Political Analysis (2003) 11:209–233
DOI: 10.1093/pan/mpg012

Causal Complexity and the Study of Politics

Bear F. Braumoeller
Department of Government, Littauer Center,
Harvard University, North Yard, Cambridge, MA 02138
e-mail: bfbraum@fas.harvard.edu

Theories that posit complex causation, or multiple causal paths, pervade the study of politics but have yet to find accurate statistical expression. To remedy this situation I derive new econometric procedures, Boolean probit and logit, based on the logic of complexity. The solution provides an answer to a puzzle in the rational deterrence literature: the divergence between theory and case-study findings, on the one hand, and the findings of quantitative studies, on the other, on the issue of the role of capabilities and willingness in the initiation of disputes. It also makes the case that different methodological traditions, rather than settling into "separate but equal" status, can instead inform and enrich one another.

1 Introduction

The logic of testing theories that posit causal complexity, or multiple causal paths, has so far been given short shrift in large-N studies of politics. Such theories are both common and prominent in the political science literature. In qualitative research, the concept of complexity and the methodological problems that it poses have been understood for some time.[1] Formal mathematical models, too, often imply Boolean hypotheses; although the terms "causal complexity" and "multiple causal paths" are rarely used to describe them, the basic logic is no different.

To date, however, no one has made a concerted effort to describe how the empirical implications of theoretical models that posit causal complexity could be captured by statistical methods. While occasional quick fixes (e.g., multiplication of independent variables) demonstrate that there is at least an awareness that standard statistical procedures are lacking, no methodology has been specifically tailored to incorporate the logic of complexity.

This article proposes a remedy. In the following pages I elaborate the concept of causal complexity and demonstrate the breadth of its applicability by discussing prominent

Author's note: A number of people were kind enough to read previous drafts and/or to offer helpful suggestions along the way; I am indebted to Chris Achen, Andrew Bennett, Colleen Braumoeller, Wendy Tam Cho, Kevin Clarke, Paul Diehl, John Freeman, Brian Gaines, Gary Goertz, Iain Johnston, Gary King, Bert Kritzer, Jim Kuklinski, Jack Levy, Everett Lipman, Daniel Little, Ellen Lust-Okar, Lisa Martin, Jim Morrow, Gerry Munck, Bob Pahre, Dan Reiter, Wesley Salmon, Anne Sartori, Theda Skocpol, Rich Snyder, and three anonymous reviewers for having done so, and to Doug Stinnett for research assistance.

[1] See especially Ragin (1987).

research traditions that incorporate its logic. The number and prominence of the theories that hypothesize causal complexity demonstrate that the development of a statistical method appropriate for complexity is a matter of great importance.

I then derive new econometric techniques, Boolean probit and logit, that are designed specifically for use as a test of theories that posit multiple paths to a given (non)outcome.[2] They provide an answer to a long-standing puzzle in the international relations literature: while IR theorists often posit that conflict occurs when at least one state has the capabilities and willingness to initiate it, and the findings of qualitative studies usually support this conclusion, quantitative studies often do not. A reexamination of some of the literature's most recent findings using the technique described herein both confirms the intuition that capabilities and willingness produce conflict and demonstrates that willingness can stem from multiple, independent sources.

More broadly, the solution proposed here represents an attempt to overcome the widespread perception that quantitative and qualitative research must proceed from fundamentally different ontological foundations. Quantitative researchers are often skeptical regarding the utility and applicability of such baroque concepts as causal complexity, INUS conditions, and the like, and qualitative (and formal) researchers accuse their large-N colleagues of being too wedded to additive, linear-in-variable formulations that constitute prosaic models of causation. Too often the result is a "separate but equal" policy. This article seeks genuine accommodation.

2 Causal Complexity

Concrete definitions of causal complexity are difficult to come by, perhaps because the concept is so slippery.[3] Jervis (1997, p. 35) describes a situation in which "the effect of one variable or characteristic can depend on which others are present." Ragin (1987, p. 20) refers to cases in which "an outcome results from several different combinations of conditions."

Ragin's solution to the problem of dealing with causal complexity, qualitative comparative analysis, is of necessity more concrete, and examining its workings reveals the essence of causal complexity: multiple causes interact with one another to produce effects, and the manner in which they interact is described by the logical operators "and" and "or" (pp. 89–93). For example, Ragin describes the mobilization of ethnic groups as being the result of either large size *and* economic advance *or* a strong linguistic base *and* high relative wealth (pp. 142–143).

This formulation reveals an interesting point: many diverse concepts can be understood as special cases of causal complexity. They include

- Multiple conjunctural causation: X_1 and X_2 and X_3 produce Y (Ragin 1987).
- Substitutability: X_1 or X_2 or X_3 produces Y (Most and Starr 1984; Cioffi-Revilla and Starr 1995).[4]

[2] Software for implementing Boolean probit and logit in Stata and R can be downloaded from the page devoted to this issue on the *Political Analysis* Web site (http://pan.oupjournals.org).

