This unique book is the first comprehensive guide to the discovery, analysis, and evaluation of natural experiments—an increasingly popular methodology in the social sciences. Thad Dunning provides an introduction to key issues in causal inference, including model specification, and emphasizes the importance of strong research design over complex statistical analysis. Surveying many examples of standard natural experiments, regression-discontinuity designs, and instrumental-variables designs, Dunning highlights both the strengths and potential weaknesses of these methods, aiding researchers in better harnessing the promise of natural experiments while avoiding the pitfalls. Dunning also demonstrates the contribution of qualitative methods to natural experiments and proposes new ways to integrate qualitative and quantitative techniques. Chapters complete with exercises, and appendices covering specialized topics such as cluster-randomized natural experiments, make this an ideal teaching tool as well as a valuable book for professional researchers.

Thad Dunning is Associate Professor of Political Science at Yale University and a research fellow at Yale’s Institution for Social and Policy Studies and the Whitney and Betty MacMillan Center for International and Area Studies. He has written on a range of methodological topics, including impact evaluation, econometric corrections for selection effects, and multi-method research in the social sciences, and his first book, *Crude Democracy: Natural Resource Wealth and Political Regimes* (Cambridge University Press, 2008), won the Best Book Award from the Comparative Democratization Section of the American Political Science Association.
Strategies for Social Inquiry

Natural Experiments in the Social Sciences: A Design-Based Approach

Editors
Colin Elman, Maxwell School of Syracuse University
John Gerring, Boston University
James Mahoney, Northwestern University

Editorial Board
Bear Braumoeller, David Collier, Francesco Guala, Peter Hedström, Theodore Hopf, Uskali Maki, Rose McDermott, Charles Ragin, Theda Skocpol, Peter Spiegler, David Waldner, Lisa Wedeen, Christopher Winship

This new book series presents texts on a wide range of issues bearing upon the practice of social inquiry. Strategies are construed broadly to embrace the full spectrum of approaches to analysis, as well as relevant issues in philosophy of social science.

Published Titles
John Gerring, Social Science Methodology: A Unified Framework, 2nd edition
Michael Coppedge, Democratization and Research Methods
Carsten Q. Schneider and Claudius Wagemann, Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis

Forthcoming Titles
Diana Kapiszewski, Lauren M. MacLean and Benjamin L. Read, Field Research in Political Science
Jason Seawright, Multi-Method Social Science: Combining Qualitative and Quantitative Tools
Natural Experiments in the Social Sciences

A Design-Based Approach

Thad Dunning
Dedicated to the memory of David A. Freedman
## Contents

- **Detailed table of contents**
- **List of figures**
- **List of tables**
- **List of boxes**
- **Preface and acknowledgements**

### 1 Introduction: why natural experiments?

**Part I Discovering natural experiments**

- 2 Standard natural experiments
- 3 Regression-discontinuity designs
- 4 Instrumental-variables designs

**Part II Analyzing natural experiments**

- 5 Simplicity and transparency: keys to quantitative analysis
- 6 Sampling processes and standard errors
- 7 The central role of qualitative evidence

**Part III Evaluating natural experiments**

- 8 How plausible is as-if random?
- 9 How credible is the model?
- 10 How relevant is the intervention?
<table>
<thead>
<tr>
<th>Part IV</th>
<th>Conclusion</th>
<th>311</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Building strong designs through multi-method research</td>
<td>313</td>
</tr>
</tbody>
</table>

References 338
Index 353
# Detailed table of contents

<table>
<thead>
<tr>
<th>Preface and acknowledgements</th>
<th>page xvii</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Introduction: why natural experiments?</td>
<td>1</td>
</tr>
<tr>
<td>1.1 The problem of confounders</td>
<td>5</td>
</tr>
<tr>
<td>1.1.1 The role of randomization</td>
<td>6</td>
</tr>
<tr>
<td>1.2 Natural experiments on military conscription and land titles</td>
<td>8</td>
</tr>
<tr>
<td>1.3 Varieties of natural experiments</td>
<td>15</td>
</tr>
<tr>
<td>1.3.1 Contrast with quasi-experiments and matching</td>
<td>18</td>
</tr>
<tr>
<td>1.4 Natural experiments as design-based research</td>
<td>21</td>
</tr>
<tr>
<td>1.5 An evaluative framework for natural experiments</td>
<td>27</td>
</tr>
<tr>
<td>1.5.1 The plausibility of as-if random</td>
<td>27</td>
</tr>
<tr>
<td>1.5.2 The credibility of models</td>
<td>28</td>
</tr>
<tr>
<td>1.5.3 The relevance of the intervention</td>
<td>29</td>
</tr>
<tr>
<td>1.6 Critiques and limitations of natural experiments</td>
<td>32</td>
</tr>
<tr>
<td>1.7 Avoiding conceptual stretching</td>
<td>34</td>
</tr>
<tr>
<td>1.8 Plan for the book, and how to use it</td>
<td>35</td>
</tr>
<tr>
<td>1.8.1 Some notes on coverage</td>
<td>37</td>
</tr>
</tbody>
</table>

