

Interaction Models

The radical interdependence between humans who live together makes virtually all human behavior conditional. The behavior of individuals is conditional upon the expectations of those around them, and those expectations are conditional upon the rules (institutions) and norms (culture) constructed to monitor, reward, and punish different behaviors. As a result, nearly all hypotheses about humans are conditional – conditional upon the resources they possess, the institutions they inhabit, or the cultural practices that tell them how to behave. *Interaction Models* provides a standalone, accessible overview of how interaction models, which are frequently used across the social and natural sciences, capture the intuition behind conditional claims and context dependence. It also addresses the simple specification and interpretation errors that are, unfortunately, commonplace. By providing a comprehensive and unified introduction to the use and critical evaluation of interaction models, this book shows how they can be used to test theoretically derived claims of conditionality.

WILLIAM ROBERTS CLARK is the author of *Capitalism, Not Globalism, Principles of Comparative Politics*, and numerous journal articles. With Sona and Matt Golder he was awarded the Brian Barry Prize by the British Academy. He has taught at six leading research universities and is currently President of the European Political Science Association.

MATT GOLDER is in the top 2 per cent of the most cited scientists worldwide, and his article with William Clark, “Understanding Interaction Models: Improving Empirical Analyses,” is in the top 10 most cited articles in political science (http://charlesbreton.ca/assets/PS_Top10_2020.pdf). He is the winner of the GESIS Klingemann Prize and the Brian Barry Prize from the British Academy.

Methodological Tools in the Social Sciences

SERIES EDITORS

Paul M. Kellstedt, *Professor of Political Science, Texas A&M University*
Guy D. Whitten, *Professor of Political Science and Director of the European Union Center at Texas A&M University*

The Methodological Tools in the Social Sciences series is comprised of accessible, stand-alone treatments of methodological topics encountered by social science researchers. The focus is on practical instruction for applying methods, for getting the methods right. The authors are leading researchers able to provide extensive examples of applications of the methods covered in each book. The books in the series strike a balance between the theory underlying and the implementation of the methods. They are accessible and discursive, and make technical code and data available to aid in replication and extension of the results, as well as enabling scholars to apply these methods to their own substantive problems. They also provide accessible advice on how to present results obtained from using the relevant methods.

OTHER BOOKS IN THE SERIES

Eric Neumayer and Thomas Plümper, *Robustness Tests for Quantitative Research*

Interaction Models

Specification and Interpretation

William Roberts Clark

Texas A&M University

Matt Golder

Pennsylvania State University



CAMBRIDGE
UNIVERSITY PRESS

Cambridge University Press & Assessment
978-1-108-41671-9 — Interaction Models
William Roberts Clark, Matt Golder
Frontmatter
[More Information](#)



CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom
One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India
103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781108416719

DOI: 10.1017/9781108241762

© William Roberts Clark and Matt Golder 2023

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press & Assessment.

First published 2023

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Names: Clark, William Roberts, 1962– author. | Golder, Matt, author.

Title: An introduction to interaction models / William Roberts Clark, Matt Golder.

Description: New York : Cambridge University Press, [2023] |

Series: MTSS methodological tools in the social sciences

Identifiers: LCCN 2023006720 (print) | LCCN 2023006721 (ebook) |

ISBN 9781108416719 (hardback) | ISBN 9781108404082 (paperback) |

ISBN 9781108241762 (epub)

Subjects: LCSH: Social sciences—Statistical methods. | Political science—Statistical methods. | Regression analysis. | Qualitative research—Methodology.

Classification: LCC HA29 .C596 2023 (print) | LCC HA29 (ebook) |

DDC 519.2—dc23/eng/20230505

LC record available at <https://lcn.loc.gov/2023006720>

LC ebook record available at <https://lcn.loc.gov/2023006721>

ISBN 978-1-108-41671-9 Hardback

ISBN 978-1-108-40408-2 Paperback

Cambridge University Press & Assessment has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Cambridge University Press & Assessment
978-1-108-41671-9 — Interaction Models
William Roberts Clark, Matt Golder
Frontmatter
[More Information](#)

To our most important students:
Meaghan, Brian, Liam, Cameron, and Sean.

