear functions of parameters	158
Illustrative example	159
4.9 Omission of relevant variables and inclusion of irrelevant variables	160
Omission of relevant variables	161
Example 1: Demand for food in the United States	163
Example 2: Production functions and management bias	163
Inclusion of irrelevant variables	164
4.10 Degrees of freedom and \bar{R}^2	165
4.11 Tests for stability	169
The analysis of variance test	169
Example 1: Stability of the demand for food function	170
Example 2: Stability of production functions	171
Predictive tests for stability	174
Illustrative example	174
4.12 The LR, W, and LM tests	176
Illustrative example	177
Summary	178
Exercises	180
Appendix 4.1: The multiple regression model in matrix notation	187
Appendix 4.2: Nonlinear regressions	193
Appendix 4.3: Large-sample theory	196
Data sets	202
References	207

Part II	Violation of the Assumptions of the Basic Regression Model	209
CHAPTER 5		
Heteroskedasticity		
		211
What is in this Chapter?		211
5.1	Introduction	211
	Illustrative example	212
5.2	Detection of heteroskedasticity	214
	Illustrative example	214
	Some other tests	215
	Illustrative example	217
	An intuitive justification for the Breusch–Pagan test	218
5.3	Consequences of heteroskedasticity	219
	Estimation of the variance of the OLS estimator under heteroskedasticity	221
5.4	Solutions to the heteroskedasticity problem	221
	Illustrative example	223
5.5	Heteroskedasticity and the use of deflators	224
	Illustrative example: The density gradient model	226
5.6	Testing the linear versus log-linear functional form	228
	The Box–Cox test	229
	The BM test	230
	The PE test	231
Summary		231
Exercises		232
Appendix: Generalized least squares		235
References		237
CHAPTER 6		
Autocorrelation		
		239
What is in this Chapter?		239
6.1	Introduction	239
6.2	The Durbin–Watson test	240
	Illustrative example	241
6.3	Estimation in levels versus first differences	242
	Some illustrative examples	243
6.4	Estimation procedures with autocorrelated errors	246
	Iterative procedures	248
	Grid-search procedures	249
	Illustrative example	250
6.5	Effect of AR(1) errors on OLS estimates	250
6.6	Some further comments on the DW test	254
	The von Neumann ratio	255
	The Berenblut–Webb test	255
6.7	Tests for serial correlation in models with lagged dependent variables	257
	Durbin's <i>h</i> -test	258
	Durbin's alternative test	258
	Illustrative example	258

6.8	A general test for higher-order serial correlation: The LM test	259
6.9	Strategies when the DW test statistic is significant	261
	Autocorrelation caused by omitted variables	261
	Serial correlation due to misspecified dynamics	263
	The Wald test	264
	Illustrative example	265
*6.10	Trends and random walks	266
	Spurious trends	268
	Differencing and long-run effects: The concept of cointegration	270
*6.11	ARCH models and serial correlation	271
6.12	Some comments on the DW test and Durbin's <i>h</i> -test and <i>t</i> -test	272
	Summary	273
	Exercises	274
	References	276

CHAPTER 7

Multicollinearity	279	
What is in this Chapter?	279	
7.1 Introduction	279	
7.2 Some illustrative examples	280	
7.3 Some measures of multicollinearity	283	
7.4 Problems with measuring multicollinearity	286	
7.5 Solutions to the multicollinearity problem: Ridge regression	290	
7.6 Principal component regression	292	
7.7 Dropping variables	297	
7.8 Miscellaneous other solutions	300	
	Using ratios or first differences	300
	Using extraneous estimates	300
	Getting more data	301
Summary	302	
Exercises	302	
Appendix: Linearly dependent explanatory variables	304	
References	311	

CHAPTER 8

Dummy Variables and Truncated Variables	313	
What is in this Chapter?	313	
8.1 Introduction	313	
8.2 Dummy variables for changes in the intercept term	314	
	Illustrative example	317
	Two more illustrative examples	317
8.3 Dummy variables for changes in slope coefficients	319	
8.4 Dummy variables for cross-equation constraints	322	
8.5 Dummy variables for testing stability of regression coefficients	324	
8.6 Dummy variables under heteroskedasticity and autocorrelation	327	
8.7 Dummy dependent variables	329	

