

## Contents

|                 |  |                |
|-----------------|--|----------------|
|                 | <i>Detailed table of contents</i>                          | <i>page</i> ix |
|                 | <i>List of figures</i>                                     | xiv            |
|                 | <i>List of tables</i>                                      | xv             |
|                 | <i>List of boxes</i>                                       | xvi            |
|                 | <i>Preface and acknowledgements</i>                        | xvii           |
| <b>1</b>        | Introduction: why natural experiments?                     | 1              |
| <hr/>           |  |                |
| <b>Part I</b>   | <b>Discovering natural experiments</b>                     | 39             |
| <hr/>           |  |                |
| <b>2</b>        | Standard natural experiments                               | 41             |
| <b>3</b>        | Regression-discontinuity designs                           | 63             |
| <b>4</b>        | Instrumental-variables designs                             | 87             |
| <hr/>           |  |                |
| <b>Part II</b>  | <b>Analyzing natural experiments</b>                       | 103            |
| <hr/>           |  |                |
| <b>5</b>        | Simplicity and transparency: keys to quantitative analysis | 105            |
| <b>6</b>        | Sampling processes and standard errors                     | 165            |
| <b>7</b>        | The central role of qualitative evidence                   | 208            |
| <hr/>           |  |                |
| <b>Part III</b> | <b>Evaluating natural experiments</b>                      | 233            |
| <hr/>           |  |                |
| <b>8</b>        | How plausible is as-if random?                             | 235            |
| <b>9</b>        | How credible is the model?                                 | 256            |
| <b>10</b>       | How relevant is the intervention?                          | 289            |

**viii**      **Contents**

---

|                |   |     |
|----------------|---|-----|
| <b>Part IV</b> | <b>Conclusion</b>                                     | 311 |
| <b>11</b>      | Building strong designs through multi-method research | 313 |
|                | <i>References</i>                                     | 338 |
|                | <i>Index</i>  | 353 |

## Detailed table of contents

|               |  |                  |
|---------------|--|------------------|
|               | <i>Preface and acknowledgements</i>                          | <i>page xvii</i> |
| 1             | <b>Introduction: why natural experiments?</b>                | 1                |
| 1.1           | The problem of confounders                                   | 5                |
| 1.1.1         | The role of randomization                                    | 6                |
| 1.2           | Natural experiments on military conscription and land titles | 8                |
| 1.3           | Varieties of natural experiments                             | 15               |
| 1.3.1         | Contrast with quasi-experiments and matching                 | 18               |
| 1.4           | Natural experiments as design-based research                 | 21               |
| 1.5           | An evaluative framework for natural experiments              | 27               |
| 1.5.1         | The plausibility of as-if random                             | 27               |
| 1.5.2         | The credibility of models                                    | 28               |
| 1.5.3         | The relevance of the intervention                            | 29               |
| 1.6           | Critiques and limitations of natural experiments             | 32               |
| 1.7           | Avoiding conceptual stretching                               | 34               |
| 1.8           | Plan for the book, and how to use it                         | 35               |
| 1.8.1         | Some notes on coverage                                       | 37               |
| <hr/>         |  |                  |
| <b>Part I</b> | <b>Discovering natural experiments</b>                       | 39               |
| 2             | <b>Standard natural experiments</b>                          | 41               |
| 2.1           | Standard natural experiments in the social sciences          | 43               |
| 2.2           | Standard natural experiments with true randomization         | 48               |
| 2.2.1         | Lottery studies  | 49               |
| 2.3           | Standard natural experiments with as-if randomization        | 53               |
| 2.3.1         | Jurisdictional borders                                       | 57               |
| 2.3.2         | Redistricting and jurisdiction shopping                      | 59               |
| 2.4           | Conclusion   | 60               |
|               | Exercises  | 61               |

|          |  |            |
|----------|--|------------|
| <b>x</b> | <b>Detailed table of contents</b>  |            |
| <b>3</b> | <b>Regression-discontinuity designs</b>  | <b>63</b>  |
| 3.1      | The basis of regression-discontinuity analysis                                   | 63         |
| 3.2      | Regression-discontinuity designs in the social sciences                          | 68         |
| 3.2.1    | Population- and size-based thresholds  | 72         |
| 3.2.2    | Near-winners and near-losers of close elections                                  | 77         |
| 3.2.3    | Age as a regression discontinuity  | 79         |
| 3.2.4    | Indices  | 80         |
| 3.3      | Variations on regression-discontinuity designs                                   | 81         |
| 3.3.1    | Sharp versus fuzzy regression discontinuities                                    | 81         |
| 3.3.2    | Randomized regression-discontinuity designs                                      | 82         |
| 3.3.3    | Multiple thresholds  | 83         |
| 3.4      | Conclusion   | 84         |
|          | Exercises  | 85         |
| <b>4</b> | <b>Instrumental-variables designs</b>  | <b>87</b>  |
| 4.1      | Instrumental-variables designs: true experiments                                 | 91         |
| 4.2      | Instrumental-variables designs: natural experiments                              | 92         |
| 4.2.1    | Lotteries  | 94         |
| 4.2.2    | Weather shocks   | 95         |
| 4.2.3    | Historical or institutional variation induced by deaths                          | 97         |
| 4.3      | Conclusion   | 101        |
|          | Exercises  | 102        |
|          | <b>Part II Analyzing natural experiments</b>                                     | <b>103</b> |
| <b>5</b> | <b>Simplicity and transparency: keys to quantitative analysis</b>                | <b>105</b> |
| 5.1      | The Neyman model   | 107        |
| 5.1.1    | The average causal effect  | 109        |
| 5.1.2    | Estimating the average causal effect   | 112        |
| 5.1.3    | An example: land titling in Argentina  | 115        |
| 5.1.4    | Key assumptions of the Neyman model  | 118        |
| 5.1.5    | Analyzing standard natural experiments   | 121        |
| 5.2      | Analyzing regression-discontinuity designs                                       | 121        |
| 5.2.1    | Two examples: Certificates of Merit and digital democratization                  | 123        |
| 5.2.2    | Defining the study group: the question of bandwidth                              | 127        |
| 5.2.3    | Is the difference-of-means estimator biased in regression-discontinuity designs? | 128        |