Contents

List of Figures	page xiii
List of Tables	xviii
Preface	xxi
Acknowledgments	xxvii

1 Introduction	1
1.1 Resource Endowments	2
1.2 Institutions	4
1.3 Culture/Identity	7
1.4 Plan of the Book	10

PART I The Fundamentals

2 Theories and Their Conditional Implications	15
2.1 How Do I Know When My Theory Posits a Conditional Relationship?	15
2.2 The Difference between Moderation and Mediation	18
2.3 How Does a Multiplicative Term Capture Conditionality?	21
2.4 Moving beyond the Interaction Effect	28
2.5 The Symmetry of Interaction and Its Implications for Theory Testing	36
2.6 Five Key Predictions	44
2.7 Exercises	50
3 Interaction Model Specification	53
3.1 Include All Constitutive Terms	54
3.2 Why Is It Important to Include All of the Constitutive Terms?	56
3.2.1 Multicollinearity	60
3.3 An Alternative Model Specification When <i>One</i> of the Modifying Variables Is Discrete	64
3.3.1 Can I Just Split My Sample?	72
3.4 A Different Alternative Model Specification when <i>Both</i> Modifying Variables Are Discrete	78

3.5	Exercises	89
4	Interpreting Quantities of Interest	98
4.1	Think about Effects in terms of Derivatives and Differences	99
4.2	Calculate Appropriate Measures of Uncertainty	105
4.2.1	Effects and Measures of Uncertainty	105
4.2.2	Predicted Values and Measures of Uncertainty	111
4.2.3	Substantive Significance	114
4.3	Key Quantities of Interest and Some Prototypical Results	115
4.4	Exercises	123
5	Three Substantive Applications	126
5.1	When X and Z Are Both Discrete: Gender, Race, and Support for the Republican Party	126
5.2	When X Is Continuous and Z Is Discrete: Ideology, Race, and Support for Barack Obama	146
5.3	When X and Z Are Both Continuous: Demand and Supply Effects on Women's Legislative Representation	175
5.4	Exercises	191

PART II More Complex Forms of Conditionality

6	When We Have More Than One Modifying Variable	195
6.1	When the Modifying Effects of Z and W Are Independent	196
6.1.1	Substantive Application: Gender, Education, Age, and Support for Feminism	213
6.2	When the Modifying Effects of Z and W Are Dependent	225
6.2.1	An Alternative Interaction Model When the Modifying Variables Are All Discrete	249
6.2.2	Substantive Application: The Impact of Demand, Supply, and Regime Type on Women's Legislative Representation	254
6.3	Exercises	269
7	When an Independent Variable Interacts with Itself	273
7.1	Polynomial Regression Models	274
7.1.1	Substantive Application: The Impact of Party Ideology on Campaign Sentiment	288
7.1.2	Using Polynomial Regression to Model Non-linear Interaction Effects	293
7.2	Possible Threshold Effects	301

Contents

xi

7.2.1	Piecewise Linear Regression Models	301
7.2.2	Switching Regression Models	306
7.3	Exercises	309

PART III Interactions and Limited Dependent Variables

8	Interactions and Dichotomous Dependent Variables	315
8.1	The Linear Probability Model	315
8.2	The Basis for Logit and Probit Models	317
8.2.1	The Pure Probability Approach	317
8.2.2	The Latent Variable Approach	320
8.2.3	The Random Utility Approach	321
8.3	Interpretation and Interaction Effects in “Additive” Logit and Probit Models	323
8.3.1	Coefficients	323
8.3.2	Marginal Effects on Probabilities	325
8.3.3	Predicted Probabilities	329
8.3.4	Differences in Predicted Probabilities	331
8.3.5	Odds Ratios	334
8.3.6	Interaction Effects	335
8.4	Interpretation and Interaction Effects in “Interactive” Logit and Probit Models	347
8.4.1	Coefficients	350
8.4.2	Marginal Effects on Probabilities	351
8.4.3	Predicted Probabilities	355
8.4.4	Differences in Predicted Probabilities	356
8.4.5	Odds Ratios	359
8.4.6	Interaction Effects	361
8.5	Measures of Uncertainty	371
8.5.1	The Method of Simulated Coefficients	372
8.5.2	The Bootstrap Method	377
8.5.3	The Delta Method	377
8.6	Substantive Application: Determinants of Pre-electoral Coalition Formation	379
8.7	Exercises	397
9	Interactions and Ordered Dependent Variables	401
9.1	A Latent Variable Approach	402
9.2	Interpretation	405
9.2.1	Coefficients	405
9.2.2	Marginal Effects on Probabilities	408
9.2.3	Differences in Predicted Probabilities	411