[3] In systems theory, this topic goes by the name of "equifinality," a property present in any system in which similar ends can be achieved via different means. In the social science literature, this term is sometimes used (see, e.g., Bennett and George 1997), though "causal complexity" is more common. The philosophy of science literature refers to "causal chains" or "causal ropes," both of which have a prominent and well-established lineage dating back at least to Venn (1866). I find the analogy of causal strands appealing but prefer to take no position on the "chains vs. ropes" (i.e., continuous vs. discrete causation) debate; hence paths.

[4] See also the February 2000 special issue of the *Journal of Conflict Resolution* (44:1) devoted to substitutability in foreign policy.

- Contexts: X_2 produces Y, but only in the presence of X_1 (Goertz 1994).
- Necessary and sufficient conditions: X_1 and X_2 produce Y, X_1 or X_2 produces Y (Dion 1998; Braumoeller and Goertz 2000).[5]
- INUS conditions:[6] (X_1 and X_2) or (X_3 and X_4) produce Y (Mackie 1965).

Complex causation is problematic for the vast majority of existing statistical techniques. The problem is that complexity implies a particular form of nonadditivity that arises from the nature of the cumulation of the impact of the independent variables on the dependent variable: the presence or absence of one independent variable mitigates—or in the extreme, nullifies—the impact of another.

There are some exceptions to this generalization, but very few.[7] In the context of basic random utility models,[8] Achen (2002, p. 439) notes that power logit (Prentice 1976) is implied whenever each of a given number of distinct, equiprobable conditions is necessary for an outcome, and scobit (Nagler 1994) is implied whenever each of a given number of distinct, equiprobable conditions is sufficient for an outcome. The advantages of Boolean probit or logit over either power logit or scobit in modeling complexity in random utility models are twofold: first, they permit the researcher to model *combinations* of Boolean conditions, and second, the extremely restrictive assumption that each of the conditions is equiprobable is relaxed.

In order to understand the advantages of the technique fully, of course, it must first be derived, which requires a more thorough understanding of complex causation. Conjunctural causation and substitutability are the fundamental building blocks of complexity, so I will begin by examining their logic.

2.1 *The Logic of Complexity*

In the case of substitutability, the marginal impact of a change in one independent variable on the probability that the dependent variable will occur decreases as the level of the other independent variables increase (and vice versa). The following are a few examples: Thomas Schelling pointed out that either a strong deterrent or a strong defense could prevent war (1966, p. 78); John Mueller claims that either nuclear weapons or the lessons of World War II would have produced stability during the Cold War (and hence that nuclear weapons were "essentially irrelevant"—1988, p. 58); and James Morrow posits that pursuit of either autonomy goals or security goals can lead to the decision to seek an alliance (1991, p. 905). The

---

[5]Typical formulations would suggest that necessary and sufficient conditions are not statistical problems because they admit no counterexamples and that, as a result, the statistical procedure described hereinafter is inapplicable; Braumoeller and Goertz (2000, p. 848) outline four reasons to expect otherwise.

[6]Mackie coined this term as a definition of a cause; it refers to "an insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result" (p. 245, emphasis in original). In the above example, X_1–X_4 are all INUS conditions. I am indebted to Catharina Wrede Braden for goading me into thinking about how such conditions could be captured by the techniques described below.

[7]The situation envisioned—the (non)occurrence of a particular event—is quite general, in the hopes that the technique used to model it will be useful to a broad audience. Techniques designed for more specific applications exist elsewhere. For example, Western (1998) applies a Bayesian hierarchical model to situations in which the effects of independent variables differ across countries in a pooled cross-sectional time-series study. Manton et al. (1976, 1991) utilize a methodology appropriate for multiple causes of human mortality. The application is interesting but the methodology does not generalize particularly well. One of its core assumptions (humans must eventually die, one way or another) is far too restrictive for the study of most political phenomena: dyads may or may not eventually go to war, states may or may not democratize, citizens may or may not revolt, etc.

[8]A *random utility model* is one in which an actor's utility is hypothesized to depend on a set of known components, captured by the independent variables in the model, as well as on a large number of small and unrelated unknown components, which by virtue of the Central Limit Theorem will tend to sum to a normally distributed error term.

literature on war in particular is replete with substitutable causes, including competition, diffidence, or the desire for glory (Hobbes 1651, p. 70); miscalculation or loss of control (Levy 1983, p. 86); incentives to misrepresent capabilities and resolve or inability to commit credibly to an agreement (Fearon 1995); and the host of scenarios listed chapter by chapter ("War as an Accident," "Death-Watch and Scapegoat Wars," etc.) in the study by Blainey (1973).[9]

What all of these examples have in common is that, as X_1 increases, variation in X_n ($n \neq 1$) has a decreasing impact on Y. In Mueller's case, for example, as the lessons of World War II become increasingly compelling, war among the Great Powers becomes increasingly unlikely and variation in ownership of nuclear weapons has a decreasing impact on peace.[10] The relationship among the independent variables is not additive, so these expectations are at best only approximated by an additive model.