|              |  |     |
|--------------|--|-----|
| 5.2.4        | Modeling functional form   | 133 |
| 5.2.5        | Fuzzy regression discontinuities   | 134 |
| 5.3          | Analyzing instrumental-variables designs   | 135 |
| 5.3.1        | Natural experiments with noncompliance   | 136 |
| 5.3.2        | An example: the effect of military service   | 143 |
| 5.3.3        | The no-Defiers assumption  | 148 |
| 5.3.4        | Fuzzy regression-discontinuities as instrumental-variables designs                                 | 149 |
| 5.3.5        | From the Complier average effect to linear regression  | 150 |
| 5.4          | Conclusion   | 153 |
| Appendix 5.1 | Instrumental-variables estimation of the Complier average causal effect                            | 154 |
| Appendix 5.2 | Is the difference-of-means estimator biased in regression-discontinuity designs (further details)? | 158 |
|              | Exercises  | 160 |
| 6            | Sampling processes and standard errors   | 165 |
| 6.1          | Standard errors under the Neyman urn model   | 166 |
| 6.1.1        | Standard errors in regression-discontinuity and instrumental-variables designs                     | 173 |
| 6.2          | Handling clustered randomization   | 175 |
| 6.2.1        | Analysis by cluster mean: a design-based approach  | 179 |
| 6.3          | Randomization inference: Fisher's exact test   | 186 |
| 6.4          | Conclusion   | 191 |
| Appendix 6.1 | Conservative standard errors under the Neyman model  | 192 |
| Appendix 6.2 | Analysis by cluster mean   | 195 |
|              | Exercises  | 201 |
| 7            | The central role of qualitative evidence   | 208 |
| 7.1          | Causal-process observations in natural experiments   | 210 |
| 7.1.1        | Validating as-if random: treatment-assignment CPOs   | 212 |
| 7.1.2        | Verifying treatments: independent-variable CPOs  | 219 |
| 7.1.3        | Explaining effects: mechanism CPOs   | 222 |
| 7.1.4        | Interpreting effects: auxiliary-outcome CPOs   | 224 |
| 7.1.5        | Bolstering credibility: model-validation CPOs  | 225 |
| 7.2          | Conclusion   | 228 |
|              | Exercises  | 230 |

**xii**      **Detailed table of contents**

---

|                 |  |            |
|-----------------|--|------------|
| <b>Part III</b> | <b>Evaluating natural experiments</b>  | <b>233</b> |
| <b>8</b>        | <b>How plausible is as-if random?</b>  | <b>235</b> |
| 8.1             | Assessing as-if random   | 236        |
| 8.1.1           | The role of balance tests  | 239        |
| 8.1.2           | Qualitative diagnostics  | 243        |
| 8.2             | Evaluating as-if random in regression-discontinuity and instrumental-variables designs | 244        |
| 8.2.1           | Sorting at the regression-discontinuity threshold: conditional density tests           | 245        |
| 8.2.2           | Placebo tests in regression-discontinuity designs                                      | 246        |
| 8.2.3           | Treatment-assignment CPOs in regression-discontinuity designs                          | 248        |
| 8.2.4           | Diagnostics in instrumental-variables designs  | 248        |
| 8.3             | A continuum of plausibility  | 249        |
| 8.4             | Conclusion   | 252        |
|                 | Exercises  | 254        |
| <b>9</b>        | <b>How credible is the model?</b>  | <b>256</b> |
| 9.1             | The credibility of causal and statistical models                                       | 258        |
| 9.1.1           | Strengths and limitations of the Neyman model  | 259        |
| 9.1.2           | Linear regression models   | 263        |
| 9.2             | Model specification in instrumental-variables regression                               | 269        |
| 9.2.1           | Control variables in instrumental-variables regression                                 | 277        |
| 9.3             | A continuum of credibility   | 278        |
| 9.4             | Conclusion: how important is the model?  | 283        |
| Appendix 9.1    | Homogeneous partial effects with multiple treatments and instruments                   | 285        |
|                 | Exercises  | 287        |
| <b>10</b>       | <b>How relevant is the intervention?</b>   | <b>289</b> |
| 10.1            | Threats to substantive relevance   | 293        |
| 10.1.1          | Lack of external validity  | 293        |
| 10.1.2          | Idiosyncrasy of interventions  | 297        |
| 10.1.3          | Bundling of treatments   | 300        |
| 10.2            | A continuum of relevance   | 303        |
| 10.3            | Conclusion   | 306        |
|                 | Exercises  | 309        |

**xiii** **Detailed table of contents**

---

|                |  |     |
|----------------|--|-----|
| <b>Part IV</b> | <b>Conclusion</b>  | 311 |
| 11             | Building strong designs through multi-method research    | 313 |
| 11.1           | The virtues and limitations of natural experiments       | 315 |
| 11.2           | A framework for strong research designs                  | 318 |
| 11.2.1         | Conventional observational studies and true experiments  | 319 |
| 11.2.2         | Locating natural experiments                             | 321 |
| 11.2.3         | Relationship between dimensions                          | 324 |
| 11.3           | Achieving strong design: the importance of mixed methods | 326 |
| 11.4           | A checklist for natural-experimental research            | 328 |
|                | <i>References</i>  | 338 |
|                | <i>Index</i>   | 353 |