CONTENTS

LIST OF TABLES xvii

1 Introduction 1

1.1 Types of panel variables and data 1

1.2 Virtues of panel data: Transformations 3

1.3 Panel data versus experimental data 8

1.4 Other virtues of panel data and some limitations 9

1.5 Overview 11

2 Regression analysis: Fixed effects models 14

2.1 Simple regression model: One-way heterogeneity 15

2.1.1 Individual-specific intercepts and coefficients 15

2.1.2 Individual-specific intercepts, common coefficient 16

2.1.3 Homogeneous benchmark model 19

2.1.4 How are the estimators related? 21

2.2 Multiple regression model: One-way heterogeneity 23

2.2.1 Individual-specific intercepts and coefficients 23

2.2.2 Individual-specific intercepts, common coefficients 25

2.2.3 Homogeneous benchmark model 27

2.2.4 How are the estimators related? 30

2.3 Simple regression model: Two-way heterogeneity 33

2.3.1 Individual- and period-specific intercepts 33

2.3.2 Homogeneous benchmark model 36

2.3.3 How are the estimators related? 39

2.4 Multiple regression model: Two-way heterogeneity 41

2.4.1 An excursion into Kronecker-products: Definition 41

2.4.2 Matrix formulae with Kronecker-products: Examples 42

2.4.3 Panel data 'operators': Bilinear and quadratic forms 43

2.4.4 Individual- and period-specific intercepts 46

2.4.5 Homogeneous benchmark model 49

2.4.6 How are the estimators related? 50

2.5 Testing for fixed heterogeneity 53

2.5.1 One-way intercept and coefficient heterogeneity 54

2.5.2 Two-way intercept heterogeneity 55

Appendix 2A. Properties of GLS 56
Appendix 2B. Kronecker-product operations: Examples 61

3 Regression analysis: Random effects models 65
3.1 One-way variance component model 66
3.1.1 Basic model 66
3.1.2 Some implications 67
3.2 GLS with known variance components 68
3.2.1 The GLS problem and its reformulation 68
3.2.2 GLS as OLS on transformed data 70
3.2.3 Four estimators and a synthesis 71
3.3 GLS with unknown variance components 74
3.3.1 The problem 74
3.3.2 'Estimation' of variance components from disturbances 75
3.3.3 Synthesis: Stepwise estimation 76
3.3.4 Testing for random one-way heterogeneity 78
3.3.5 Disturbance serial correlation and heteroskedasticity 79
3.4 Maximum Likelihood estimation 81
3.4.1 The problem 81
3.4.2 Simplifying the log-likelihood 82
3.4.3 Stepwise solution 83
3.5 Two-way variance component model 84
3.5.1 Basic assumptions 84
3.5.2 Some implications 85
3.6 GLS with known variance components 87
3.6.1 The problem and its reformulation 87
3.6.2 GLS as OLS on transformed data 88
3.6.3 Five estimators and a synthesis 89
3.7 GLS with unknown variance components 92
3.7.1 The problem 92
3.7.2 'Estimation' of variance components from disturbances 93
3.7.3 Synthesis: Stepwise estimation 95
3.7.4 Testing for random two-way heterogeneity 95
3.8 Maximum Likelihood estimation 96
3.8.1 The problem 96
3.8.2 Simplifying the log-likelihood 96
3.8.3 Stepwise solution 97
Appendix 3A. Two theorems related to GLS Estimation 99
Appendix 3B. Efficiency in the one-regressor case 100

4 Regression analysis with heterogeneous coefficients 107
4.1 Introduction 107
4.2 Fixed coefficient models 108
4.2.1 Single-regressor models without intercept 108
4.2.2 Synthesis 113
4.2.3 Including intercepts 114
4.2.4 Multi-regressor models 115
4.3 Random coefficient models 116
4.3.1 Single-regressor models and their estimators 117
4.3.2 Multi-regressor model and its GLS estimator 123
4.3.3 Coefficient prediction 127
Appendix 4A. Matrix inversion and matrix products: Useful results 132
Appendix 4B. A reinterpretation of the GLS estimator 134