9.2.4	Odds Ratios	412
9.2.5	Interaction Effects	414
9.3	Substantive Application: Ideology, Race, and Presidential Approval of Barack Obama	418
9.4	Exercises	429
10	Interactions and Unordered Dependent Variables	433
10.1	A Random Utility Approach	434
10.2	Data Structure and Interactions	437
10.3	Interpretation	444
10.3.1	Coefficients	445
10.3.2	Marginal Effects on Probabilities	447
10.3.3	Differences in Predicted Probabilities	452
10.3.4	Odds Ratios	453
10.3.5	Interaction Effects	458
10.4	Substantive Application: Policy Preferences, Gender, and Party Support	462
10.5	Exercises	489
	Appendix A. Basic Properties of Variances	493
	Appendix B. Marginal Effects and Variances for Various Linear- Interactive Models	495
	Appendix C. Calculating the Smallest Standard Error for the Marginal Effect of X on Y	498
	Appendix D. Calculating the Values of the Modifying Variable Z at which the Bounds of the Confidence Interval for the Marginal Effect of X Equal 0	500
D.1	Substantive Example: Marginal Effect Plot for <i>Demand</i>	502
	References	506
	Solutions	524
	Alphabetical Index	578

Figures

2.1	Possible causal relationships: <i>Moderation</i>	page 17
2.2	Possible causal relationships: <i>Mediation</i>	19
2.3	Graphical illustration of a linear-additive model consistent with the <i>Unconditional Aid Hypothesis</i>	22
2.4	Graphical illustration of a linear-additive model consistent with Eq. 2.2 where $\gamma_0, \gamma_1, \gamma_2 > 0$	23
2.5	Graphical illustration of a linear-interactive model that's consistent with the <i>Conditional Aid Hypothesis</i>	25
2.6	Marginal effect plot for <i>Aid</i> that's consistent with the <i>Conditional Aid Hypothesis</i> , $\beta_1 = 0$ and $\beta_3 > 0$	29
2.7	(a)–(c) Three very different marginal effect plots for <i>Aid</i> that each have the same slope, $\beta_3 = 0.04$	31
2.8	(a)–(j) Ten possible conditional relationships between <i>X</i> and <i>Y</i> in a theory positing interaction between <i>X</i> and <i>Z</i>	35
2.9	The symmetry of interaction	36
2.10	(a)–(c) Three conditional relationships among <i>Aid</i> , <i>Good Policy</i> , and <i>Growth</i> consistent with the marginal effect plot for <i>aid</i> in Figure 2.7a (assuming $\beta_0 = 4$, $\beta_1 = 0$, and $\beta_3 = 0.04$)	42
2.11	(a)–(b) Marginal effect plots consistent with hypotheses $H_{X Z}$ and $H_{Z X}$	47
3.1	The consequences of omitting a constitutive term	58
3.2	(a)–(b) Predicted values and effects from the standard and alternative interaction models	84
4.1	(a)–(b) Marginal effect plots consistent with hypotheses $H_{X Z}$ and $H_{Z X}$	119
4.2	(a)–(f) Plots of the marginal effect of <i>X</i> on <i>Y</i> across the observed range of <i>Z</i> reflecting several prototypical sets of empirical results	120
5.1	The conditional effects of gender and race on support for the Republican Party in the 2016 US presidential elections	136

5.2	Predicted values and the conditional effects of gender and race on support for the Republican Party in the 2016 US presidential elections	139
5.3	The conditional effect of ideological incongruence on support for Barack Obama in the 2012 US presidential elections	156
5.4	The conditional effect of <i>Black</i> on <i>Feeling Thermometer Obama</i> across various values of <i>Ideological Incongruence</i>	161
5.5	The conditional effect of <i>Black</i> on <i>Feeling Thermometer Obama</i> across the observed range of <i>Ideological Incongruence</i>	163
5.6	The conditional effects of <i>Ideological Incongruence</i> and <i>Black</i> on support for Obama	167
5.7	The conditional effects of <i>Ideological Distance</i> and <i>Black</i> on support for Obama (with controls)	174
5.8	The conditional effects of <i>Demand</i> and <i>Supply</i> on the level of women's legislative representation	181
5.9	Visualizing the relationship between <i>Supply</i> , <i>Demand</i> , and <i>Women's Representation</i>	185
6.1	(a)–(b) Two modifying variables: Independent and dependent modifying effects	196
6.2	A 2-D marginal effect plot for X when $W = W^*$	202
6.3	A 2-D marginal effect plot for X consistent with hypothesis $H_{X Z,W}$ when Z is continuous and W is dichotomous	203
6.4	A 2-D marginal effect plot for X consistent with hypothesis $H_{X Z,W}$ when Z and W are both dichotomous	205
6.5	The marginal effect of X when Z and W are both dichotomous	206
6.6	A 3-D plot of the marginal effect of X on Y consistent with hypothesis $H_{X Z,W}$ (assuming $\beta_1 = 4$, $\beta_4 = 0.4$, and $\beta_2 = -0.2$)	207
6.7	A marginal effect plot for Z consistent with hypothesis $H_{Z X}$	210
6.8	A marginal effect plot for W consistent with hypothesis $H_{W X}$	211
6.9	The conditional effects of gender on support for feminism across age and educational attainment	220
6.10	Differences in support for feminism between women and men	223
6.11	Visualizing a fully interactive relationship between X , Z , and W	227
6.12	A different 2-D marginal effect plot for X when $W = W^*$	237

