I

INTRODUCTION

Social scientists study complex phenomena. This complexity requires a wide variety of quantitative methods and tools for empirical analyses. Often, however, social scientists might begin with interest in identifying the simple impact of some variable(s), $\boldsymbol{X}$, on some dependent variable, $\boldsymbol{Y}$. Political scientists might study the effect of socioeconomic status on an individual's level of political participation or the effect of partisanship on a legislator's voting behavior. Scholars of comparative politics might be interested in the effect of electoral rules such as multimember versus single-member districts on the party composition of legislatures. Scholars of international relations might study the effect of casualties on the duration of military conflict. Psychologists might study the effect of personality traits on an individual's willingness to obey authority or the effect of an experimental manipulation of background noise on an individual's ability to solve a problem. Economists might investigate the effect of education on labor-market earnings or the effect of fiscal policy on macroeconomic growth. Sociologists might examine the effect of the number of years an immigrant has lived in a host country on his or her level of cultural and linguistic assimilation. Each of these examples posits a simple relationship between some independent variable and a dependent variable.

One of the simplest empirical model specifications for these types of queries is the linear-additive model. The linear-additive model proposes that a dependent variable has a linear-additive, that is, a simple, constant, unconditional, relationship with a set of independent variables. For each unit increase in an independent variable, the linear-additive model assumes that the dependent variable responds in the same way, under any conditions. Much of the quantitative analysis in print across the social sciences exemplifies this approach.

Such linear-additive approaches address what might be described as a "first generation" question, where researchers seek to establish whether some relationship exists between an independent variable, $\boldsymbol{X}$, and a dependent variable, $\boldsymbol{Y}$. A "second generation" question adds an additional layer of complexity, asking not simply whether some relationship exists between an independent variable and a dependent variable but under what conditions and in what manner such a relationship exists: for example, under what conditions is the relationship greater or lesser? Thus, this slightly more complex question posits that the effect of some variable, $\boldsymbol{X}$, on the dependent variable, $\boldsymbol{Y}$, depends upon a third (set of) independent variable(s), $\boldsymbol{Z}$.[1]

One could imagine adding such a layer of complexity to each of the preceding examples. For example, the political scientist studying the effect of socioeconomic status on political participation might suspect that this effect depends upon the level of party mobilization in an election—the participatory gains from socioeconomic status might be attenuated when political parties do more to mobilize citizens at all levels. The effect of a legislator's partisanship on his or her votes surely depends upon whether bills have bipartisan or partisan sponsorship. The effect of multimember districts on the party composition of legislatures likely depends on a third variable, societal fragmentation. The effect of casualties on the duration of military conflict might depend on domestic economic conditions. The psychologists might expect the effects of certain personality traits on individuals' willingness to obey authority to increase, and of others to decrease, with age, and the effect of background noise on problem-solving ability might depend on how well rested the subject is. The economist studying the returns to education might expect booming macroeconomic conditions to magnify, and slumping ones to dampen, the effect of education on labor-market earnings; and the one studying fiscal policy would predict zero real-growth effects when the public expected policies and nonzero effects only when policies were unexpected. Finally, the sociologist studying immigrant assimilation might expect the years lived in the host country to have a greater effect for immigrants from source countries with smaller diasporas than for immigrants from source countries with

1. For expositional ease and clarity, the discussion that follows primarily focuses on a single variable, x, and a single variable, z, as they relate to a single dependent variable, y. The general claims extend naturally to vectors $\boldsymbol{X}$, $\boldsymbol{Z}$, and $\boldsymbol{Y}$.

larger diasporas, the former perhaps being forced to assimilate more quickly. Social scientists often evaluate such hypotheses using the linear-interactive, or multiplicative, term.[2]

Interaction terms are hardly new to social-science research; indeed, their use is now almost common. Given the growing attention to the roles of institutions and institutional contexts in politics, economics, and society, and the growing attention to how context more generally (e.g., information environments, neighborhood composition, social networks) conditions the influence of individual-level characteristics on behavior and attitudes, interactive hypotheses should perhaps become even more common. However, despite occasional constructive pedagogical treatises on interaction usage in the past, a commonly known, accepted, and followed methodology for using and interpreting interaction terms continues to elude social scientists. Partly as a consequence, misinterpretation and substantive and statistical confusion remain rife. Sadly, Friedrich's (1982) summary of the state of affairs could still serve today:

> while multiplicative terms are widely identified as a way to assess interaction in data, the extant literature is short on advice about how to interpret their results and long on caveats and disclaimers regarding their use. (798)

This book seeks to redress this and related persistent needs. Our discussion assumes working knowledge of the linear-additive regression model.[3] Chapter 2 begins our discussion of modeling and interpreting

2. Scholars also refer to the *interactive term* as the *multiplicative* or *product term,* or the *moderator variable,* depending on the discipline. We use *interactive term* and *multiplicative term* interchangeably. In the field of psychology, distinctions are made between *mediator* and *moderator* variables (Baron and Kenny 1986). The distinction is similar to that made in other disciplines, including sometimes in political science, between *intervening* and *interactive* variables, but this terminology is not consistently applied across disciplines and sometimes not even within disciplines. Our discussion applies to *moderator* and *interactive* variables, which Baron and Kenny (1986) define as "a qualitative . . . or quantitative . . . variable that affects the direction and/or strength of the relation between an independent or predictor variable and a dependent or criterion variable" (1174). We reiterate that interactive terms apply when scholars theorize that z affects the *existence or magnitude of the relationship* between x and y, not when scholars believe that some variable z affects the *level* of some variable x that in turn relates to y. This latter argument represents z as a mediating or intervening variable, and an interaction term is not the appropriate way to model it. Instead, mediation is more appropriately modeled by linear-additive regression in various sorts of path analysis; moderation implies interactions.

