

Analysis of Panel Data

Third Edition

Cheng Hsiao

University of Southern California

ggg **CAMBRIDGE**
UNIVERSITY PRESS

Contents

<i>Preface to the Third Edition</i>	<i>page xvii</i>
<i>Preface to the Second Edition</i>	xix
<i>Preface to the First Edition</i>	xxi
1 Introduction	1
1.1 Introduction	1
1.2 Advantages of Panel Data	4
1.3 Issues Involved in Utilizing Panel Data	10
1.3.1 Unobserved Heterogeneity across Individuals and over Time	10
1.3.2 Incidental Parameters and Multidimensional Statistics	13
1.3.3 Sample Attrition	13
1.4 Outline of the Monograph	14
2 Homogeneity Tests for Linear Regression Models (Analysis of Covariance)	17
2.1 Introduction	17
2.2 Analysis of Covariance	18
2.3 An Example	24
3 Simple Regression with Variable Intercepts	31
3.1 Introduction	31
3.2 Fixed-Effects Models: Least-Squares Dummy Variable Approach	34
3.3 Random Effects Models: Estimation of Variance- Components Models	39
3.3.1 Covariance Estimation	40
3.3.2 Generalized Least-Squares (GLS) Estimation	41
3.3.3 Maximum-Likelihood Estimation	45

3.4	Fixed Effects or Random Effects	47
3.4.1	An Example	47
3.4.2	Conditional Inference or Unconditional (Marginal) Inference	48
3.5	Tests for Misspecification	56
3.6	Models with Time- and/or Individual-Invariant Explanatory Variables and Both Individual- and Time-Specific Effects	58
3.6.1	Estimation of Models with Individual-Specific Variables	58
3.6.2	Estimation of Models with Both Individual and Time Effects	61
3.7	Heteroscedasticity and Autocorrelation	64
3.7.1	Heteroscedasticity	64
3.7.2	Models with Serially Correlated Errors	65
3.7.3	Heteroscedasticity Autocorrelation Consistent Estimator for the Covariance Matrix of the CV Estimator	68
3.8	Models with Arbitrary Error Structure – Chamberlain n -Approach	69
	Appendix 3A: Consistency and Asymptotic Normality of the Minimum-Distance Estimator	75
	Appendix 3B: Characteristic Vectors and the Inverse of the Variance-Covariance Matrix of a Three-Component Model	77
4	Dynamic Models with Variable Intercepts	80
4.1	Introduction	80
4.2	The CV Estimator	82
4.3	Random-Effects Models	84
4.3.1	Bias in the OLS Estimator	85
4.3.2	Model Formulation	86
4.3.3	Estimation of Random-Effects Models	89
4.3.4	Testing Some Maintained Hypotheses on Initial Conditions	106
4.3.5	Simulation Evidence	107
4.4	An Example	108
4.5	Fixed-Effects Models	111
4.5.1	Transformed Likelihood Approach	112
4.5.2	Minimum Distance Estimator	114
4.5.3	Relations between the Likelihood-Based Estimator and the GMM	116
4.5.4	Issues of Random versus Fixed-Effects Specification	119

4.6	Estimation of Dynamic Models with Arbitrary Serial Correlations in the Residuals	121
4.7	Models with Both Individual- and Time-Specific Additive Effects	122
	Appendix 4A: Derivation of the Asymptotic Covariance Matrix of Feasible MDE	129
	Appendix 4B: Large N and T Asymptotics	130
5	Static Simultaneous-Equations Models	136
5.1	Introduction	136
5.2	Joint Generalized Least-Squares Estimation Technique	140
5.3	Estimation of Structural Equations	144
5.3.1	Estimation of a Single Equation in the Structural Model	144
5.3.2	Estimation of the Complete Structural System	149
5.4	Triangular System	152
5.4.1	Identification	153
5.4.2	Estimation	155
5.4.3	An Example	162
	Appendix 5A	164
6	Variable-Coefficient Models	167
6.1	Introduction	167
6.2	Coefficients that Vary over Cross-Sectional Units	170
6.2.1	Fixed-Coefficient Model	170
6.2.2	Random-Coefficient Model	172
6.3	Coefficients that Vary over Time and Cross-Sectional Units	180
6.3.1	The Model	180
6.3.2	Fixed-Coefficient Model	182
6.3.3	Random-Coefficient Model	183
6.4	Coefficients that Evolve over Time	186
6.4.1	The Model	186
6.4.2	Predicting β_t by the Kalman Filter	188
6.4.3	Maximum-Likelihood Estimation	191
6.4.4	Tests for Parameter Constancy	192
6.5	Coefficients that Are Functions of Other Exogenous Variables	193
6.6	A Mixed Fixed-and Random-Coefficients Model	196
6.6.1	Model Formulation	196
6.6.2	A Bayes Solution	198
6.6.3	Random or Fixed Differences?	201
6.7	Dynamic Random-Coefficients Models	206
6.8	Two Examples	212