List of Figures

xv

6.13	A different 2-D marginal effect plot for X consistent with hypothesis $H_{X Z,W}$ when Z is continuous and W is dichotomous (0/1)	238
6.14	A different 2-D marginal effect plot for X consistent with hypothesis $H_{X Z,W}$ when Z and W are both dichotomous	241
6.15	The marginal effect of X when Z and W are both dichotomous	242
6.16	A 3-D plot of the marginal effect of X on Y consistent with hypothesis $H_{X Z,W}$ (assuming $\beta_1 = 15$, $\beta_4 = 7$, $\beta_5 = 3$, and $\beta_7 = -0.6$)	243
6.17	Evaluating <i>Demand Hypothesis I</i>	261
6.18	The marginal effect of <i>Demand</i>	265
6.19	The substantive effect of <i>Demand</i> on women's legislative representation	267
7.1	(a)–(c) Three possible non-linear relationships between X and Y	275
7.2	Different combinations of signs for β_1 and β_2 in a quadratic model	282
7.3	The difference between a derivative and a difference when the effect of X is non-linear	286
7.4	(a)–(c) Party left-right ideology and positive campaign sentiment	292
7.5	(a)–(b) An example of a linear interaction effect	296
7.6	(a)–(c) Three possible non-linear relationships for how the marginal effect of X on Y varies with the value of Z	298
7.7	Another possible non-linear relationship for how the marginal effect of X on Y varies with the value of Z	300
7.8	(a)–(b) Two possible continuous piecewise linear relationships between X and Y	302
7.9	(a)–(b) Two non-continuous piecewise linear relationships between X and Y	307
8.1	The cumulative standard logistic (smooth) and normal (dashed) distributions	319
8.2	Predicted probability that $Y = 1$ from the logit model in Eq. 8.29 where $\beta_0 = -0.4$, $\beta_1 = 0.2$, $\beta_2 = -0.6$, and $\beta_3 = -0.2$.	330
8.3	The effect of a one-unit increase in X_1 based on the logit model in Eq. 8.28 where $\beta_0 = -0.4$, $\beta_1 = 0.2$, $\beta_2 = -0.6$, and $\beta_3 = -0.2$	333

8.4	Different signs for the compression-based interaction effect between X_1 and X_2 on $Pr(Y)$ in the additive logit model shown in Eq. 8.40	337
8.5	Logit models with and without appreciable compression in the observed range $[0, 8]$ for X_1 and X_2	344
8.6	Different quantities from the logit model in Eq. 8.51 where $\beta_0 = -0.4$, $\beta_1 = 0.4$, $\beta_2 = -0.6$, $\beta_3 = -0.2$, and $\beta_4 = 0.5$	353
8.7	The effect of a one-unit increase in X_1 in terms of differences in probability	358
8.8	The effect of a one-unit increase in X_1 in terms of the odds of Y	360
8.9	Interaction effects from the logit model in Eq. 8.51 where $\beta_0 = -0.4$, $\beta_1 = 0.4$, $\beta_2 = -0.6$, $\beta_3 = -0.2$, and $\beta_4 = 0.5$	371
8.10	The conditional effects of <i>Disproportionality</i> and <i>Polarization</i> on the propensity of pre-electoral coalition formation	385
8.11	Expected coalition size and the propensity of pre-electoral coalition formation	387
8.12	(a)–(c) The effect of <i>Disproportionality</i> on the probability of pre-electoral coalition formation	389
8.13	The conditional effect of <i>Disproportionality</i> on the probability of pre-electoral coalition formation	391
8.14	The modifying effect of party system polarization on the effect of electoral system disproportionality on the probability of pre-electoral coalition formation	393
8.15	The effect of a one-standard deviation increase in <i>Polarization</i> on the odds of pre-electoral coalition formation	394
8.16	(a)–(c) The effect of <i>Coalition Size</i> on the probability of pre-electoral coalition formation	396
9.1	Mapping the unobserved latent dependent variable Y^* onto the observed ordered choice dependent variable Y	403
9.2	Probabilities in the ordered logit model	404
9.3	Effect of increasing X_1 on the probabilities from an ordered logit model when the marginal effect of X_1 on Y^* is positive	407
9.4	The conditional effects of <i>Liberal-Conservative</i> and <i>Black on Obama Approval</i> *	423

