

# CONTENTS

	<b>PREFACE</b>	xv
CHAPTER 1	<b>Introduction</b>	1
	1.1 Drawing Inferences from Intuitions, Anecdotes, and Correlations	2
	1.2 Experiments as a Solution to the Problem of Unobserved Confounders	5
	1.3 Experiments as Fair Tests	7
	1.4 Field Experiments	8
	1.5 Advantages and Disadvantages of Experimenting in Real-World Settings	13
	1.6 Naturally Occurring Experiments and Quasi-Experiments	15
	1.7 Plan of the Book	17
	Suggested Readings	18
	Exercises	18
CHAPTER 2	<b>Causal Inference and Experimentation</b>	21
	2.1 Potential Outcomes	21
	2.2 Average Treatment Effects	23
	2.3 Random Sampling and Expectations	26
	2.4 Random Assignment and Unbiased Inference	30
	2.5 The Mechanics of Random Assignment	36
	2.6 The Threat of Selection Bias When Random Assignment Is Not Used	37
	2.7 Two Core Assumptions about Potential Outcomes	39
	2.7.1 Excludability	39
	2.7.2 Non-Interference	43



	Summary	44
	Suggested Readings	46
	Exercises	46
CHAPTER 3	<b>Sampling Distributions, Statistical Inference, and Hypothesis Testing</b>	51
	3.1 Sampling Distributions	52
	3.2 The Standard Error as a Measure of Uncertainty	54
	3.3 Estimating Sampling Variability	59
	3.4 Hypothesis Testing	61
	3.5 Confidence Intervals	66
	3.6 Sampling Distributions for Experiments That Use Block or Cluster Random Assignment	71
	3.6.1 Block Random Assignment	71
	3.6.1.1 Matched Pair Design	77
	3.6.1.2 Summary of the Advantages and Disadvantages of Blocking	79
	3.6.2 Cluster Random Assignment	80
	Summary	85
	Suggested Readings	86
	Exercises	86
	Appendix 3.1: Power	93
CHAPTER 4	<b>Using Covariates in Experimental Design and Analysis</b>	95
	4.1 Using Covariates to Rescale Outcomes	96
	4.2 Adjusting for Covariates Using Regression	102
	4.3 Covariate Imbalance and the Detection of Administrative Errors	105
	4.4 Blocked Randomization and Covariate Adjustment	109
	4.5 Analysis of Block Randomized Experiments with Treatment Probabilities That Vary by Block	116
	Summary	121
	Suggested Readings	123
	Exercises	123



CHAPTER 5	<b>One-Sided Noncompliance</b>	131
	5.1 New Definitions and Assumptions	134
	5.2 Defining Causal Effects for the Case of One-Sided Noncompliance	137
	5.2.1 The Non-Interference Assumption for Experiments That Encounter Noncompliance	138
	5.2.2 The Excludability Assumption for One-Sided Noncompliance	140
	5.3 Average Treatment Effects, Intent-to-Treat Effects, and Complier Average Causal Effects	141
	5.4 Identification of the CACE	143
	5.5 Estimation	149
	5.6 Avoiding Common Mistakes	152
	5.7 Evaluating the Assumptions Required to Identify the CACE	155
	5.7.1 Non-Interference Assumption	155
	5.7.2 Exclusion Restriction	156
	5.8 Statistical Inference	157
	5.9 Designing Experiments in Anticipation of Noncompliance	161
	5.10 Estimating Treatment Effects When Some Subjects Receive “Partial Treatment”	164
	Summary	165
	Suggested Readings	167
	Exercises	168
CHAPTER 6	<b>Two-Sided Noncompliance</b>	173
	6.1 Two-Sided Noncompliance: New Definitions and Assumptions	175
	6.2 ITT, $ITT_D$ , and CACE under Two-Sided Noncompliance	179
	6.3 A Numerical Illustration of the Role of Monotonicity	181
	6.4 Estimation of the CACE: An Example	185
	6.5 Discussion of Assumptions	189
	6.5.1 Monotonicity	190
	6.5.2 Exclusion Restriction	191
	6.5.3 Random Assignment	192
	6.5.4 Design Suggestions	192



---

	6.6 Downstream Experimentation	193
	Summary	204
	Suggested Readings	206
	Exercises	206
CHAPTER 7	<b>Attrition</b>	211
	7.1 Conditions Under Which Attrition Leads to Bias	215
	7.2 Special Forms of Attrition	219
	7.3 Redefining the Estimand When Attrition Is Not a Function of Treatment Assignment	224
	7.4 Placing Bounds on the Average Treatment Effect	226
	7.5 Addressing Attrition: An Empirical Example	230
	7.6 Addressing Attrition with Additional Data Collection	236
	7.7 Two Frequently Asked Questions	241
	Summary	243
	Suggested Readings	244
	Exercises	244
	Appendix 7.1: Optimal Sample Allocation for Second-Round Sampling	248
CHAPTER 8	<b>Interference between Experimental Units</b>	253
	8.1 Identifying Causal Effects in the Presence of Localized Spillover	256
	8.2 Spatial Spillover	260
	8.2.1 Using Nonexperimental Units to Investigate Spillovers	264
	8.3 An Example of Spatial Spillovers in Two Dimensions	264
	8.4 Within-Subjects Design and Time-Series Experiments	273
	8.5 Waitlist Designs (Also Known as Stepped-Wedge Designs)	276
	Summary	281
	Suggested Readings	283
	Exercises	283
CHAPTER 9	<b>Heterogeneous Treatment Effects</b>	289
	9.1 Limits to What Experimental Data Tell Us about Treatment Effect Heterogeneity	291



