Contents

	Preface	xxi
	Acknowledgments	xxix
I	**INTRODUCTION AND BACKGROUND**	1
1	**Introduction**	3
1.1	Causal Relationships and Ceteris Paribus Analysis	3
1.2	Stochastic Setting and Asymptotic Analysis	4
	1.2.1 Data Structures	4
	1.2.2 Asymptotic Analysis	7
1.3	Some Examples	7
1.4	Why Not Fixed Explanatory Variables?	9
2	**Conditional Expectations and Related Concepts in Econometrics**	13
2.1	Role of Conditional Expectations in Econometrics	13
2.2	Features of Conditional Expectations	14
	2.2.1 Definition and Examples	14
	2.2.2 Partial Effects, Elasticities, and Semielasticities	15
	2.2.3 Error Form of Models of Conditional Expectations	18
	2.2.4 Some Properties of Conditional Expectations	19
	2.2.5 Average Partial Effects	22
2.3	Linear Projections	25
	Problems	27
	Appendix 2A	30
	2.A.1 Properties of Conditional Expectations	30
	2.A.2 Properties of Conditional Variances and Covariances	32
	2.A.3 Properties of Linear Projections	34
3	**Basic Asymptotic Theory**	37
3.1	Convergence of Deterministic Sequences	37
3.2	Convergence in Probability and Boundedness in Probability	38
3.3	Convergence in Distribution	40
3.4	Limit Theorems for Random Samples	41
3.5	Limiting Behavior of Estimators and Test Statistics	42
	3.5.1 Asymptotic Properties of Estimators	42
	3.5.2 Asymptotic Properties of Test Statistics	45
	Problems	47

II	**LINEAR MODELS**	51
4	**Single-Equation Linear Model and Ordinary Least Squares Estimation**	53
4.1	Overview of the Single-Equation Linear Model	53
4.2	Asymptotic Properties of Ordinary Least Squares	55
	4.2.1 Consistency	56
	4.2.2 Asymptotic Inference Using Ordinary Least Squares	59
	4.2.3 Heteroskedasticity-Robust Inference	60
	4.2.4 Lagrange Multiplier (Score) Tests	62
4.3	Ordinary Least Squares Solutions to the Omitted Variables Problem	65
	4.3.1 Ordinary Least Squares Ignoring the Omitted Variables	65
	4.3.2 Proxy Variable–Ordinary Least Squares Solution	67
	4.3.3 Models with Interactions in Unobservables: Random Coefficient Models	73
4.4	Properties of Ordinary Least Squares under Measurement Error	76
	4.4.1 Measurement Error in the Dependent Variable	76
	4.4.2 Measurement Error in an Explanatory Variable	78
	Problems	82
5	**Instrumental Variables Estimation of Single-Equation Linear Models**	89
5.1	Instrumental Variables and Two-Stage Least Squares	89
	5.1.1 Motivation for Instrumental Variables Estimation	89
	5.1.2 Multiple Instruments: Two-Stage Least Squares	96
5.2	General Treatment of Two-Stage Least Squares	98
	5.2.1 Consistency	98
	5.2.2 Asymptotic Normality of Two-Stage Least Squares	101
	5.2.3 Asymptotic Efficiency of Two-Stage Least Squares	103
	5.2.4 Hypothesis Testing with Two-Stage Least Squares	104
	5.2.5 Heteroskedasticity-Robust Inference for Two-Stage Least Squares	106
	5.2.6 Potential Pitfalls with Two-Stage Least Squares	107
5.3	IV Solutions to the Omitted Variables and Measurement Error Problems	112
	5.3.1 Leaving the Omitted Factors in the Error Term	112
	5.3.2 Solutions Using Indicators of the Unobservables	112
	Problems	115