List of Figures	xvii
9.5 The marginal effect of <i>Liberal-Conservative</i> on the probability of different approval categories for Obama among Whites and Blacks	425
9.6 The effect of being Black on the odds of an approval category greater than some category k versus lower approval categories	429
10.1 The conditional effect of ideological distance on voter utility by gender in the 1992 legislative elections in the United Kingdom	474
10.2 The conditional effect of being female on voter utility for the Conservative Party during the 1992 legislative elections in the United Kingdom I	476
10.3 The conditional effect of being female on voter utility for the Conservative Party during the 1992 legislative elections in the United Kingdom II	477
10.4 (a)–(f) The effect of being female on the probability of voting for the various parties in the 1992 British general elections	486
10.5 The effect of being female on the odds of voting for the Conservatives rather than Labour in the 1992 British general elections	487
D.1 The conditional effects of <i>Demand</i> on the level of women’s legislative representation	503
S.1 The effect of <i>Symbolic Racism</i> on <i>Oppose Race Policies</i>	543
S.2 The effect of <i>Anger</i> on <i>Oppose Race Policies</i>	544
S.3 The sign of the coefficient on X^3	549
S.4 Cubic regression model, the marginal effect of X , and the modifying effect of X	552

Tables

3.1	Comparing the standard and alternative interaction models when one modifying variable, <i>Z</i> , is dichotomous: Five key predictions	<i>page 69</i>
3.2	Comparing the standard and alternative interaction models when both modifying variables, <i>X</i> and <i>Z</i> , are dichotomous: Five key predictions	87
3.3	Social divisions, electoral system permissiveness, and the number of parties	91
3.4	Political responsiveness of niche and mainstream parties	92
3.5	Symbolic racism, anger, and opposition to race-conscious policies I	94
3.6	Gender, race, and support for the Republican Party in the United States	96
4.1	Symbolic racism, anger, and opposition to race-conscious policies II	124
5.1	Gender, race, and support for the Republican Party in the United States at the 2016 presidential elections	131
5.2	Key predictions, derivatives, and quantities of interest	150
5.3	Ideology, race, and support for Barack Obama in the 2012 US presidential elections	151
5.4	The conditional effect of <i>Black</i> on support for Obama at different values of <i>Ideological Incongruence</i>	159
5.5	Ideology, race, and support for Barack Obama in the 2012 US presidential elections (with controls)	173
5.6	Demand, supply, and women's legislative representation	179
5.7	Symbolic racism, anger, and opposition to race-conscious policies III	192
6.1	Gender, educational attainment, age, and support for feminism I	219

List of Tables	xix
6.2 Comparing the standard and alternative interaction models when the modifying variables X , Z , and W are dichotomous: Nineteen key predictions	253
6.3 Demand, supply, regime type, and women's legislative representation I	258
6.4 Gender, educational attainment, age, and support for feminism II	270
6.5 Demand, supply, regime type, and women's legislative representation II	271
7.1 Party left-right ideology and positive campaign sentiment	291
8.1 Calculating quantities of interest via simulation	376
8.2 Determinants of pre-electoral coalition formation I	384
8.3 Determinants of pre-electoral coalition formation II	398
8.4 Determinants of pre-electoral coalition formation III	400
9.1 Ideology, race, and presidential approval for Barack Obama in 2012	422
9.2 The effect of a one-unit increase in <i>Liberal-Conservative</i> on the probability of different approval categories for Obama among Whites and Blacks	427
9.3 Prejudicial attitudes, elite speech, and expressed prejudice	432
10.1 Incomplete data structure for a multinomial logit model	438
10.2 Example data structure for a multinomial logit model I	439
10.3 Example data structure for a multinomial logit model II	442
10.4 Determinants of party vote choice in the 1992 British general election I	472
10.5 The marginal effect of increasing <i>Government Intervention Distance</i> for the Conservative Party on the probability of voting for different parties in the 1992 United Kingdom legislative elections	481
10.6 The effect of a two-unit increase in <i>Government Intervention Distance</i> for the Conservative Party on the probability of voting for different parties in the 1992 United Kingdom legislative elections	482
10.7 Determinants of party vote choice in the 1992 British general election II	490
B.1 Marginal effects and variances for various linear-interactive models	496
B.2 Marginal effects and variances for various linear-interactive models (quadratic terms)	497

