

Detailed Contents

List of Tables	xiii
List of Figures	xvii
Preface	xix
1. What the Book Is About	xix
2. New in the Second Edition	xx
3. Acknowledgments	xxii
About the Authors	xxiii
1 Introduction	1
1.1 Observational Studies	3
1.2 History and Development	4
1.3 Randomized Experiments	6
1.3.1 Fisher's Randomized Experiment	6
1.3.2 Types of Randomized Experiments and Statistical Tests	10
1.3.3 Critiques of Social Experimentation	11
1.4 Why and When a Propensity Score Analysis Is Needed	11
1.5 Computing Software Packages	16
1.6 Plan of the Book	17
2 Counterfactual Framework and Assumptions	21
2.1 Causality, Internal Validity, and Threats	22
2.2 Counterfactuals and the Neyman-Rubin Counterfactual Framework	23
2.3 The Ignorable Treatment Assignment Assumption	29
2.4 The Stable Unit Treatment Value Assumption	33
2.5 Methods for Estimating Treatment Effects	34
2.5.1 Design of Observational Study	34
2.5.2 The Seven Models	35
2.5.3 Other Balancing Methods	38
2.5.4 Instrumental Variables Estimator	38
2.5.5 Regression Discontinuity Designs	42
2.6 The Underlying Logic of Statistical Inference	43
2.7 Types of Treatment Effects	48

2.8 Treatment Effect Heterogeneity	53
2.8.1 The Importance of Studying Treatment Effect Heterogeneity	53
2.8.2 Checking the Plausibility of the Unconfoundedness Assumption	55
2.8.3 A Methodological Note About the Hausman Test of Endogeneity	56
2.8.4 Tests of Treatment Effect Heterogeneity	57
2.8.5 Example	59
2.9 Heckman's Econometric Model of Causality	62
2.10 Conclusion	65

3 Conventional Methods for Data Balancing 67

3.1 Why Is Data Balancing Necessary? A Heuristic Example	68
3.2 Three Methods for Data Balancing	71
3.2.1 The Ordinary Least Squares Regression	71
3.2.2 Matching	75
3.2.3 Stratification	76
3.3 Design of the Data Simulation	77
3.4 Results of the Data Simulation	79
3.5 Implications of the Data Simulation	86
3.6 Key Issues Regarding the Application of OLS Regression	93
3.7 Conclusion	94

4 Sample Selection and Related Models 95

4.1 The Sample Selection Model	95
4.1.1 Truncation, Censoring, and Incidental Truncation	96
4.1.2 Why Is It Important to Model Sample Selection?	99
4.1.3 Moments of an Incidentally Truncated Bivariate Normal Distribution	100
4.1.4 The Heckman Model and Its Two-Step Estimator	101
4.2 Treatment Effect Model	105
4.3 Overview of the Stata Programs and Main Features of <i>treatreg</i>	107
4.4 Examples	112
4.4.1 Application of the Treatment Effect Model to Analysis of Observational Data	112
4.4.2 Evaluation of Treatment Effects From a Program With a Group Randomization Design	118
4.4.3 Running the Treatment Effect Model After Multiple Imputations of Missing Data	125
4.5 Conclusion	127

5 Propensity Score Matching and Related Models 129

5.1 Overview	130
5.2 The Problem of Dimensionality and the Properties of Propensity Scores	134
5.3 Estimating Propensity Scores	137
5.3.1 Binary Logistic Regression	137
5.3.2 Strategies to Specify a Correct Model—Predicting Propensity Scores	140
5.3.3 Hirano and Imbens's Method for Specifying Predictors Relying on Predetermined Critical t Values	142
5.3.4 Generalized Boosted Modeling	144
5.4 Matching	145

5.4.1 Greedy Matching	145
5.4.2 Optimal Matching	148
5.4.3 Fine Balance	152
5.5 Postmatching Analysis	153
5.5.1 Multivariate Analysis After Greedy Matching	153
5.5.2 Computing Indices of Covariate Imbalance	154
5.5.3 Outcome Analysis Using the Hodges-Lehmann Aligned Rank Test After Optimal Matching	155
5.5.4 Regression Adjustment Based on Sample Created by Optimal Pair Matching	156
5.5.5 Regression Adjustment Using Hodges-Lehmann Aligned Rank Scores After Optimal Matching	157
5.6 Propensity Score Matching With Multilevel Data	157
5.6.1 Overview of Statistical Approaches to Multilevel Data	158
5.6.2 Perspectives Extending the Propensity Score Analysis to the Multilevel Modeling	160
5.6.3 Estimation of the Propensity Scores Under the Context of Multilevel Modeling	161
5.6.4 Multilevel Outcome Analysis	164
5.7 Overview of the Stata and R Programs	166
5.8 Examples	172
5.8.1 Greedy Matching and Subsequent Analysis of Hazard Rates	173
5.8.2 Optimal Matching	183
5.8.3 Post-Full Matching Analysis Using the Hodges-Lehmann Aligned Rank Test	190
5.8.4 Post-Pair Matching Analysis Using Regression of Difference Scores	191
5.8.5 Multilevel Propensity Score Analysis	191
5.8.6 Comparison of Rand- <i>gbm</i> and Stata's <i>boost</i> Algorithms	198
5.9 Conclusion	201