6	**Additional Single-Equation Topics**	123
6.1	Estimation with Generated Regressors and Instruments	123
	6.1.1 Ordinary Least Squares with Generated Regressors	123
	6.1.2 Two-Stage Least Squares with Generated Instruments	124
	6.1.3 Generated Instruments and Regressors	125
6.2	Control Function Approach to Endogeneity	126
6.3	Some Specification Tests	129
	6.3.1 Testing for Endogeneity	129
	6.3.2 Testing Overidentifying Restrictions	134
	6.3.3 Testing Functional Form	137
	6.3.4 Testing for Heteroskedasticity	138
6.4	Correlated Random Coefficient Models	141
	6.4.1 When Is the Usual IV Estimator Consistent?	142
	6.4.2 Control Function Approach	145
6.5	Pooled Cross Sections and Difference-in-Differences Estimation	146
	6.5.1 Pooled Cross Sections over Time	146
	6.5.2 Policy Analysis and Difference-in-Differences Estimation	147
	Problems	152
	Appendix 6A	157
7	**Estimating Systems of Equations by Ordinary Least Squares and Generalized Least Squares**	161
7.1	Introduction	161
7.2	Some Examples	161
7.3	System Ordinary Least Squares Estimation of a Multivariate Linear System	166
	7.3.1 Preliminaries	166
	7.3.2 Asymptotic Properties of System Ordinary Least Squares	167
	7.3.3 Testing Multiple Hypotheses	172
7.4	Consistency and Asymptotic Normality of Generalized Least Squares	173
	7.4.1 Consistency	173
	7.4.2 Asymptotic Normality	175
7.5	Feasible Generalized Least Squares	176
	7.5.1 Asymptotic Properties	176
	7.5.2 Asymptotic Variance of Feasible Generalized Least Squares under a Standard Assumption	180

	7.5.3 Properties of Feasible Generalized Least Squares with (Possibly Incorrect) Restrictions on the Unconditional Variance Matrix	182
7.6	Testing the Use of Feasible Generalized Least Squares	183
7.7	Seemingly Unrelated Regressions, Revisited	185
	7.7.1 Comparison between Ordinary Least Squares and Feasible Generalized Least Squares for Seemingly Unrelated Regressions Systems	185
	7.7.2 Systems with Cross Equation Restrictions	188
	7.7.3 Singular Variance Matrices in Seemingly Unrelated Regressions Systems	189
7.8	Linear Panel Data Model, Revisited	191
	7.8.1 Assumptions for Pooled Ordinary Least Squares	191
	7.8.2 Dynamic Completeness	194
	7.8.3 Note on Time Series Persistence	196
	7.8.4 Robust Asymptotic Variance Matrix	197
	7.8.5 Testing for Serial Correlation and Heteroskedasticity after Pooled Ordinary Least Squares	198
	7.8.6 Feasible Generalized Least Squares Estimation under Strict Exogeneity	200
	Problems	202
8	**System Estimation by Instrumental Variables**	207
8.1	Introduction and Examples	207
8.2	General Linear System of Equations	210
8.3	Generalized Method of Moments Estimation	213
	8.3.1 General Weighting Matrix	213
	8.3.2 System Two-Stage Least Squares Estimator	216
	8.3.3 Optimal Weighting Matrix	217
	8.3.4 The Generalized Method of Moments Three-Stage Least Squares Estimator	219
8.4	Generalized Instrumental Variables Estimator	222
	8.4.1 Derivation of the Generalized Instrumental Variables Estimator and Its Asymptotic Properties	222
	8.4.2 Comparison of Generalized Method of Moment, Generalized Instrumental Variables, and the Traditional Three-Stage Least Squares Estimator	224