## Part I Discovering natural experiments

| 2 Standard natural experiments | 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 |
3 Regression-discontinuity designs
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

4 Instrumental-variables designs
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

Part II Analyzing natural experiments

5 Simplicity and transparency: keys to quantitative analysis 105
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
# Part III Evaluating natural experiments

## 8 How plausible is as-if random?

8.1 Assessing as-if random

8.1.1 The role of balance tests

8.1.2 Qualitative diagnostics

8.2 Evaluating as-if random in regression-discontinuity and instrumental-variables designs

8.2.1 Sorting at the regression-discontinuity threshold: conditional density tests

8.2.2 Placebo tests in regression-discontinuity designs

8.2.3 Treatment-assignment CPOs in regression-discontinuity designs

8.2.4 Diagnostics in instrumental-variables designs

8.3 A continuum of plausibility

8.4 Conclusion

Exercises

## 9 How credible is the model?

9.1 The credibility of causal and statistical models

9.1.1 Strengths and limitations of the Neyman model

9.1.2 Linear regression models

9.2 Model specification in instrumental-variables regression

9.2.1 Control variables in instrumental-variables regression

9.3 A continuum of credibility

9.4 Conclusion: how important is the model?

Appendix 9.1 Homogeneous partial effects with multiple treatments and instruments

Exercises

## 10 How relevant is the intervention?

10.1 Threats to substantive relevance

10.1.1 Lack of external validity

10.1.2 Idiosyncrasy of interventions

10.1.3 Bundling of treatments

10.2 A continuum of relevance

10.3 Conclusion

Exercises
<table>
<thead>
<tr>
<th>Part IV</th>
<th>Conclusion</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Building strong designs through multi-method research</td>
<td>313</td>
</tr>
<tr>
<td>11.1</td>
<td>The virtues and limitations of natural experiments</td>
<td>315</td>
</tr>
<tr>
<td>11.2</td>
<td>A framework for strong research designs</td>
<td>318</td>
</tr>
<tr>
<td>11.2.1</td>
<td>Conventional observational studies and true experiments</td>
<td>319</td>
</tr>
<tr>
<td>11.2.2</td>
<td>Locating natural experiments</td>
<td>321</td>
</tr>
<tr>
<td>11.2.3</td>
<td>Relationship between dimensions</td>
<td>324</td>
</tr>
<tr>
<td>11.3</td>
<td>Achieving strong design: the importance of mixed methods</td>
<td>326</td>
</tr>
<tr>
<td>11.4</td>
<td>A checklist for natural-experimental research</td>
<td>328</td>
</tr>
</tbody>
</table>

References 338
Index 353
Figures

1.1 Natural experiments in political science and economics  page 2
1.2 Typology of natural experiments  31
3.1 Examples of regression discontinuities  66
5.1 The Neyman model  113
5.2 A regression-discontinuity design  124
5.3 Potential and observed outcomes in a regression-discontinuity design  129
5.4 Noncompliance under the Neyman model  140
6.1 Clustered randomization under the Neyman model  177
8.1 Plausibility of as-if random assignment  250
9.1 Credibility of statistical models  280
10.1 Substantive relevance of intervention  303
11.1 Strong research designs  318
11.2 A decision flowchart for natural experiments  329
### Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Death rates from cholera by water-supply source</td>
<td>13</td>
</tr>
<tr>
<td>2.1</td>
<td>Typical “standard” natural experiments</td>
<td>44</td>
</tr>
<tr>
<td>2.2</td>
<td>Standard natural experiments with true randomization</td>
<td>45</td>
</tr>
<tr>
<td>2.3</td>
<td>Standard natural experiments with as-if randomization</td>
<td>46</td>
</tr>
<tr>
<td>3.1</td>
<td>Selected sources of regression-discontinuity designs</td>
<td>69</td>
</tr>
<tr>
<td>3.2</td>
<td>Examples of regression-discontinuity designs</td>
<td>70</td>
</tr>
<tr>
<td>4.1</td>
<td>Selected sources of instrumental-variables designs</td>
<td>90</td>
</tr>
<tr>
<td>4.2</td>
<td>Selected instrumental-variables designs (true experiments)</td>
<td>92</td>
</tr>
<tr>
<td>4.3</td>
<td>Selected instrumental-variables designs (natural experiments)</td>
<td>93</td>
</tr>
<tr>
<td>4.4</td>
<td>Direct and indirect colonial rule in India</td>
<td>98</td>
</tr>
<tr>
<td>5.1</td>
<td>The effects of land titles on children’s health</td>
<td>117</td>
</tr>
<tr>
<td>5.2</td>
<td>Social Security earnings in 1981</td>
<td>145</td>
</tr>
<tr>
<td>6.1</td>
<td>Potential outcomes under the strict null hypothesis</td>
<td>187</td>
</tr>
<tr>
<td>6.2</td>
<td>Outcomes under all randomizations, under the strict null hypothesis</td>
<td>188</td>
</tr>
</tbody>
</table>
Boxes

4.1 The intention-to-treat principle \hspace{1cm} page 88
5.1 The Neyman model and the average causal effect \hspace{1cm} 111
5.2 Estimating the average causal effect \hspace{1cm} 115
6.1 Standard errors under the Neyman model \hspace{1cm} 170
6.2 Code for analysis by cluster means \hspace{1cm} 181
Preface and acknowledgements

Natural experiments have become ubiquitous in the social sciences. From standard natural experiments to regression-discontinuity and instrumental-variables designs, our leading research articles and books more and more frequently reference this label. For professional researchers and students alike, natural experiments are often recommended as a tool for strengthening causal claims.