8.8	The linear probability model and the linear discriminant function	329
	The linear probability model	329
	The linear discriminant function	332
8.9	The probit and logit models	333
	Illustrative example	335
	The problem of disproportionate sampling	336
	Prediction of effects of changes in the explanatory variables	337
	Measuring goodness of fit	338
	Illustrative example	340
8.10	Truncated variables: The tobit model	343
	Some examples	344
	Method of estimation	345
	Limitations of the tobit model	346
	The truncated regression model	347
	Summary	349
	Exercises	350
	References	352

CHAPTER 9

	Simultaneous Equation Models	355
	What is in this Chapter?	355
9.1	Introduction	355
9.2	Endogenous and exogenous variables	357
9.3	The identification problem: Identification through reduced form	357
	Illustrative example	360
9.4	Necessary and sufficient conditions for identification	362
	Illustrative example	364
9.5	Methods of estimation: The instrumental variable method	365
	Measuring R^2	368
	Illustrative example	368
9.6	Methods of estimation: The two-stage least squares method	371
	Computing standard errors	373
	Illustrative example	375
9.7	The question of normalization	378
*9.8	The limited-information maximum likelihood method	379
	Illustrative example	380
*9.9	On the use of OLS in the estimation of simultaneous equation models	380
	Working's concept of identification	382
	Recursive systems	384
	Estimation of Cobb–Douglas production functions	385
*9.10	Exogeneity and causality	386
	Weak exogeneity	389
	Superexogeneity	389
	Strong exogeneity	389
	Granger causality	390

Granger causality and exogeneity	390
Tests for exogeneity	391
9.11 Some problems with instrumental variable methods	392
Summary	392
Exercises	394
Appendix	396
References	400

CHAPTER 10

Diagnostic Checking, Model Selection, and Specification Testing	401
What is in this Chapter?	401
10.1 Introduction	401
10.2 Diagnostic tests based on least squares residuals	402
Tests for omitted variables	402
Tests for ARCH effects	404
10.3 Problems with least squares residuals	404
10.4 Some other types of residual	405
Predicted residuals and studentized residuals	406
Dummy variable method for studentized residuals	407
BLUS residuals	407
Recursive residuals	408
Illustrative example	409
10.5 DFFITS and bounded influence estimation	411
Illustrative example	413
10.6 Model selection	414
Hypothesis-testing search	415
Interpretive search	416
Simplification search	416
Proxy variable search	416
Data selection search	417
Post-data model construction	417
Hendry's approach to model selection	418
10.7 Selection of regressors	419
Theil's \bar{R}^2 criterion	421
Criteria based on minimizing the mean-squared error of prediction	421
Akaike's information criterion	422
10.8 Implied F-ratios for the various criteria	423
Bayes' theorem and posterior odds for model selection	425
10.9 Cross-validation	427
10.10 Hausman's specification error test	428
An application: Testing for errors in variables or exogeneity	430
Some illustrative examples	431
An omitted variable interpretation of the Hausman test	433
10.11 The Plosser–Schwert–White differencing test	435
10.12 Tests for nonnested hypotheses	436

The Davidson and MacKinnon test	437
The encompassing test	439
A basic problem in testing nonnested hypotheses	440
Hypothesis testing versus model selection as a research strategy	440
10.13 Nonnormality of errors	440
Tests for normality	441
10.14 Data transformations	441
Summary	442
Exercises	444
Appendix	446
References	447
CHAPTER 11	
Errors in Variables	451
What is in this Chapter?	451
11.1 Introduction	451
11.2 The classical solution for a single-equation model with one explanatory variable	452
11.3 The single-equation model with two explanatory variables	455
Two explanatory variables: One measured with error	455
Illustrative example	459
Two explanatory variables: Both measured with error	460
11.4 Reverse regression	463
11.5 Instrumental variable methods	465
11.6 Proxy variables	468
Coefficient for the proxy variable	470
11.7 Some other problems	471
The case of multiple equations	471
Correlated errors	472
Summary	473
Exercises	474
References	476
Part III Special Topics	479
CHAPTER 12	
Introduction to Time-Series Analysis	481
What is in this Chapter?	481
12.1 Introduction	481
12.2 Two methods of time-series analysis: Frequency domain and time domain	482
12.3 Stationary and nonstationary time series	482
Strict stationarity	483
Weak stationarity	483
Properties of autocorrelation function	484
Nonstationarity	484
12.4 Some useful models for time series	485
Purely random process	485