9.2	Bounding Var ( $\tau_i$ ) and Testing for Heterogeneity	292
9.3	Two Approaches to the Exploration of Heterogeneity: Covariates and Design	296
9.3.1	Assessing Treatment-by-Covariate Interactions	296
9.3.2	Caution Is Required When Interpreting Treatment-by-Covariate Interactions	299
9.3.3	Assessing Treatment-by-Treatment Interactions	303
9.4	Using Regression to Model Treatment Effect Heterogeneity	305
9.5	Automating the Search for Interactions	310
	Summary	310
	Suggested Readings	312
	Exercises	313
CHAPTER 10	<b>Mediation</b>	319
10.1	Regression-Based Approaches to Mediation	322
10.2	Mediation Analysis from a Potential Outcomes Perspective	325
10.3	Why Experimental Analysis of Mediators Is Challenging	328
10.4	Ruling Out Mediators?	330
10.5	What about Experiments That Manipulate the Mediator?	331
10.6	Implicit Mediation Analysis	333
	Summary	336
	Suggested Readings	338
	Exercises	338
	Appendix 10.1: Treatment Postcards Mailed to Michigan Households	343
CHAPTER 11	<b>Integration of Research Findings</b>	347
11.1	Estimation of Population Average Treatment Effects	350
11.2	A Bayesian Framework for Interpreting Research Findings	353
11.3	Replication and Integration of Experimental Findings: An Example	358
11.4	Treatments That Vary in Intensity: Extrapolation and Statistical Modeling	366



---

	Summary	377
	Suggested Readings	378
	Exercises	379
CHAPTER 12	<b>Instructive Examples of Experimental Design</b>	383
	12.1 Using Experimental Design to Distinguish between Competing Theories	384
	12.2 Oversampling Subjects Based on Their Anticipated Response to Treatment	387
	12.3 Comprehensive Measurement of Outcomes	393
	12.4 Factorial Design and Special Cases of Non-Interference	395
	12.5 Design and Analysis of Experiments In Which Treatments Vary with Subjects' Characteristics	400
	12.6 Design and Analysis of Experiments In Which Failure to Receive Treatment Has a Causal Effect	406
	12.7 Addressing Complications Posed by Missing Data	410
	Summary	414
	Suggested Readings	415
	Exercises	416
CHAPTER 13	<b>Writing a Proposal, Research Report, and Journal Article</b>	425
	13.1 Writing the Proposal	426
	13.2 Writing the Research Report	435
	13.3 Writing the Journal Article	440
	13.4 Archiving Data	442
	Summary	444
	Suggested Readings	445
	Exercises	445
APPENDIX A	<b>Protection of Human Subjects</b>	447
	A.1 Regulatory Guidelines	447
	A.2 Guidelines for Keeping Field Experiments within Regulatory Boundaries	449



APPENDIX B <b>Suggested Field Experiments for Class Projects</b>	453
B.1 Crafting Your Own Experiment	453
B.2 Suggested Experimental Topics for Practicum Exercises	455
<b>REFERENCES</b>	461
<b>INDEX</b>	479

For more than a decade, we have taught a one-semester course on experimental research methods to undergraduate and graduate students in the social sciences.

Although readings and discussion sometimes address experiments conducted in the lab, the course focuses primarily on “field” experiments: studies conducted in natural settings in which subjects are allocated randomly to treatment and control groups. Students read research articles that illustrate key principles of experimental design or analysis, and class time is devoted to explaining these principles. Students often find the material engaging and even inspiring, but the fact that they read selections from a broad research literature rather than a textbook means that even very talented students frequently fail to assimilate important terms, concepts, and techniques.

Our aim in writing this book is to provide a systematic introduction to experimentation that also conveys the excitement of encountering and conducting primary research. Each chapter weaves abstract principles together with examples drawn from a wide range of social science disciplines: criminology, economics, education, political science, social psychology, and sociology. The exercises at the end of each chapter invite students to reflect on abstract problems of research design and to analyze data from (or inspired by) important experiments. Our aim is to alert readers to the vast range of experimental applications and opportunities for future investigation.

Developing expertise as an experimental researcher is part technical training and part apprenticeship. The former requires the reader to think about experimentation in abstract terms. What inferences can be drawn from an experiment, and under what conditions might these inferences be jeopardized? Any explanation of abstract principles must inevitably invoke statistical terminology, because the language of statistics brings precision and generality. The presentation in this book presupposes that the reader has at some point taken a one- or two-semester introduction to statistical inference and regression. Recognizing that the reader’s memory of statistical principles may be hazy, the book continually defines, explains, and illustrates. In an effort to make the presentation accessible, we have freely renamed arcane terms of art.

Our aim throughout is to use naming conventions that convey the intuition behind the idea or procedure. The term “external validity,” for example, is replaced by “generalizability.” We also depart from the academic convention of using scholars’ names to refer to ideas or procedures. The term “external value bounds,” for example, replaces the term “interval bounds.” References are provided so that recognition of key ideas receive appropriate credit.