5 Regression analysis with unidimensional variables 136
5.1 Introduction 136
5.2 Properties of moment matrices in unidimensional variables 137
5.2.1 Basics 137
5.2.2 Implications for regression analysis 138
5.3 One-way fixed effects models 139
5.3.1 Simple model with two regressors 139
5.3.2 The general case 143
5.4 One-way random effects models 145
5.4.1 Simple model with two regressors 145
5.4.2 The general case 147
5.5 Two-way models 150
5.5.1 Fixed effects two-regressor model 150
5.5.2 Random effects two-regressor model 154

6 Latent heterogeneity correlated with regressors 157
6.1 Introduction 157
6.2 Models with only two-dimensional regressors 158
6.2.1 A simple case 158
6.2.2 Generalization 161
6.3 Models with time-invariant regressors 164
6.3.1 A simple case 164
6.3.2 Generalization 166
6.3.3 The Wu–Hausman test 168
6.4 Exploiting instruments: Initial attempts 171

6.4.1 A simple example 171
6.4.2 Generalization 172
6.4.3 Reinterpretation using a recursive system 173
6.5 Exploiting instruments: Further steps 175
6.5.1 Choosing instruments: General remarks 176
6.5.2 Estimation in case of exact identification 177
6.5.3 Estimation in case of overidentification 181
Appendix 6A. Reinterpretation: Block-recursive system 185
Appendix 6B. Proof of consistency in case of exact identification 187

7 Measurement errors 189
7.1 Introduction 189
7.2 A homogeneous model 190
7.2.1 Model setup: Basic probability limits 191
7.2.2 Basic (disaggregate) estimators 192
7.2.3 Aggregate estimators 194
7.2.4 Combining inconsistent estimators 198
7.2.5 Difference estimators 199
7.3 Extension I: Heterogeneity in the equation 202
7.3.1 Model 202
7.3.2 Between and within estimation 203
7.3.3 Combining inconsistent estimators 204
7.4 Extension II: Heterogeneity in the error 206
7.4.1 Model 206
7.4.2 Estimation 206
7.4.3 Combining inconsistent estimators 209
7.5 Extension III: Memory in the error 209
7.5.1 One-component error with memory 210
7.5.2 Three-component specification 212
7.6 Generalized Method of Moments estimators 215
7.6.1 Generalities on the GMM 215
7.6.2 GMM-estimation of the equation in differences 217
7.6.3 Extensions and modifications 222
7.7 Concluding remarks 223
Appendix 7A. Asymptotics for aggregate estimators 224

8 Dynamic models 227
8.1 Introduction 227
8.2 Fixed effects AR-models 229

8.2.1 A simple model 229
8.2.2 Within estimation: Other difference transformations 230
8.2.3 Equation in differences: Simple estimation by level IVs 234
8.2.4 Equation in differences: System GMM with level IVs 236
8.2.5 Equation in levels: Simple estimation by difference IVs 238
8.2.6 Equation in levels: System GMM with difference IVs 238
8.2.7 *Panel unit roots and panel co-integration* 240
8.2.8 Bias correction 242
8.3 Fixed effects AR-models with exogenous variables 242
8.3.1 Model 242
8.3.2 Equation in differences: System GMM with level IVs 244
8.3.3 Equation in levels: System GMM with difference IVs 246
8.3.4 GMM estimation: AR-model versus EIV-model 249
8.3.5 Extensions and modifications 250
8.4 Random effects AR(1) models 250
8.4.1 A simple AR(1) model 250
8.4.2 Including exogenous variables 252
8.5 Conclusion 254
Appendix 8A. Within estimation of the AR coefficient: Asymptotics 255
Appendix 8B. Autocovariances and correlograms of y_{it} and Δy_{it} 257