xx **List of Tables**

S.1	Social divisions, electoral system permissiveness, and the number of parties	530
S.2	Political responsiveness of niche and mainstream parties	532
S.3	Symbolic racism, anger, and opposition to race-conscious policies I	533
S.4	Gender, race, and support for the Republican Party in the United States	534
S.5	Symbolic racism, anger, and opposition to race-conscious policies II	536
S.6	Gender, educational attainment, age, and support for feminism III	546
S.7	Demand, supply, regime type, and women's legislative representation III	547
S.8	The marginal effect of <i>Prejudice</i> on the probability of different values of <i>Expressed Prejudice</i>	565
S.9	The effect of <i>Prejudiced Speech</i> on the probability of different values of <i>Expressed Prejudice</i>	568



Preface

This book is about how to use interaction models to test the conditional implications of our social science theories. As we argue in more detail in Chapter 1, the radical interdependence of humans who live together means that many of the hypotheses that we can derive from our theories across a broad array of topics throughout the social sciences are likely to be context dependent. This explains why conditional claims such as “an increase in X is associated with an increase in Y when condition Z is present, but not otherwise” are so ubiquitous across the various fields in political science, economics, sociology, psychology, and other social science disciplines. It’s well established that the intuition behind conditional claims and context dependence can be captured by interaction models (Wright, 1976; Friedrich, 1982; Aiken and West, 1991). Unfortunately, the implementation of interaction models is often flawed, and inferential errors are common.

Our mutual interest in interaction models is long-standing. We’ve been discussing the appropriate specification and interpretation of interaction models with each other for over two decades now. Our continued interest in interaction models is often met with bemusement by many of our colleagues, especially those trained in more computationally complex or “cutting-edge” quantitative methods. We suspect that their bemusement stems from the widespread belief, probably arising from the fact that interaction models are often introduced early on in someone’s quantitative methods training, that interaction models are easy and well-understood. When we started writing our first methods paper on interaction models in the early 2000s, a senior methodology colleague openly questioned its utility. “Everybody already knows how to deal with interaction models” was the basic response. We were just wasting our time. He was, of course, trying to be helpful. Our own experience reviewing papers and reading published articles, though, made us much less sanguine about the quality of social science research based on the use of interaction models. We ignored our colleague’s advice and went ahead with our paper anyway.

To motivate it, we conducted a survey of the literature in our home discipline of political science where we systematically examined the use of interaction models in articles published between 1998 and 2002 by

the three leading non-specialized political science journals: the *American Journal of Political Science*, the *American Political Science Review*, and the *Journal of Politics*. Contrary to the beliefs of our senior methodology colleague, we found that just ten percent of the articles that we identified as using an interaction model followed all four of the basic “best practice” recommendations that we outlined in our paper. The results from our survey suggested that there was considerable potential for inferential errors in the articles we examined, something that was troubling as many had been written by the discipline’s leading figures and had gone on to generate substantial research agendas. The message was clear. Although we as a discipline may well have thought we knew how to use interaction models, we obviously didn’t. At least when it came to common practice. Our experience as reviewers and communication with others told us that political science was not unique in this regard.

Our paper, which contained a simple checklist of dos and don’ts for using interaction models, was published in *Political Analysis* in 2006 (Brambor, Clark and Golder, 2006). It quickly became clear that there was huge pent-up demand across the social sciences for advice on how to improve empirical analyses involving interaction models. Evidence for this comes from the fact that our paper is the third most cited article in political science published in the 2000s and, according to some measures, in the top ten of all political science articles ever published (http://charlesbreton.ca/assets/PS_Top10_2020.pdf). Other recent publications that offer additional advice on interaction models have also proven to be very influential (Ai and Norton, 2003; Kam and Franzese, 2007; Berry, DeMeritt and Esarey, 2010; Berry, Golder and Milton, 2012; Hainmueller, Mummolo and Zu, 2019). The result is that the quality of social science research that uses interaction models has improved significantly in recent years. For example, we’re now much less likely to see scholars inappropriately omit “constitutive terms” from their interaction models or incorrectly interpret these terms as capturing unconditional effects. And scholars are much more likely to employ graphical techniques such as “marginal effect plots” to evaluate their conditional claims and provide substantively meaningful information about the effects of their variables.