8.5	Testing Using Generalized Method of Moments	226
	8.5.1 Testing Classical Hypotheses	226
	8.5.2 Testing Overidentification Restrictions	228
8.6	More Efficient Estimation and Optimal Instruments	229
8.7	Summary Comments on Choosing an Estimator	232
	Problems	233
9	**Simultaneous Equations Models**	239
9.1	Scope of Simultaneous Equations Models	239
9.2	Identification in a Linear System	241
	9.2.1 Exclusion Restrictions and Reduced Forms	241
	9.2.2 General Linear Restrictions and Structural Equations	245
	9.2.3 Unidentified, Just Identified, and Overidentified Equations	251
9.3	Estimation after Identification	252
	9.3.1 Robustness-Efficiency Trade-off	252
	9.3.2 When Are 2SLS and 3SLS Equivalent?	254
	9.3.3 Estimating the Reduced Form Parameters	255
9.4	Additional Topics in Linear Simultaneous Equations Methods	256
	9.4.1 Using Cross Equation Restrictions to Achieve Identification	256
	9.4.2 Using Covariance Restrictions to Achieve Identification	257
	9.4.3 Subtleties Concerning Identification and Efficiency in Linear Systems	260
9.5	Simultaneous Equations Models Nonlinear in Endogenous Variables	262
	9.5.1 Identification	262
	9.5.2 Estimation	266
	9.5.3 Control Function Estimation for Triangular Systems	268
9.6	Different Instruments for Different Equations	271
	Problems	273
10	**Basic Linear Unobserved Effects Panel Data Models**	281
10.1	Motivation: Omitted Variables Problem	281
10.2	Assumptions about the Unobserved Effects and Explanatory Variables	285
	10.2.1 Random or Fixed Effects?	285
	10.2.2 Strict Exogeneity Assumptions on the Explanatory Variables	287
	10.2.3 Some Examples of Unobserved Effects Panel Data Models	289

10.3 Estimating Unobserved Effects Models by Pooled Ordinary Least Squares 291
10.4 Random Effects Methods 291
10.4.1 Estimation and Inference under the Basic Random Effects Assumptions 291
10.4.2 Robust Variance Matrix Estimator 297
10.4.3 General Feasible Generalized Least Squares Analysis 298
10.4.4 Testing for the Presence of an Unobserved Effect 299
10.5 Fixed Effects Methods 300
10.5.1 Consistency of the Fixed Effects Estimator 300
10.5.2 Asymptotic Inference with Fixed Effects 304
10.5.3 Dummy Variable Regression 307
10.5.4 Serial Correlation and the Robust Variance Matrix Estimator 310
10.5.5 Fixed Effects Generalized Least Squares 312
10.5.6 Using Fixed Effects Estimation for Policy Analysis 315
10.6 First Differencing Methods 315
10.6.1 Inference 315
10.6.2 Robust Variance Matrix 318
10.6.3 Testing for Serial Correlation 319
10.6.4 Policy Analysis Using First Differencing 320
10.7 Comparison of Estimators 321
10.7.1 Fixed Effects versus First Differencing 321
10.7.2 Relationship between the Random Effects and Fixed Effects Estimators 326
10.7.3 Hausman Test Comparing Random Effects and Fixed Effects Estimators 328
Problems 334

11 More Topics in Linear Unobserved Effects Models 345

11.1 Generalized Method of Moments Approaches to the Standard Linear Unobserved Effects Model 345
11.1.1 Equivalance between GMM 3SLS and Standard Estimators 345
11.1.2 Chamberlain's Approach to Unobserved Effects Models 347
11.2 Random and Fixed Effects Instrumental Variables Methods 349
11.3 Hausman and Taylor–Type Models 358
11.4 First Differencing Instrumental Variables Methods 361

11.5	Unobserved Effects Models with Measurement Error	365
11.6	Estimation under Sequential Exogeneity	368
	11.6.1 General Framework	368
	11.6.2 Models with Lagged Dependent Variables	371
11.7	Models with Individual-Specific Slopes	374
	11.7.1 Random Trend Model	375
	11.7.2 General Models with Individual-Specific Slopes	377
	11.7.3 Robustness of Standard Fixed Effects Methods	382
	11.7.4 Testing for Correlated Random Slopes	384
	Problems	387
III	**GENERAL APPROACHES TO NONLINEAR ESTIMATION**	395
12	**M-Estimation, Nonlinear Regression, and Quantile Regression**	397
12.1	Introduction	397
12.2	Identification, Uniform Convergence, and Consistency	401
12.3	Asymptotic Normality	405
12.4	Two-Step M-Estimators	409
	12.4.1 Consistency	410
	12.4.2 Asymptotic Normality	411
12.5	Estimating the Asymptotic Variance	413
	12.5.1 Estimation without Nuisance Parameters	413
	12.5.2 Adjustments for Two-Step Estimation	418
12.6	Hypothesis Testing	420
	12.6.1 Wald Tests	420
	12.6.2 Score (or Lagrange Multiplier) Tests	421
	12.6.3 Tests Based on the Change in the Objective Function	428
	12.6.4 Behavior of the Statistics under Alternatives	430
12.7	Optimization Methods	431
	12.7.1 Newton-Raphson Method	432
	12.7.2 Berndt, Hall, Hall, and Hausman Algorithm	433
	12.7.3 Generalized Gauss-Newton Method	434
	12.7.4 Concentrating Parameters out of the Objective Function	435
12.8	Simulation and Resampling Methods	436
	12.8.1 Monte Carlo Simulation	436
	12.8.2 Bootstrapping	438