The converse of this scenario is a situation in which only the combination of X_1 and X_2 is believed to have a causal impact on Y—that is, there are multiple paths to the *non*occurrence of the dependent variable. For example, Matthew Evangelista has demonstrated that transnational relations and domestic political structures combined to produce security policy in Russia and the Soviet Union. Change required both a domestic "window of opportunity" and transnational policy entrepreneurs willing to jump through it (Evangelista 1995).[11] Other prominent examples include the "democratic peace" proposition [the assertion that high levels of democracy in both i and j produce peace between the two; see, e.g., Doyle (1983a, b)] and Arthur Lupia and Kaare Strom's finding that coalition dissolution can be expected only in the presence of a critical event, a majority preference for an election over the status quo, and a universal preference for election over "best offer" (Lupia and Strom 1995, p. 655).

In each of these cases, the impact of changes in X_1 on Y depends on X_2 and vice versa. Absent one factor, changes in the other have no impact. In the case of Evangelista's argument, two paths to nonoccurrence of policy change were present: either the absence of a window or the absence of an entrepreneur would have precluded it. Here, again, additivity is a nonsensical assumption. The assumption of additivity, therefore, does not always (or perhaps even usually) reflect the manner in which we and our theories envision the process of causation.[12] Regardless of which type of complexity is in the offing, the effects of the independent variables on the dependent variable, though cumulative, are simply not additive.

Moreover, theories that incorporate causal complexity are quite common. Tables 1 and 2 summarize the hypotheses listed above as well as others. The tables demonstrate that theories that posit causal complexity can be found throughout the discipline. The three main empirical subfields are all represented. The citations span nearly three and a half

[9] Jim Morrow (personal communication) has called into question the meaningfulness of the concept of substitutability by pointing out that events with different sets of causes are often in the end classified as different types of events, in the same way that (for example) snakes and legless lizards are now classified as separate animals because of their distinct evolutionary origins. The caveat is worth noting, though my own belief is that, in political science, wars and revolutions and votes will continue to be seen as similar enough to be studied in the aggregate even if they are found to have multiple sources. In any event, whether or not *some* sets of events are eventually broken up into more meaningful subsets, I find it difficult to conclude, a priori, that *all* will be. That said, I should also note that the econometric difficulties posed by substitutability disappear if the different "paths" to a given outcome do produce qualitatively different kinds of phenomena that can be separated empirically and analyzed separately.

[10] It should be noted that, logically, the converse should be true as well: in the presence of nuclear weapons, the lessons of World War II are, or were, irrelevant.

[11] The origins of the entrepreneur-window analogy are akin to those of the Folk Theorem in game theory: it seems that everyone who uses it can cite an earlier usage.

[12] This point is hardly new—see, e.g., Hoffmann (1959, esp. p. 359) and a rejoinder by Alker Jr. (1966). It has simply been ignored, or perhaps forgotten.

Table 1 Hypotheses that posit substitutability

Substitutable causes	*Effect*	*Source*
International relations		
Competition, diffidence, *or* desire for glory	War	Hobbes (1651, p. 70)
Pursuit of either autonomy goals *or* security goals	Decision to seek alliance	Morrow (1991, p. 905)
Deterrence *or* defense	Prevention of conquest	Schelling (1966, p. 78)
Political isolation *or* geographical encirclement	Civilian intervention in military affairs	Posen (1984, p. 79)
Distinguishability of offensive vs. defensive weaponry *or* defense dominance	Amelioration of the security dilemma	Jervis (1978, pp. 186–187)
Miscalculation *or* loss of control	War	Levy (1983, p. 86)
Lessons of World War II *or* existence of nuclear weapons	Stability during the Cold War	Mueller (1988, p. 58)
Similar attitudes *or* group pressure and loyalty	Agreement in voting in U.N. General Assembly	Lijphart (1963, p. 904)
Comparative politics		
Bourgeois revolution, failed bourgeois revolution, *or* communism	Transition from preindustrialism to modernity	Moore Jr. (1966)
Challenge to legitimacy of old regime *or* challenge to nature of regime	Regime transition	Shain and Linz (1995, p. 3)
Agreed reform within leadership, controlled opening to opposition, *or* collapse of authoritarian regime	Democratization	Colomer (1991)
Lack of resources, adverse political opportunity structure, *or* lack of information	Failure of woman suffrage movements	Banaszak (1996, p. 217)
Deference *or* secularism	Departures from class-based voting (e.g., “Working-class Tory” vote)	McKenzie and Silver (1967)
Exit *or* voice	Recovery of state or organization from lapse	Hirschman (1970)
American politics		
High levels of political information *or* centrality of groups *qua* groups as objects in belief systems	High levels of constraint in belief systems	Converse (1964, pp. 234–238)
Ignorance, indifference, dissatisfaction, *or* inactivity	Nonvoting	Ragsdale and Rusk (1993, pp. 723–724)
Uncertainty *or* institutional pressures	Supreme Court Justices’ choice to change position on a decision	Maltzman and Wahlbeck (1996, p. 581)
Poverty, truancy, *or* exploitation by parents	Juvenile involvement in street trades	Goldmark (1904, p. 424)
High levels of information *or* use of candidate gender as proxy for social information	Individual’s voting decision	McDermott (1997)