While this progress is substantial and obviously welcome, simple specification and inferential errors with interaction models remain stubbornly commonplace. One thing we’ve noticed in recent years is that scholars will often cite our paper or one of the others offering methodological advice on the use of interaction models (perhaps because they feel that reviewers require this) but fail to actually implement the recommended practices. Indeed, it’s not unusual to see scholars cite our 2006 paper to support

practices that we explicitly note are unnecessary or, worse, inappropriate. This is disappointing.

In general, widespread confusion persists about certain aspects of interaction models. As an example, there are, as we'll show, alternative ways of specifying the exact same interaction model when one or more of the interacting variables is discrete. Scholars are often unaware of this equivalence or what it means for interpretation. The result can be that some researchers aren't aware that they're estimating an interaction model (Reingold, Haynie and Widner, 2020). As another example, few scholars appreciate the fact that there are two distinct sources of interaction when we estimate an interaction model with a limited dependent variable and the issues that this poses for interpretation and hypothesis testing. As we'll demonstrate, these confounding sources of interaction relate to the *variable-specific* interaction that arises from the inclusion of an explicit interaction term and the *compression-based* interaction that results from the inherent non-linearity that links the outcome variable to the independent variables in these types of models. Considerable uncertainty also surrounds exactly how to specify and interpret interaction models when we move beyond evaluating the simple two-variable interactions typically examined in pedagogical pieces on the use of interaction models to test more complex claims of conditionality. What's the best way to evaluate conditional claims when more than two variables interact or when a variable interacts with itself? What's the best way to present the results in these settings? It's also the case, in our experience, that scholars frequently fail to present all of the quantities of interest necessary to test their conditional claims or expose their theories to as strong an empirical test as is possible given the available data. This often has to do with a failure to think through all of the key predictions that can be derived from their underlying theory.

Our Approach

Our goal in writing this book is to provide a comprehensive and unified introduction to the use of interaction models and how they can be used to test theoretically-derived claims of conditionality. We take an “empirical implications of theoretical models (EITM) style” approach where we emphasize the importance of closely integrating the theoretical and empirical components of social science research. Throughout the book, for example, we always try to make a strong connection between the conditional implications that can be derived from our theories and the interaction model specification and quantities of interest necessary to fully evaluate our conditional claims. A consequence is that we discuss theory

and its implications for empirical analysis much more than is usually the case in methods pieces dealing with interaction models.

The fact that the book is designed to help scholars to better use interaction models in situations where the primary goal is theory or hypothesis testing rather than, say, prediction is reflected in the types of interaction models and estimation techniques that we cover. To be specific, our book deliberately focuses on the types of regression-based parametric models that are most commonly used by applied researchers interested in theory testing and drawing inferences. There are, of course, other more sophisticated estimation techniques, such as neural networks (Zeng, 1999; Beck, King and Zeng, 2000), generalized additive models (Hastie and Tibshirani, 1986; Beck and Jackman, 1998), kernel regularized least squares (Hainmueller and Hazlett, 2014), tree-based models (Quinlan, 1986; Green and Kern, 2012; Montgomery and Olivella, 2018), and support vector machines (Vapnik, 1995, 1998; D’Orazio et al., 2014), that can be used to capture highly complex forms of interaction and conditionality. These methods, though, many of which come out of the statistical and machine learning literature (Hastie, Tibshirani and Friedman, 2017) and depart from the familiar regression framework, are arguably more suited to model-fitting problems related to prediction and classification rather than problems related to theory testing and inference. It can often be challenging to meaningfully interpret exactly what’s going on in these models and derive the types of quantities of interest that applied scholars require to test their theoretical claims.

Who Is This Book For?

Our book is written for people who are interested in formulating contextual theories and testing conditional or “context-dependent” hypotheses using quantitative methods. Given the ubiquity of conditional relationships in the study of human behavior, we suspect that scholars from across the social sciences will find something of value in reading this book – or, at least, that’s our hope. We’ve assumed that readers have some grounding in the basics of statistical inference and ordinary least squares (OLS) regression but not much beyond that. This means that our book could be used by most graduate students, faculty, and researchers in the social sciences. We suspect that it could also be used by some advanced undergraduate students who’ve taken at least one class in quantitative methods.