3. For a refresher on the linear-additive regression model, the interested reader might consult Achen (1982).

interactive hypotheses. This chapter emphasizes how interactive terms are essential for testing common and important classes of theories in social science and provides several theoretical examples in this regard.

In chapter 3, we offer advice on connecting theoretical propositions that suggest interactive relationships to empirical models that enable the researcher to test those interactive hypotheses. We then show which standard statistical tests (certain common *t*- and *F*-tests) speak to which of the specific hypotheses that are typically nested in interactive arguments. We discuss a generic approach to interpreting the estimation results of interactive models and illustrate its application across an array of different types of interactive relationships where different types and numbers of variables are involved. We also address the presentation of interaction effects. In all cases, we urge researchers to go beyond merely reporting individual *coefficients* and standard-error estimates. Instead, we strongly suggest graphical or tabular presentation of results, including *effect-line* graphs or *conditional-coefficient* tables, complete with standard errors, confidence intervals, or significance levels of those *effects* or *conditional coefficients*. We discuss and provide examples of several types of graphs that facilitate interpretation of interaction effects, including effect-line plots, scatter plots, and box plots. We also provide instructions on how to construct these plots and tables with statistical software commonly used in social science, in addition to specific mathematical formulas for their elements. Our approach underscores the importance of understanding the elementary logic and mathematics underlying models that use interactive terms, rather than simply providing a set of commands for the user to enter mechanically. If students and scholars understand the foundations of this generic approach, then they will be well equipped to apply and extend it to any new theoretical problems and empirical analyses.

In chapter 4, we consider certain general-practice rules for modeling interactions that some previous methodological treatments advise and social scientists often follow. We suggest that some scholars may be misinterpreting these rules, and we argue that such general rules should never substitute for a solid understanding of the simple mathematical structure of interaction terms. For example, "centering" the variables to be interacted, as several methods texts advise, alters nothing important statistically and nothing at all substantively. Furthermore, the common admonition that one must include both x and z if the model contains an xz term is an often-advisable philosophy-of-science guideline—as an application of Occam's razor (that the simplest explanation is to be preferred) and, as a practical matter, such inclusion is usually a much safer adage than exclusion—but it is neither logically nor statistically necessary and not always advisable, much less required.

Chapter 5 discusses some more technical concerns often expressed regarding interactive models. First, we discuss the question of pooled-sample versus separate-sample estimation that arises in every social-science discipline. We show that estimating interactive effects in separate samples is essentially equivalent to estimating them in a pooled sample but that pooled-sample estimation is more flexible and facilitates statistical comparisons even if one might prefer separate-sample estimation for convenience in preliminary analyses. The chapter then discusses nonlinear models. Although all of our preceding discussion addresses multiplicative terms exclusively in the context of linear-regression models, statistical research in social science increasingly employs qualitative or limited dependent-variable models or other models beyond linear ones. We show first that most of the discussion regarding linear-regression models holds for nonlinear models, and then we provide specific guidance for the special case of interactive terms in two commonly used nonlinear models: probit and logit. Finally, we address random-coefficient and hierarchical models. As Western (1998) notes, using multiplicative terms alone to capture the dependence on z of x's effect on y (and vice versa) implicitly assumes that the dependence is deterministic. Yet this dependence is surely as stochastic as any other empirical relationship we might posit in social science, and so we should perhaps model it as such. Many researchers take this need to incorporate a stochastic element as demanding the use of random-coefficient models. Others go further to claim that cross-level interaction terms—that is, those involving variables at a microlevel (e.g., individual characteristics in a survey) and at a more macrolevel (e.g., characteristics of that individual's state of residence)—that do not allow such stochastic elements may be biased. As a consequence, a growing number of scholars recommend the use of hierarchical linear models (HLM) or first-stage separate-sample estimation of microlevel factors' effects followed by second-stage estimation of macrolevel and macrolevel-conditional effects from the first-stage estimates. Actually, separate-sample versus pooled-sample estimation and whether one must apply two-stage or HLM techniques in multilevel data are related issues, and, as we show, under some conditions, the simple multiplicative term sacrifices little relative to these more complicated approaches. Moreover, steps of intermediate complexity can allay those concerns (not quite fully, but likely sufficiently) under a wide array of circumstances. Thus, some of these concerns are, strictly speaking, well founded, but they do not amount to serious practical problems for social scientists as often as one might have supposed.

Finally, chapter 6 provides a summary of our advice for researchers seeking to formulate, estimate, test, and present interactive hypotheses in empirical research.