Table 2 Hypotheses that posit conjunctural causation

Conjunctural causes	*Effect*	*Source*
International relations		
Democracy of i *and* democracy of j	Peace between i and j	Doyle (1983a, b)
Positive expected utility of conflict for i *and* for j	War between i and j	Bueno de Mesquita (1985, p. 165)
Dissatisfaction *and* power parity	Great power war	Organski and Kugler (1980)
Capabilities *and* resolve of defender	Deterrence success	Stein (1987, p. 333)
Domestic strife *and* threat to military interests	Diversionary aggression	Dassel and Reinhardt (1999)
Rough equality of strength *and* bilateral antagonisms	Arms races	Huntington (1958)
Existence of market coalition with global influence *and* existence of alternative cognitive framework	Regime change	Cowhey (1990, p. 173)
Comparative politics		
Critical event, majority preference for election over status quo, *and* universal preference for election over best offer	Coalition dissolution	Lupia and Strom (1995, p. 655)
Social, economic, and external factors *and* desire of political leaders	Democratization	Huntington (1991, p. 108)
Declining institutional incentives *and* increasing social ties	Electoral stabilization	Bartolini and Mair (1990, p. 303)
Tactical freedom (solidarity and autonomy) of peasants *and* relaxation of state coercion	Peasant revolt	Skocpol (1979)
Scarcity, desire by prospective right-holders, *and* advantages to enforcement of respect for right by right-grantors	Creation of property rights	Riker and Sened (1991)
Existence of human rights network *and* decision of network to focus on particular country	Changes in human rights practices within that country	Sikkink (1993, p. 436)
American politics		
High political awareness and message intensity *and* few predispositions	Mass opinion change	Zaller (1992, p. 157)
Racial heterogeneity *and* lack of geographical proximity	Opposition to intermarriage	Kinder and Mendelberg (1995, p. 419)
Perceived needs *and* political ties	Contacting elected officials	Hirlinger (1992, p. 556)
Presidential initiative *and* ability of President to alter context of Congressional action	Success of President's programs	Peterson (1990)
Discovery of negative information *and* expansion of conflict to committee and media	Defeat of Presidential nomination in Senate	Krutz et al. (1998)
Race, welfare status, *and* perceived effort of proposed recipient	White conservative antagonism toward poverty programs	Sniderman et al. (1996, pp. 40–41)
Relative youth *and* self-identification as gay	Support for disruptive behavior among AIDS activists	Jennings and Andersen (1996, p. 329)

centuries. The hypotheses are derived from both verbal and mathematical theories, and they are examined in both case studies and large-N research. Further examples could be listed, but to do so would merely belabor the point: complexity pervades theorizing in political science.

3 A Solution

The solution I propose to the problem of statistical estimation in the presence of complexity flows directly from the logics of conjunctural causation and substitutability. I will begin by discussing hypothetical examples of both forms of complexity in which the values of the independent variables correspond directly to probabilities and no statistical estimation is necessary; the point will be to make the logic of such situations as clear as possible. I will then generalize to the sorts of independent variables to which social science researchers have become accustomed—that is, those with values that do not correspond directly to known probabilities, making the estimation of coefficients necessary. I will model these probabilities as points somewhere along a cumulative normal curve [denoted by $\Phi(\cdot)$]; let the location of the points be determined by the independent variables (X), their coefficients (β), and a constant term (α); calculate the dependent variable (Y) as a function of those estimated probabilities; and utilize maximum likelihood estimation to obtain parameter estimates.[13] The result will be a new technique—Boolean probit—appropriate for use in testing theories that posit causal complexity.

It would also be possible to combine logit rather than probit curves, simply by replacing $\Phi(\cdot)$ with $\frac{e^{(\cdot)}}{1+e^{(\cdot)}}$. In fact, I utilize Boolean logit in the reanalysis below. However, the probit notation is somewhat more compact, so I use it in the derivation that follows for the sake of clarity.

Following the derivation, I demonstrate the strengths and weaknesses of the Boolean probit and logit procedures in the context of empirical analysis. The main strengths that will be demonstrated are that the procedures provide a way for statistical modelers to test theories that posit even fairly convoluted forms of complexity; that they correctly model causal processes of interest, and in so doing report findings that might very well have been missed by existing techniques; and that they provide a vehicle for theoretical synthesis. There are also two drawbacks, both pertaining to the shape of the likelihood function to be maximized. The first is that complex multidimensional analyses of this sort occasionally produce convoluted likelihood functions with multiple maxima, and so care must be taken to ensure that the maximum eventually reached is global rather than local. The second is that multidimensional analyses in general have more complex and rather demanding data requirements; the absence of sufficient information can produce plateaus in the likelihood function, rendering estimates of a particular coefficient essentially useless. Neither of these problems is unique to these procedures, of course, but they can be expected to crop up more often than they would in standard logit and probit analyses.

3.1 *Conjunctural Causation*

In order to model a process of conjunctural causation we need

1. to model the probabilities that each of the relevant causal factors will occur and then
2. to multiply those probabilities together to obtain the overall probability that the event in question will occur.

[13] It is worth noting that the dependent variable is assumed to measure the occurrence or nonoccurrence of a given event. In practice, the vast majority of complex hypotheses refer to such outcomes (see, e.g., Tables 1 and 2).