We’ve attempted to make the material in the book as accessible as possible. We’ve tried, for example, to build our approach to thinking about interaction models from the ground up, starting with what we call “the fundamentals” before introducing more “complex” forms of interaction

Preface

xxv

in an incremental and systematic fashion. In contrast to many methods books, we try to show more of the intermediary steps that are involved in deriving and calculating particular quantities of interest and not just the final equations themselves. Where possible, we also try to provide the intuition behind the math. In addition, we offer practical advice on how to state conditional hypotheses and present quantities of interest in the most effective manner. Call-out boxes keep track of particularly important points. Throughout the book, we provide detailed substantive applications showing how each of the techniques and cases that we cover can be implemented in practice. The replication code, written in both Stata and R, for all of the substantive applications and exercises is available online at <http://mattgolder.com/>. We've tried to minimize pre-written packages and commands in our code in order to make the calculations underlying the methods clearer.¹

How to Use This Book

We believe that our book would be ideal for a class that focuses on how to formulate contextual theories and test context-dependent hypotheses. Of course, we recognize that such a class, while likely to be valuable and interesting, is rarely taught, even in our own departments and institutions. Given this, we suspect that our book will most likely be used as supplementary reading in various quantitative methods classes and as a professional reference for applied researchers. In terms of quantitative methods classes, the material covered in *Part I: The Fundamentals* is particularly well-suited to an introductory class on regression analysis dealing with ordinary least squares estimation. Depending on the level of this class, the material introduced in *Part II: More Complex Forms of Conditionality* may also be appropriate. This additional material would certainly fit well in an advanced regression analysis class where students start to move beyond the simple examples typically covered in an introductory class. The material examined in *Part III: Interactions and Limited Dependent Variables* will be of greatest value in a standard maximum likelihood estimation class that introduces students to a variety of limited dependent variable models or in a class that specifically focuses on discrete choice models. As you can see, we think that our book will be of use to students in several of the classes that have usually made up a significant part of the traditional “methods sequence” in graduate programs in the social sciences. Given this,

¹ The figures that appear in this book were created using the PGFPLOTS package in L^AT_EX (Feuersänger, 2010). The code for them can also be found online at <http://mattgolder.com/>.

xxvi **Preface**

we suspect that students may want to obtain a copy of our book in the first year of graduate school so that they can come back to it again and again as they progress through their required quantitative methods classes.

The book can also be used as a professional reference for applied researchers who seek advice on how to appropriately specify and interpret an interaction model for a particular substantive application. Over the years, we've been contacted by researchers from around the world who've been looking for advice on how to test specific types of conditional claims. They're usually reaching out to us because their particular conditional claim involves some kind of deviation from the scenario covered in the typical methods piece on interaction models where there's an interaction between two continuous independent variables and we have a continuous dependent variable. Perhaps one or more of their independent variables is discrete, there are more than two variables interacting, or there's some kind of limited dependent variable involved. If our book had been completed earlier, we could have simply referred these researchers to the relevant sections. Indeed, it's partly because so many scholars have contacted us with these sorts of requests over the years that we thought a book like this would be useful. We certainly hope that it will be.

Acknowledgments

We've had productive conversations about the use and misuse of interaction models with many scholars over the years. These conversations have been important for clarifying our thinking about how to appropriately test claims of conditionality. In this regard, we're particularly grateful for the discussions we've had with Neal Beck, William Berry, Rob Franzese, and Jonathan Nagler. Numerous individuals have provided comments and useful insights or responded to one of our many queries at various points while writing this book. These include Ray Block, Emma Cohen, Scott Cook, Charles Crabtree, Yaoyao Dai, Kostanca Dhima, Ben Ferland, Gilles Godefroy, Jerg Guttman, Boyoon Lee, Howard Liu, Eric Plutzer, and Chris Zorn. We're especially grateful to Garrett Glasgow for being such a valuable sounding board for thinking about interactions in the context of discrete choice models, to Ali Kagalwala for translating our Stata code into R, and to Anil Kuleli for help with the index. Special thanks must also go to Sean Golder for patiently working on various bits of the math in the book with his Dad and to Sona Nadenichek Golder for reading more drafts of the chapters than anyone should ever have to read.

We're indebted to Paul Kellstedt and Guy Whitten, the editors of the Methodological Tools in the Social Sciences book series at Cambridge University Press, for encouraging us to write this book in the first place and their enormous patience as we made slow progress in bringing it to fruition. We're also grateful to them for organizing a book conference at Texas A&M University where we received incredibly detailed and thoughtful feedback from our reviewers, Justin Esarey and Vera Troeger, and other participants, especially Dave Armstrong.

Our greatest debt is to our families. Sona and Sean have shown incredible support and patience throughout the entire writing (and non-writing) process. They make life much, much better. Laurie Clark's support during this project's long gestation (much of it during a global pandemic) has been unwavering and unconditional. Thank you for reinforcing and facilitating the positive aspects of life, while inhibiting and mitigating the negative.

The art on the cover is "Linear Space 359" by Amie Adelman (<http://www.amieadelman.com/>)