9 Analysis of discrete response 261
9.1 Introduction 261
9.2 Binomial models: Fixed heterogeneity 262
9.2.1 Simple binomial model: Logit and probit parameterizations 263
9.2.2 Binomial model with full fixed heterogeneity 269
9.2.3 Binomial model with fixed intercept heterogeneity 270
9.2.4 Fixed intercept heterogeneity, small T and conditional ML 270
9.3 Binomial model: Random heterogeneity 275
9.3.1 Model 275
9.3.2 ML estimation 277
9.4 Multinomial model: Other extensions 278
9.4.1 Model 279
9.4.2 Likelihood function and ML estimation 281
9.4.3 ML estimation conditional on the individual effects 282
Appendix 9A. The general binomial model: ML Estimation 283
Appendix 9B. The multinomial logit model: Conditional ML estimation 285

10 Unbalanced panel data 287
10.1 Introduction 287
10.2 Basic model and notation 289
10.2.1 Formalization of the non-balance 289
10.2.2 Individual-specific, period-specific, and global means 291
10.3 The fixed effects case 292
10.3.1 The one-regressor case 292
10.3.2 Between- and OLS-estimation in the one-regressor case 294
10.3.3 The multi-regressor case 296
10.3.4 Between- and OLS-estimation in the general case 297
10.4 The random effects case 298
10.4.1 The one-regressor case 299
10.4.2 The multi-regressor case 302
10.4.3 Estimating the variance components: FGLS 303
10.5 Maximum Likelihood estimation 306
10.5.1 A convenient notation and reordering 306
10.5.2 Reformulating the model 307
10.5.3 'Estimating' the variance components from disturbances 308
10.5.4 The log-likelihood function 309
Appendix 10A. Between estimation: Proofs 312
Appendix 10B. GLS estimation: Proofs 313
Appendix 10C. Estimation of variance components: Details 315

11 Panel data with systematic unbalance 317
11.1 Introduction: Observation rules 317
11.2 Truncation and censoring: Baseline models 320
11.2.1 Basic concepts and point of departure 320
11.2.2 Truncation and truncation bias 321
11.2.3 Censoring and censoring bias 324
11.2.4 Maximum Likelihood in case of truncation 326
11.2.5 Maximum Likelihood in case of censoring 327
11.3 Truncation and censoring: Heterogeneity 328
11.3.1 Truncation: Maximum Likelihood estimation 329
11.3.2 Censoring: Maximum Likelihood estimation 330
11.4 Truncation and censoring: Cross-effects 332
11.4.1 Motivation and common model elements 333
11.4.2 Implied theoretical regressions for the observable variables 334
11.4.3 Stepwise estimation: Examples 335
11.4.4 Extensions: ML estimation—Heterogeneity 338

Appendix 11A. On truncated normal distributions 340
Appendix 11B. Partial effects in censoring models 345

12 Multi-equation models 347
12.1 Introduction 347
12.2 Regression system with one-way random effects 348
12.2.1 Model description 349
12.2.2 Matrix version: Two formulations 349
12.2.3 GLS estimation 353
12.2.4 Maximum Likelihood estimation 355
12.3 Regression system with two-way random effects 358
12.3.1 Model description 358
12.3.2 Matrix version: Two formulations 359
12.3.3 GLS estimation 362
12.4 Interdependent model: One-equation estimation 362
12.4.1 Within and between two-stage least squares 364
12.4.2 Variance components two-stage least squares 366
12.5 Interdependent model: Joint estimation 371
12.5.1 Within and between three-stage least squares 372
12.5.2 Variance components three-stage least squares 374
Appendix 12A. Estimation of error component covariance matrices 376
Appendix 12B. Matrix differentiation: Useful results 378
Appendix 12C. Estimator covariance matrices in interdependent models 379

REFERENCES 383
INDEX 394