To provide the intuition behind this point, consider a brief example. Imagine that Rosencrantz, a piano mover by profession, picks his lunch at random on his way out the door in the morning from a box containing two sausages and one banana. Once at work, he selects a coil of rope, again at random, from a room containing three ropes, one of which is frayed. By noon, Rosencrantz has hoisted the piano up six stories, at which point he takes a break to eat lunch. If he has brought a banana for lunch, he drops the peel, and Guildenstern, an innocent pedestrian, slips on it and falls, landing directly beneath the piano. If Rosencrantz has chosen the frayed rope, the rope breaks and the piano falls to the ground. If (and *only* if) both of these events occur, Guildenstern is killed.

This is an example of conjunctural causation. If the piano does not fall, Guildenstern will survive. Similarly, if Guildenstern does not slip on the peel, he will be clear of the piano when it lands and will not be injured. In combination, however, the two events produce an effect that the sum of their individual inputs would not suggest. Elementary calculations give us the probability that both events will occur, assuming that they are independent of one another: $\frac{1}{3} \times \frac{1}{3} = \frac{1}{9}$. Put intuitively, three choices of lunch and three choices of rope produce nine scenarios, only one of which (banana and frayed rope) is fatal to Guildenstern.

If we use p_y to denote the outcome, p_{x_1} to denote the probability that the first prerequisite will occur, and p_{x_2} to denote the probability that the second prerequisite will occur, the general form for calculating the probability that a process of conjunctural causation will produce an outcome is simply

$$p_y = p_{x_1} \times p_{x_2}. \tag{1}$$

Converting this theoretical expectation into an econometric model suitable for estimation requires two steps. The first is to model the probability p_{x_j} that each of the prerequisites will occur as a function of a constant (α_j), the relevant independent variable (X_j), and its parameter (β_j), where the j subscript distinguishes among different causal paths. I will utilize the standard normal cumulative distribution function $\Phi(\cdot)$ for this purpose.[14] The second is to model the probability that the outcome in question will occur in the ith observation as the product of the probabilities that each of its prerequisites will occur; so

$$\begin{array}{ccccc} p_y & = & p_{x_1} & \times & p_{x_2} \\ \downarrow & & \downarrow & & \downarrow \\ \Pr(y_i = 1 \mid \alpha, \beta, x_i) & = & \Phi(\alpha_1 + \beta_1 x_{i1}) & \times & \Phi(\alpha_2 + \beta_2 x_{i2}). \end{array}$$

Generalizing to J causal paths (to nonoccurrence of Y) gives

$$\Pr(y_i = 1 \mid \alpha, \beta, x_i) = \prod_{j=1}^{J} [\Phi(\alpha_j + \beta_j x_{ij})]. \tag{2}$$

[14] $\Phi(\alpha + \beta X) \equiv \int_{-\infty}^{\alpha+\beta X} \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}}\, du$. This, of course, is simply a probit curve. I should emphasize that functions other than probit or logit curves could just as easily be used. Logit and probit are bounded by 0 and 1, making them ideal for the estimation of probabilities. The only drawback is the requirement that $0 < p_{x_1} < 1$ and $0 < p_{x_2} < 1$; though this requirement is usually unproblematic in practice, if need be the researcher can utilize other functions that are free of it, such as the Burr distribution [see Aldrich and Nelson (1984, pp. 31–35)].

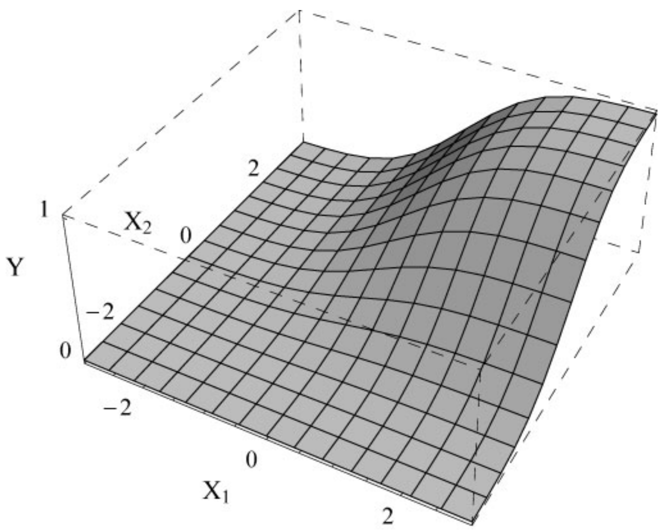

Fig. 1 Conjunctural causation: Multiple paths to failure ($Y = 0$).

The likelihood function becomes

$$L(Y \mid \alpha, \beta, X) = \prod_{i=1}^{N} \left(\prod_{j=1}^{J} [\Phi(\alpha_j + \beta_j x_{ij})] \right)^{y_i} \left(1 - \prod_{j=1}^{J} [\Phi(\alpha_j + \beta_j x_{ij})] \right)^{1-y_i}. \quad (3)$$

It is, of course, possible to complicate matters: one might hypothesize that k independent variables have an additive effect via the same causal path at the same time that other independent variables are having an effect via different causal paths. If so, the generalized likelihood function[15] becomes

$$L(Y \mid \alpha, \beta, X) = \prod_{i=1}^{N} \left(\prod_{j=1}^{J} \left[\Phi\left(\alpha_j + \sum_{k=1}^{K} \beta_{jk} x_{ijk} \right) \right] \right)^{y_i} \times \left(1 - \prod_{j=1}^{J} \left[\Phi\left(\alpha_j + \sum_{k=1}^{K} \beta_{jk} x_{ijk} \right) \right] \right)^{1-y_i}. \quad (4)$$

Figure 1 illustrates this functional form. The figure corresponds to the logic of the "falling piano" example: Guildenstern is in danger (Y, representing the probability of being killed, is high) only when the probability that Rosencrantz will choose a banana for lunch, $\Phi(\alpha_1 + \beta_1 X_1)$, and the probability that Rosencrantz will select a frayed rope, $\Phi(\alpha_2 + \beta_2 X_2)$, are high. As X_1 decreases, so does the marginal impact of changes in X_2 on Y, and vice versa.