6 Propensity Score Subclassification **203**

6.1 Overview	204
6.2 The Overlap Assumption and Methods to Address Its Violation	209
6.3 Structural Equation Modeling With Propensity Score Subclassification	210
6.3.1 The Need for Integrating SEM and Propensity Score Modeling Into One Analysis	210
6.3.2 Kaplan's (1999) Work to Integrate Propensity Score Subclassification With SEM	216
6.3.3 Conduct SEM With Propensity Score Subclassification	216
6.4 The Stratification-Multilevel Method	217
6.5 Examples	219
6.5.1 Stratification After Greedy Matching	219
6.5.2 Subclassification Followed by a Cox Proportional Hazards Model	221
6.5.3 Propensity Score Subclassification in Conjunction With SEM	234
6.6 Conclusion	236

7 Propensity Score Weighting **239**

7.1 Overview	240
7.2 Weighting Estimators	244
7.2.1 Formulas for Creating Weights to Estimate ATE and ATT	244
7.2.2 A Corrected Version of Weights Estimating ATE	245
7.2.3 Steps in Propensity Score Weighting	246

- 7.3 Examples 246
 - 7.3.1 Propensity Score Weighting With a Multiple Regression Outcome Analysis 246
 - 7.3.2 Propensity Score Weighting With a Cox Proportional Hazards Model 247
 - 7.3.3 Propensity Score Weighting With an SEM 249
 - 7.3.4 Comparison of Models and Conclusions of the Study of
the Impact of Poverty on Child Academic Achievement 252
- 7.4 Conclusion 254

8 Matching Estimators 255

- 8.1 Overview 255
- 8.2 Methods of Matching Estimators 259
 - 8.2.1 Simple Matching Estimator 260
 - 8.2.2 Bias-Corrected Matching Estimator 266
 - 8.2.3 Variance Estimator Assuming Homoscedasticity 268
 - 8.2.4 Variance Estimator Allowing for Heteroscedasticity 269
 - 8.2.5 Large Sample Properties and Correction 269
- 8.3 Overview of the Stata Program *nnmatch* 270
- 8.4 Examples 271
 - 8.4.1 Matching With Bias-Corrected and Robust Variance Estimators 271
 - 8.4.2 Efficacy Subset Analysis With Matching Estimators 276
- 8.5 Conclusion 281

9 Propensity Score Analysis With Nonparametric Regression 283

- 9.1 Overview 284
- 9.2 Methods of Propensity Score Analysis With Nonparametric Regression 286
 - 9.2.1 The Kernel-Based Matching Estimators 287
 - 9.2.2 Review of the Basic Concepts of Local Linear Regression (*lowess*) 288
 - 9.2.3 Asymptotic and Finite-Sample Properties of Kernel and Local Linear Matching 297
- 9.3 Overview of the Stata Programs *psmatch2* and *bootstrap* 298
- 9.4 Examples 299
 - 9.4.1 Analysis of Difference-in-Differences 303
 - 9.4.2 Application of Kernel-Based Matching to One-Point Data 306
- 9.5 Conclusion 308

10 Propensity Score Analysis of Categorical or Continuous Treatments: Dosage Analyses 309

- 10.1 Overview 310
- 10.2 Modeling Doses With a Single Scalar Balancing Score Estimated by an
Ordered Logistic Regression 311
- 10.3 Modeling Doses With Multiple Balancing Scores Estimated by a
Multinomial Logit Model 313
- 10.4 The Generalized Propensity Score Estimator 314
- 10.5 Overview of the Stata *gpscore* Program 321
- 10.6 Examples 328
 - 10.6.1 Modeling Doses of Treatment With Multiple Balancing Scores
Estimated by a Multinomial Logit Model 328

10.6.2 Modeling Doses of Treatment With the Generalized Propensity Score Estimator	331
10.7 Conclusion	334
11 Selection Bias and Sensitivity Analysis	335
11.1 Selection Bias: An Overview	336
11.1.1 Sources of Selection Bias	336
11.1.2 Overt Bias Versus Hidden Bias	340
11.1.3 Consequences of Selection Bias	341
11.1.4 Strategies to Correct for Selection Bias	342
11.2 A Monte Carlo Study Comparing Corrective Models	345
11.2.1 Design of the Monte Carlo Study	347
11.2.2 Results of the Monte Carlo Study	353
11.2.3 Implications	356
11.3 Rosenbaum's Sensitivity Analysis	357
11.3.1 The Basic Idea	358
11.3.2 Illustration of Wilcoxon's Signed Rank Test for Sensitivity Analysis of a Matched Pair Study	359
11.4 Overview of the Stata Program <i>rbounds</i>	369
11.5 Examples	377
11.5.1 Sensitivity Analysis of the Effects of Lead Exposure	377
11.5.2 Sensitivity Analysis for the Study Using Pair Matching	378
11.6 Conclusion	379
12 Concluding Remarks	381
12.1 Common Pitfalls in Observational Studies: A Checklist for Critical Review	381
12.2 Approximating Experiments With Propensity Score Approaches	386
12.2.1 Criticism of Propensity Score Methods	386
12.2.2 Regression and Propensity Score Approaches: Do They Provide Similar Results?	387
12.2.3 Criticism of Sensitivity Analysis (I')	390
12.2.4 Group Randomized Trials	390
12.3 Other Advances in Modeling Causality	391
12.4 Directions for Future Development	392
References	395
Index	409