12.9 Multivariate Nonlinear Regression Methods 442
12.9.1 Multivariate Nonlinear Least Squares 442
12.9.2 Weighted Multivariate Nonlinear Least Squares 444
12.10 Quantile Estimation 449
12.10.1 Quantiles, the Estimation Problem, and Consistency 449
12.10.2 Asymptotic Inference 454
12.10.3 Quantile Regression for Panel Data 459
Problems 462

13 Maximum Likelihood Methods 469

13.1 Introduction 469
13.2 Preliminaries and Examples 470
13.3 General Framework for Conditional Maximum Likelihood Estimation 473
13.4 Consistency of Conditional Maximum Likelihood Estimation 475
13.5 Asymptotic Normality and Asymptotic Variance Estimation 476
13.5.1 Asymptotic Normality 476
13.5.2 Estimating the Asymptotic Variance 479
13.6 Hypothesis Testing 481
13.7 Specification Testing 482
13.8 Partial (or Pooled) Likelihood Methods for Panel Data 485
13.8.1 Setup for Panel Data 486
13.8.2 Asymptotic Inference 490
13.8.3 Inference with Dynamically Complete Models 492
13.9 Panel Data Models with Unobserved Effects 494
13.9.1 Models with Strictly Exogenous Explanatory Variables 494
13.9.2 Models with Lagged Dependent Variables 497
13.10 Two-Step Estimators Involving Maximum Likelihood 499
13.10.1 Second-Step Estimator Is Maximum Likelihood Estimator 499
13.10.2 Surprising Efficiency Result When the First-Step Estimator Is Conditional Maximum Likelihood Estimator 500
13.11 Quasi-Maximum Likelihood Estimation 502
13.11.1 General Misspecification 502
13.11.2 Model Selection Tests 505
13.11.3 Quasi-Maximum Likelihood Estimation in the Linear Exponential Family 509

13.11.4 Generalized Estimating Equations for Panel Data 514
Problems 517
Appendix 13A 522

14 Generalized Method of Moments and Minimum Distance Estimation 525

14.1 Asymptotic Properties of Generalized Method of Moments 525
14.2 Estimation under Orthogonality Conditions 530
14.3 Systems of Nonlinear Equations 532
14.4 Efficient Estimation 538
14.4.1 General Efficiency Framework 538
14.4.2 Efficiency of Maximum Likelihood Estimator 540
14.4.3 Efficient Choice of Instruments under Conditional Moment Restrictions 542
14.5 Classical Minimum Distance Estimation 545
14.6 Panel Data Applications 547
14.6.1 Nonlinear Dynamic Models 547
14.6.2 Minimum Distance Approach to the Unobserved Effects Model 549
14.6.3 Models with Time-Varying Coefficients on the Unobserved Effects 551
Problems 555
Appendix 14A 558

IV NONLINEAR MODELS AND RELATED TOPICS 559

15 Binary Response Models 561

15.1 Introduction 561
15.2 Linear Probability Model for Binary Response 562
15.3 Index Models for Binary Response: Probit and Logit 565
15.4 Maximum Likelihood Estimation of Binary Response Index Models 567
15.5 Testing in Binary Response Index Models 569
15.5.1 Testing Multiple Exclusion Restrictions 570
15.5.2 Testing Nonlinear Hypotheses about β 571
15.5.3 Tests against More General Alternatives 571