[15] At this point, the x_{ijk} notation may befuddle even the most careful reader. As a reminder, the i subscript refers to the observation number, the j subscript refers to the causal path along which the variable exerts an effect, and the k subscript refers to the order of the variables within a given path. So, x_{612} refers to the sixth observation on the second independent variable that is hypothesized to have an effect via the first causal path. The summation symbol may seem odd to large-N methodologists, who are more accustomed to the simpler $\Phi(\beta X)$ notation introduced later in the article; I beg their indulgence here in the interests of clarity.

I should emphasize that there is nothing controversial about this model. In fact, although its use in political science is exceedingly rare,[16] it is a straightforward bivariate probit model with partial observability—a technique with a respectable pedigree in economics and econometrics (Abowd and Farber 1982; Greene 2000).[17] Perhaps because it is only touched on in most texts and is occasionally ignored,[18] the applicability of this technique to cases of multiple conjunctural causation has escaped the notice of political scientists, despite the fact that—as the derivation demonstrates—it is ideally suited to testing such theories.

Still, conjunctural causation is only part of the story: bivariate probit with partial observability is essentially a special case of the more general Boolean probit model. In the next sections I utilize the same method—working from logic to probability calculus to statistical equation—to derive entirely new tests of theories that posit complex causation.

3.2 *Substitutability*

Modeling a process of substitutability is mathematically slightly more complex, but it still follows directly from intuition. Here we need

1. to model the probabilities that each of the relevant causal factors will *not* occur and then
2. to substract the product of those probabilities from 1 to obtain the overall probability that the event in question *will* occur.

To see this point, imagine that Rosencrantz has been convicted of probabilistic manslaughter. In view of the nature of the offense, the punishment is probabilistic as well: he is sentenced to face a firing squad consisting of two expert marksmen, each of whom is given a six-chambered revolver and a single bullet. When the appointed hour arrives he is tied to a pole, blindfolded, and given a cigarette. At the commandant's order the marksmen load their weapons, spin the barrels Russian Roulette-style, take aim, and fire.

This is an example of substitutability.[19] Either of the bullets would bring about his death independently of the other. Though it may not seem obvious at first glance, each pull of the trigger decreases the marginal causal impact of the other: given that the first bullet alone would kill Rosencrantz, and that Rosencrantz cannot be killed more than once, the second bullet would produce no additional effect.[20]

---

[16] Although related models in which one dependent variable is observed if and only if the other takes on a given range of values have come into vogue recently, simple bivariate probit models with a single dependent variable in the political science literature are extremely uncommon—let alone tri- or *n*-variate models. Exceptions to this generalization include Przeworski and Vreeland (2002), who utilize bivariate probit with partial observability, and Reed (2000), who utilizes a variant designed for use with censored data.

[17] It is worth noting that Abowd and Farber's variant of the bivariate probit model was developed to capture a situation in labor economics in which the assumption of little or no interdimensional correlation ($\rho \cong 0$) is a plausible one. Such an assumption may be unjustifiable under some circumstances—for example, in game-theoretic random utility models in which joint observation by the actors of real-world factors that are relevant to utility but that are not captured in the econometric equation imply that error terms will be correlated (Smith 1999). Under such circumstances the Poirier (1980) variant, in which ρ is estimated, is to be preferred.

[18] Schmidt (1984), in his review of the first edition of Maddala (1994), refers to the book's coverage as "encyclopedic" but then takes it to task for its scant coverage of a handful of topics, one of which is Poirier's paper on partial observability in bivariate probit models.

[19] In fact, it is an example that is often utilized. See Good (1985), cited in Salmon (1998, pp. 239–240); Salmon develops the probabilistic variant, while Good's is deterministic (no chambers are left empty).

[20] It is worth mentioning that, for this reason, the deterministic version of redundant causation (six chambers, six bullets) can be problematic for counterfactual causal analyses [à la Tetlock and Belkin (1996)]: there is no doubt

What is the probability that Rosencrantz will be shot? An abstract model of such a situation, straight out of first-year statistic books, is rolling dice: if we throw two dice in an attempt to roll at least one 6, for example, there are two "paths" to rolling a 6, each one corresponding to one of the dice. Obviously, the probability of rolling at least one 6 when throwing more than one die is not obtained simply by adding up the probabilities of rolling a 6 on each die; otherwise, the probability of rolling at least one 6 when throwing seven dice would equal $\frac{7}{6}$. Rather, the probability is obtained by multiplying together the probability of *not* rolling a 6 (that is, $1-\frac{1}{6}$) on each roll and subtracting the total from 1:

$$\Pr(6 \mid \text{two rolls}) = 1 - \left[\left(1-\frac{1}{6}\right)\times\left(1-\frac{1}{6}\right)\right] = \frac{11}{36}. \tag{5}$$

Therefore, though he may find little solace in the fact, Rosencrantz has nearly a 70% probability of surviving if each marksman fires once and only once.

