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
Penn sylvania State University

CAMBRIDGE
UNIVERSITY PRESS

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

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://lccn.loc.gov/2023006720
LC ebook record available at https://lccn.loc.gov/2023006721

ISBN 978-1-108-41671-9 Hardback
ISBN 978-1-108-40408-2 Paperback

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

Contents

Figures

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

Tables

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), kernell 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

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

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 LaTeX (Feuersänger, 2010). The code for them can also be found online at http://mattgolder.com/.

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/)