15.6	Reporting the Results for Probit and Logit	573
15.7	Specification Issues in Binary Response Models	582
	15.7.1 Neglected Heterogeneity	582
	15.7.2 Continuous Endogenous Explanatory Variables	585
	15.7.3 Binary Endogenous Explanatory Variable	594
	15.7.4 Heteroskedasticity and Nonnormality in the Latent Variable Model	599
	15.7.5 Estimation under Weaker Assumptions	604
15.8	Binary Response Models for Panel Data	608
	15.8.1 Pooled Probit and Logit	609
	15.8.2 Unobserved Effects Probit Models under Strict Exogeneity	610
	15.8.3 Unobserved Effects Logit Models under Strict Exogeneity	619
	15.8.4 Dynamic Unobserved Effects Models	625
	15.8.5 Probit Models with Heterogeneity and Endogenous Explanatory Variables	630
	15.8.6 Semiparametric Approaches	632
	Problems	635
16	**Multinomial and Ordered Response Models**	643
16.1	Introduction	643
16.2	Multinomial Response Models	643
	16.2.1 Multinomial Logit	643
	16.2.2 Probabilistic Choice Models	646
	16.2.3 Endogenous Explanatory Variables	651
	16.2.4 Panel Data Methods	653
16.3	Ordered Response Models	655
	16.3.1 Ordered Logit and Ordered Probit	655
	16.3.2 Specification Issues in Ordered Models	658
	16.3.3 Endogenous Explanatory Variables	660
	16.3.4 Panel Data Methods	662
	Problems	663
17	**Corner Solution Responses**	667
17.1	Motivation and Examples	667
17.2	Useful Expressions for Type I Tobit	671

17.3 Estimation and Inference with the Type I Tobit Model 676
17.4 Reporting the Results 677
17.5 Specification Issues in Tobit Models 680
17.5.1 Neglected Heterogeneity 680
17.5.2 Endogenous Explanatory Models 681
17.5.3 Heteroskedasticity and Nonnormality in the Latent Variable Model 685
17.5.4 Estimating Parameters with Weaker Assumptions 687
17.6 Two-Part Models and Type II Tobit for Corner Solutions 690
17.6.1 Truncated Normal Hurdle Model 692
17.6.2 Lognormal Hurdle Model and Exponential Conditional Mean 694
17.6.3 Exponential Type II Tobit Model 697
17.7 Two-Limit Tobit Model 703
17.8 Panel Data Methods 705
17.8.1 Pooled Methods 705
17.8.2 Unobserved Effects Models under Strict Exogeneity 707
17.8.3 Dynamic Unobserved Effects Tobit Models 713
Problems 715

18 Count, Fractional, and Other Nonnegative Responses 723

18.1 Introduction 723
18.2 Poisson Regression 724
18.2.1 Assumptions Used for Poisson Regression and Quantities of Interest 724
18.2.2 Consistency of the Poisson QMLE 727
18.2.3 Asymptotic Normality of the Poisson QMLE 728
18.2.4 Hypothesis Testing 732
18.2.5 Specification Testing 734
18.3 Other Count Data Regression Models 736
18.3.1 Negative Binomial Regression Models 736
18.3.2 Binomial Regression Models 739
18.4 Gamma (Exponential) Regression Model 740
18.5 Endogeneity with an Exponential Regression Function 742
18.6 Fractional Responses 748

	18.6.1 Exogenous Explanatory Variables	748
	18.6.2 Endogenous Explanatory Variables	753
18.7	Panel Data Methods	755
	18.7.1 Pooled QMLE	756
	18.7.2 Specifying Models of Conditional Expectations with Unobserved Effects	758
	18.7.3 Random Effects Methods	759
	18.7.4 Fixed Effects Poisson Estimation	762
	18.7.5 Relaxing the Strict Exogeneity Assumption	764
	18.7.6 Fractional Response Models for Panel Data	766
	Problems	769
19	**Censored Data, Sample Selection, and Attrition**	777
19.1	Introduction	777
19.2	Data Censoring	778
	19.2.1 Binary Censoring	780
	19.2.2 Interval Coding	783
	19.2.3 Censoring from Above and Below	785
19.3	Overview of Sample Selection	790
19.4	When Can Sample Selection Be Ignored?	792
	19.4.1 Linear Models: Estimation by OLS and 2SLS	792
	19.4.2 Nonlinear Models	798
19.5	Selection on the Basis of the Response Variable: Truncated Regression	799
19.6	Incidental Truncation: A Probit Selection Equation	802
	19.6.1 Exogenous Explanatory Variables	802
	19.6.2 Endogenous Explanatory Variables	809
	19.6.3 Binary Response Model with Sample Selection	813
	19.6.4 An Exponential Response Function	814
19.7	Incidental Truncation: A Tobit Selection Equation	815
	19.7.1 Exogenous Explanatory Variables	815
	19.7.2 Endogenous Explanatory Variables	817
	19.7.3 Estimating Structural Tobit Equations with Sample Selection	819
19.8	Inverse Probability Weighting for Missing Data	821