In more general terms, substitutability calculations take the form

$$p_y = 1 - \left[\left(1-p_{x_1}\right)\times\left(1-p_{x_2}\right)\right]. \tag{6}$$

We can use the same logic to derive from the general probability calculus the probability that the outcome in question will occur in the ith observation:

$$\begin{array}{ccccccc} p_y & = 1 - & \left[\left(1-p_{x_1}\right)\right. & \times & & \left.\left(1-p_{x_2}\right)\right] & \\ \downarrow & \downarrow & \downarrow & & & \downarrow & \\ \Pr(y_i = 1 \mid \alpha,\beta,x_i) & = 1 - & \{[1-\Phi(\alpha_1+\beta_1 x_{i1})] & \times & & [1-\Phi(\alpha_2+\beta_2 x_{i2})]\}. & \end{array}$$

Generalizing to J causal paths (to occurrence of the dependent variable) gives

$$\Pr(y_i = 1 \mid \alpha,\beta,x_i) = 1 - \prod_{j=1}^{J}[1-\Phi(\alpha_j+\beta_j x_{ij})], \tag{7}$$

where the j subscript again distinguishes among different causal paths. The likelihood function is then derived from this probability precisely as it was in Eqs. (2)–(4).

Figure 2, which illustrates this functional form, reflects the logic of Rosencrantz's fate at the hands of his probabilistic firing squad. The probability that he will be killed is high (Y approaches 1) if the probability that either executioner's gun will fire is high (X_1 or X_2 is high) but is low otherwise. At the same time, he cannot be killed twice, even if both executioners' weapons are virtually certain to discharge. Boolean probit conforms to these expectations.

3.3 *Combinations and Permutations*

It is possible to model more complex causal phenomena by simply extending the logic outlined above. The most straightforward example would be a combination of the two examples

that one bullet would be sufficient to cause Rosencrantz's death, but subtracting either one of the two from the situation produces no change in the result. It would be a mistake, of course, to conclude that individual bullets have no causal impact on their targets' health. Kiser and Levi (1996, pp. 196–197) briefly mention problems associated with counterfactuals in the context of conjunctural causation.

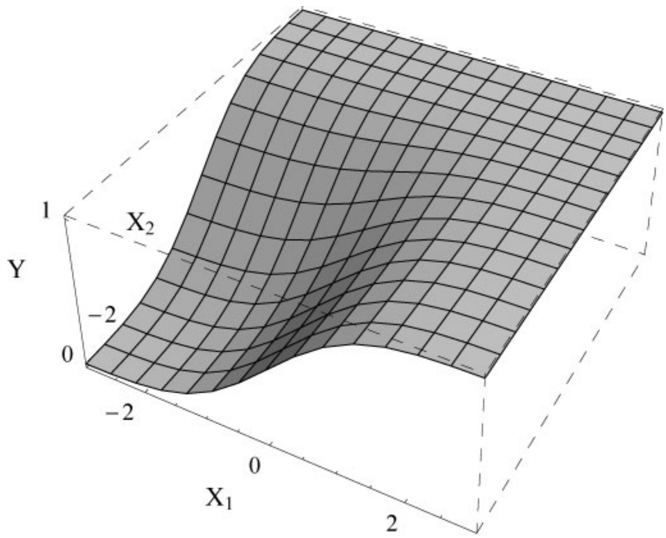

Fig. 2 Substitutability: Multiple paths to success ($Y = 1$).

already given—to wit, given that we know all of the probabilities before Rosencrantz even leaves home, what is the probability that both Rosencrantz and Guildenstern will end up dead?

The death of both men requires a banana peel, a frayed rope, and at least one executioner's bullet. Here, one case of substitutability (one of the two guns must fire) and one case of conjunctural causation (the banana peel and the frayed rope must be present) are embedded within a larger instance of conjunctural causation (Rosencrantz and Guildenstern must both be killed for the condition of interest to be realized).

The known probabilities mentioned previously yield a simple answer:

$$\left(\frac{1}{3} \times \frac{1}{3}\right) \times \left\{1 - \left[\left(1 - \frac{1}{6}\right) \times \left(1 - \frac{1}{6}\right)\right]\right\} = \frac{1}{9} \times \frac{11}{36} = \frac{11}{324}$$

or about 3.4%.