Random walk	486
Moving average process	486
Autoregressive process	488
Autoregressive moving average process	490
Autoregressive integrated moving average process	491
12.5 Estimation of AR, MA, and ARMA models	492
<i>Estimation of MA models</i>	492
Estimation of ARMA models	492
Residuals from the ARMA models	494
Testing goodness of fit	494
12.6 The Box–Jenkins approach	496
Forecasting from Box–Jenkins models	497
Illustrative example	499
Trend elimination: The traditional method	500
A summary assessment	500
Seasonality in the Box–Jenkins modeling	502
12.7 R^2 measures in time-series models	503
Summary	506
Exercises	506
Data sets	507
References	508

CHAPTER 13

Models of Expectations and Distributed Lags	509
What is in this Chapter?	509
13.1 Models of expectations	509
13.2 Naive models of expectations	510
13.3 The adaptive expectations model	512
13.4 Estimation with the adaptive expectations model	514
Estimation in the autoregressive form	514
Estimation in the distributed lag form	515
13.5 Two illustrative examples	516
13.6 Expectational variables and adjustment lags	520
13.7 Partial adjustment with adaptive expectations	524
13.8 Alternative distributed lag models: Polynomial lags	526
Finite lags: The polynomial lag	527
Illustrative example	530
Choosing the degree of the polynomial	532
13.9 Rational lags	533
13.10 Rational expectations	534
13.11 Tests for rationality	536
13.12 Estimation of a demand and supply model under rational expectations	538
Case 1	538
Case 2	539
Illustrative example	542
13.13 The serial correlation problem in rational expectations models	544

Summary	545
Exercises	547
References	548

CHAPTER 14

Vector Autoregressions, Unit Roots, and Cointegration	551
What is in this Chapter?	551
14.1 Introduction	551
14.2 Vector autoregressions	551
14.3 Problems with VAR models in practice	553
14.4 Unit roots	554
14.5 Unit root tests	555
The Dickey–Fuller tests	556
The serial correlation problem	556
The low power of unit root tests	557
The DF-GLS test	557
What are the null and alternative hypotheses in unit root tests?	558
Tests with stationarity as null	559
Confirmatory analysis	560
Panel data unit root tests	561
Structural change and unit roots	562
14.6 Cointegration	563
14.7 The cointegrating regression	564
14.8 Vector autoregressions and cointegration	567
14.9 Cointegration and error correction models	571
14.10 Tests for cointegration	571
14.11 Cointegration and testing of the REH and MEH	572
14.12 A summary assessment of cointegration	574
Summary	575
Exercises	576
References	579

CHAPTER 15

Panel Data Analysis	583
What is in this Chapter?	583
15.1 Introduction	583
15.2 The LSDV or fixed effects model	584
Illustrative example: Fixed effect estimation	585
15.3 The random effects model	586
15.4 Fixed effects versus random effects	589
Hausman test	589
Breusch and Pagan test	590
Tests for serial correlation	590
15.5 Dynamic panel data models	591
15.6 Panel data models with correlated effects and simultaneity	593
15.7 Errors in variables in panel data	595

15.8	The SUR model	597
15.9	The random coefficient model	597
	Summary	599
	References	599
CHAPTER 16		
	Small-Sample Inference: Resampling Methods	601
	What is in this Chapter?	601
16.1	Introduction	601
16.2	Monte Carlo methods	602
	More efficient Monte Carlo methods	603
	Response surfaces	603
16.3	Resampling methods: Jackknife and bootstrap	603
	Some illustrative examples	604
	Other issues relating to the bootstrap	605
16.4	Bootstrap confidence intervals	605
16.5	Hypothesis testing with the bootstrap	606
16.6	Bootstrapping residuals versus bootstrapping the data	607
16.7	Non-IID errors and nonstationary models	607
	Heteroskedasticity and autocorrelation	607
	Unit root tests based on the bootstrap	608
	Cointegration tests	608
	Miscellaneous other applications	608
	Summary	609
	References	609
	Appendix	611
	Index	621