19.9 Sample Selection and Attrition in Linear Panel Data Models 827
19.9.1 Fixed and Random Effects Estimation with Unbalanced Panels 828
19.9.2 Testing and Correcting for Sample Selection Bias 832
19.9.3 Attrition 837
Problems 845

20 Stratified Sampling and Cluster Sampling 853

20.1 Introduction 853
20.2 Stratified Sampling 854
20.2.1 Standard Stratified Sampling and Variable Probability Sampling 854
20.2.2 Weighted Estimators to Account for Stratification 856
20.2.3 Stratification Based on Exogenous Variables 861
20.3 Cluster Sampling 863
20.3.1 Inference with a Large Number of Clusters and Small Cluster Sizes 864
20.3.2 Cluster Samples with Unit-Specific Panel Data 876
20.3.3 Should We Apply Cluster-Robust Inference with Large Group Sizes? 883
20.3.4 Inference When the Number of Clusters Is Small 884
20.4 Complex Survey Sampling 894
Problems 899

21 Estimating Average Treatment Effects 903

21.1 Introduction 903
21.2 A Counterfactual Setting and the Self-Selection Problem 904
21.3 Methods Assuming Ignorability (or Unconfoundedness) of Treatment 908
21.3.1 Identification 911
21.3.2 Regression Adjustment 915
21.3.3 Propensity Score Methods 920
21.3.4 Combining Regression Adjustment and Propensity Score Weighting 930
21.3.5 Matching Methods 934

21.4	Instrumental Variables Methods	937
	21.4.1 Estimating the Average Treatment Effect Using IV	937
	21.4.2 Correction and Control Function Approaches	945
	21.4.3 Estimating the Local Average Treatment Effect by IV	951
21.5	Regression Discontinuity Designs	954
	21.5.1 The Sharp Regression Discontinuity Design	954
	21.5.2 The Fuzzy Regression Discontinuity Design	957
	21.5.3 Unconfoundedness versus the Fuzzy Regression Discontinuity	959
21.6	Further Issues	960
	21.6.1 Special Considerations for Responses with Discreteness or Limited Range	960
	21.6.2 Multivalued Treatments	961
	21.6.3 Multiple Treatments	964
	21.6.4 Panel Data	968
	Problems	975
22	**Duration Analysis**	983
22.1	Introduction	983
22.2	Hazard Functions	984
	22.2.1 Hazard Functions without Covariates	984
	22.2.2 Hazard Functions Conditional on Time-Invariant Covariates	988
	22.2.3 Hazard Functions Conditional on Time-Varying Covariates	989
22.3	Analysis of Single-Spell Data with Time-Invariant Covariates	991
	22.3.1 Flow Sampling	992
	22.3.2 Maximum Likelihood Estimation with Censored Flow Data	993
	22.3.3 Stock Sampling	1000
	22.3.4 Unobserved Heterogeneity	1003
22.4	Analysis of Grouped Duration Data	1010
	22.4.1 Time-Invariant Covariates	1011
	22.4.2 Time-Varying Covariates	1015
	22.4.3 Unobserved Heterogeneity	1017

22.5 Further Issues 1018
22.5.1 Cox's Partial Likelihood Method for the Proportional Hazard Model 1018
22.5.2 Multiple-Spell Data 1018
22.5.3 Competing Risks Models 1019
Problems 1019

References 1025
Index 1045