6.8.1	Liquidity Constraints and Firm Investment Expenditure	212
6.8.2	Aggregate versus Disaggregate Analysis	217
6.9	Correlated Random-Coefficients Models	220
6.9.1	Introduction	220
6.9.2	Identification with Cross-Sectional Data	221
6.9.3	Estimation of the Mean Effects with Panel Data	223
	Appendix 6A: Combination of Two Normal Distributions	228
7	Discrete Data	230
7.1	Introduction	230
7.2	Some Discrete-Response Models for Cross-Sectional Data	230
7.3	Parametric Approach to Static Models with Heterogeneity	235
7.3.1	Fixed-Effects Models	236
7.3.2	Random-Effects Models	242
7.4	Semiparametric Approach to Static Models	246
7.4.1	Maximum Score Estimator	247
7.4.2	A <i>Root-N</i> Consistent Semiparametric Estimator	249
7.5	Dynamic Models	250
7.5.1	The General Model	250
7.5.2	Initial Conditions	252
7.5.3	A Conditional Approach	255
7.5.4	State Dependence versus Heterogeneity	261
7.5.5	Two Examples	264
7.6	Alternative Approaches for Identifying State Dependence	270
7.6.1	Bias-Adjusted Estimator	270
7.6.2	Bounding Parameters	274
7.6.3	Approximate Model	276
8	Sample Truncation and Sample Selection	281
8.1	Introduction	281
8.2	An Example – Nonrandomly Missing Data	292
8.2.1	Introduction	292
8.2.2	A Probability Model of Attrition and Selection Bias	292
8.2.3	Attrition in the Gary Income-Maintenance Experiment	296
8.3	Tobit Models with Random Individual Effects	298
8.4	Fixed-Effects Estimator	299
8.4.1	Pairwise Trimmed Least-Squares and Least Absolute Deviation Estimators for Truncated and Censored Regressions	299

8.4.2	A Semiparametric Two-Step Estimator for the Endogenously Determined Sample Selection Model	311
8.5	An Example: Housing Expenditure	313
8.6	Dynamic Tobit Models	317
8.6.1	Dynamic Censored Models	317
8.6.2	Dynamic Sample Selection Models	324
9	Cross-Sectionally Dependent Panel Data	327
9.1	Issues of Cross-Sectional Dependence	327
9.2	Spatial Approach	329
9.2.1	Introduction	329
9.2.2	Spatial Error Model	332
9.2.3	Spatial Lag Model	333
9.2.4	Spatial Error Models with Individual-Specific Effects	334
9.2.5	Spatial Lag Model with Individual-Specific Effects	335
9.2.6	Spatial Dynamic Panel Data Models	336
9.3	Factor Approach	337
9.4	Group Mean Augmented (Common Correlated Effects) Approach to Control the Impact of Cross-Sectional Dependence	342
9.5	Test of Cross-Sectional Independence	344
9.5.1	Linear Model	344
9.5.2	Limited Dependent-Variable Model	348
9.5.3	An Example – A Housing Price Model of China	350
9.6	A Panel Data Approach for Program Evaluation	352
9.6.1	Introduction	352
9.6.2	Definition of Treatment Effects	352
9.6.3	Cross-Sectional Adjustment Methods	354
9.6.4	Panel Data Approach	359
10	Dynamic System	369
10.1	Panel Vector Autoregressive Models	370
10.1.1	"Homogeneous" Panel VAR Models	370
10.1.2	Heterogeneous Vector Autoregressive Models	377
10.2	Cointegrated Panel Models and Vector Error Correction	379
10.2.1	Properties of Cointegrated Processes	379
10.2.2	Estimation	381
10.3	Unit Root and Cointegration Tests	386
10.3.1	Unit Root Tests	386
10.3.2	Tests of Cointegration	394
10.4	Dynamic Simultaneous Equations Models	397
10.4.1	The Model	397

10.4.2	Likelihood Approach	398
10.4.3	Method of Moments Estimator	401
11	Incomplete Panel Data	403
11.1	Rotating or Randomly Missing Data	403
11.2	Pseudo-Panels (or Repeated Cross-Sectional Data)	408
11.3	Pooling of Single Cross-Sectional and Single Time Series Data	411
11.3.1	Introduction	411
11.3.2	The Likelihood Approach to Pooling Cross-Sectional and Time Series Data	413
11.3.3	An Example	416
11.4	Estimating Distributed Lags in Short Panels	418
11.4.1	Introduction	418
11.4.2	Common Assumptions	419
11.4.3	Identification Using Prior Structure on the Process of the Exogenous Variable	421
11.4.4	Identification Using Prior Structure on the Lag Coefficients	425
11.4.5	Estimation and Testing	428
12	Miscellaneous Topics	430
12.1	Duration Model	430
12.2	Count Data Model	438
12.3	Panel Quantile Regression	445
12.4	Simulation Methods	448
12.5	Data with Multilevel Structures	453
12.6	Errors of Measurement	455
12.7	Nonparametric Panel Data Models	461
13	A Summary View	464
13.1	Benefits of Panel Data	464
13.1.1	Increasing Degrees of Freedom and Lessening the Problem of Multicollinearity	464
13.1.2	Identification and Discrimination between Competing Hypotheses	465
13.1.3	Reducing Estimation Bias	467
13.1.4	Generating More Accurate Predictions for Individual Outcomes	468
13.1.5	Providing Information on Appropriate Level of Aggregation	468
13.1.6	Simplifying Computation and Statistical Inference	469

13.2	Challenges for Panel Data Analysis	469
13.2.1	Modeling Unobserved Heterogeneity	469
13.2.2	Controlling the Impact of Unobserved Heterogeneity in Nonlinear Models	470
13.2.3	Modeling Cross-Sectional Dependence	471
13.2.4	Multidimensional Asymptotics	472
13.2.5	Sample Attrition	472
13.3	A Concluding Remark	473
	<i>References</i>	475
	<i>Author Index</i>	507
	<i>Subject Index</i>	513