Surprisingly, we lack a comprehensive guide to this type of research design. Finding a useful and viable natural experiment is as much an art as a science. Thus, an extensive survey of examples—grouped and discussed to highlight how and why they provide the leverage they do—may help scholars to use natural experiments effectively in their substantive research. Just as importantly, awareness of the obstacles to successful natural experiments may help scholars maximize their promise while avoiding their pitfalls. There are significant challenges involved in the analysis and interpretation of natural-experimental data. Moreover, the growing popularity of natural experiments can lead to conceptual stretching, as the label is applied to studies that do not very credibly bear the hallmarks of this research design. Discussion of both the strengths and limitations of natural experiments may help readers to evaluate and bolster the success of specific applications. I therefore hope that this book will provide a resource for scholars and students who want to conduct or critically consume work of this type.

While the book is focused on natural experiments, it is also a primer on design-based research in the social sciences more generally. Research that depends on ex post statistical adjustment (such as cross-country regressions) has recently come under fire; there has been a commensurate shift of focus toward design-based research—in which control over confounding variables comes primarily from research design, rather than model-based statistical adjustment. The current enthusiasm for natural experiments reflects this renewed emphasis on design-based research. Yet, how should such research be conducted and evaluated? What are the key assumptions
behind design-based inference, and what causal and statistical models are appropriate for this style of research? And can such design-based approaches help us make progress on big, important substantive topics, such as the causes and consequences of democracy or socioeconomic development? Answering such questions is critical for sustaining the credibility and relevance of design-based research.

Finally, this book also highlights the potential payoffs from integrating qualitative and quantitative methods. “Bridging the divide” between approaches is a recurring theme in many social sciences. Yet, strategies for combining multiple methods are not always carefully explicated; and the value of such combinations is sometimes presumed rather than demonstrated. This is unfortunate: at least with natural experiments, different methods do not just supplement but often require one another. I hope that this book can clarify the payoffs of mixing methods and especially of the “shoe-leather” research that, together with strong designs, makes compelling causal inference possible.

This book grows out of discussions with many colleagues, students, and especially mentors. I am deeply fortunate to have met David Freedman, to whom the book is dedicated, while finishing my Ph.D. studies at the University of California at Berkeley. His impact on this book will be obvious to readers who know his work; I only wish that he were alive to read it. While he is greatly missed, he left behind an important body of research, with which every social scientist who seeks to make causal inferences should grapple.

I would also like to thank several other mentors, colleagues, and friends. David Collier’s exemplary commitment to the merger of qualitative and quantitative work has helped me greatly along the way; this book grew out of a chapter I wrote for the second edition of his book, *Rethinking Social Inquiry*, co-edited with Henry Brady. Jim Robinson, himself a prominent advocate of natural-experimental research designs, continues to influence my own substantive and methodological research. I would especially like to thank Don Green and Dan Posner, both great friends and colleagues, who read and offered detailed and incisive comments on large portions of the manuscript. Colin Elman organized a research workshop at the Institute for Qualitative and Multi-Method Research at Syracuse University, where John Gerring and David Waldner served as very discerning discussants, while Macartan Humphreys and Alan Jacobs convoked a book event at the University of British Columbia, at which Anjali Bohlken, Chris Kam, and Ben Nyblade each perceptively dissected individual chapters. I am grateful to all of the participants in these two events. For helpful conversations and suggestions, I also thank Jennifer Bussell, Colin Elman, Danny Hidalgo,
Macartan Humphreys, Jim Mahoney, Ken Scheve, Jay Seawright, Jas Sekhon, Rocío Titiunik, and David Waldner. I have been privileged to teach courses and workshops on related material to graduate students at the Institute for Qualitative and Multi-Method Research and at Yale, where Natalia Bueno, Germán Feierherd, Nikhar Gaikwad, Malte Lierl, Pia Raffler, Steve Rosenzweig, Luis Schiumerini, Dawn Teele, and Guadalupe Tuñón, among others, offered insightful reactions. I have also enjoyed leading an annual short course on multi-method research at the American Political Science Association with David Collier and Jay Seawright. These venues have provided a valuable opportunity to improve my thinking on the topics discussed in this book, and I thank all the participants in those workshops and courses for their feedback.

I would also like to thank my editor on this book, John Haslam of Cambridge University Press, as well as Carrie Parkinson, Ed Robinson, and Jim Thomas, for their gracious shepherding of this book to completion. I am particularly grateful to Colin Elman, John Gerring, and Jim Mahoney, who approached me about writing this book for their Strategies of Social Inquiry series. For their steadfast love and support, my deepest gratitude goes to my family.