The procedure for generalizing to a situation in which probabilities are unknown but can be modeled as functions of sets of independent variables is precisely the same in this situation as it was in the previous ones. If we use p_{x_1} to denote the probability that Rosencrantz will choose a banana, p_{x_2} to denote the probability that Rosencrantz will choose a frayed rope, p_{x_3} to denote the probability that the first executioner's gun will fire and p_{x_4} to denote the probability that the second executioner's gun will fire, p_y (the probability that both Rosencrantz and Guildenstern will end up dead) can be calculated as

$$p_y = \left(p_{x_1} \times p_{x_2}\right) \times \left\{1 - \left[\left(1 - p_{x_3}\right) \times \left(1 - p_{x_4}\right)\right]\right\}. \tag{8}$$

Parameterization is substantially more complex but, as before, flows directly from the logic of the example. Using l and m to denote secondary causal paths and additive variables (analogous to j and k, respectively), we have

$$\Pr(y_i = 1 \mid \alpha, \beta, x_i) = \left(\prod_{j=1}^{J}\left[\Phi\left(\alpha_j + \sum_{k=1}^{K}\beta_{jk}x_{ijk}\right)\right]\right) \times \left(1 - \prod_{l=1}^{L}\left[1 - \Phi\left(\alpha_l + \sum_{m=1}^{M}\beta_{lm}x_{ilm}\right)\right]\right). \tag{9}$$

The likelihood function can then be drawn up in precisely the same manner as the previous likelihood functions and maximized to produce parameter estimates.

More examples would be pointless, as the basic principle should now be clear: if a theorist attempts to explain the occurrence of an event in the language of complex causation, the resulting hypotheses should be capable of expression in terms of a basic probability calculus. A likelihood function can then be derived and maximized to produce Boolean probit (or logit) estimates. There is, in principle, no limit to how convoluted either procedure could become, although in practice the so-called curse of dimensionality will place increasingly high demands on the data as the number of dimensions increases.[21]

In the next section I use these insights and techniques to solve a long-standing puzzle in the study of international security: the curious disjuncture between theory and statistical evidence in the deterrence literature that relates capabilities and resolve to conflict.

4 The Puzzle: Capabilities, Resolve, and Conflict

The idea that conflict occurs when at least one party possesses both the capabilities and the willingness to initiate it is one of the most durable intuitions in the field of international relations. Generally, "capabilities" is used to refer to the ability to engage in military conflict with a reasonable chance of success, whereas "willingness" refers to the importance of the issues at stake to the state in question.[22] Schelling (1966) devoted the first two chapters of his seminal *Arms and Influence* to the topics of capabilities ("The Diplomacy of Violence") and willingness ("The Art of Commitment"). The two concepts are built into the literature on expected utility and war: the probability of victory or defeat is directly related to capabilities, positive or negative utility determines whether or not a state is willing to fight, and expected utility is typically expressed as their product (Bueno de Mesquita 1981). Game theorists focus primarily on willingness—and in particular the problem of commitment, or conveying willingness credibly (Morrow 1999)—though both capabilities and willingness permeate the rationalist literature on war. Even George and Smoke (1974), who critique the rationalist literature on deterrence, cannot escape the centrality of willingness and capabilities, which appear as the first and third factors listed (respectively) in their reformulation of deterrence theory (pp. 523–531). At times, terminology hides the pervasiveness of the two concepts; to Most and Starr (1989), for example, capabilities would be a form of "opportunity."

Curiously, however, the quantitative literature has largely avoided explicit tests of this formulation—so much so that, in Geller and Singer's book-length summary of quantitative

[21] Bellman (1961) was, to my knowledge, the first to use this phrase, though it has now become quite common. It refers here to the exponential growth, as n increases, of the number of regions in an n-dimensional dataspace that must contain data if all parameters are to be estimated. Simple crosstabulation provides a clear illustration: 2×2 crosstabs require data in four regions (squares); $2 \times 2 \times 2$ crosstabs require data in eight regions; and so on.

[22] "Willingness" and "resolve" are often used interchangeably; I prefer the former because the latter is also used to refer to the outcome of the decision process to which capabilities and willingness are inputs.

studies of war (Geller and Singer 1998), capabilities and "status quo orientation" (the nearest thing to willingness) occupy entirely separate sections (pp. 68–76 and 89–92, respectively).[23] The few tests that have been conducted often report results that are either inconsistent or at odds with intuition: either one of the two elements plays a disproportionately small role in determining outcomes, or one is found not to matter at all. Balance of capabilities is found to be only partially relevant (immediate balance of forces matters, long-term balance does not) to the question of conflict initiation in both Huth and Russett (1984) and Huth (1988), for example, but are found to be entirely irrelevant in Huth (1996). One might also expect, in cases of disputes, to find that the resulting settlement reflects the capabilities and interests of the actors involved (e.g., Fearon 1995), but this intuition is also unsupported by the quantitative literature: Maoz (1983) examines both capabilities and willingness as predictors of the outcomes of militarized interstate disputes and finds superior capabilities to be unrelated to dispute outcome. Perhaps because a wider variety of indicators are utilized, willingness tends to fare a bit better, but there are still significant discrepancies among studies regarding which indicators make a significant difference (see, e.g., the different findings in Huth and Russett (1984) vs. Huth (1988) regarding the significance of military and economic ties between defender and protégé). Finally, although Bueno de Mesquita (1981) finds that his data support the claim that expected utility is necessary for war initiation "within the bounds of measurement error" (p. 130), Braumoeller and Goertz (2000, pp. 850–852) examine the measurement-error claim and find the conclusion to be problematic.

How can large-N findings be both so inconsistent and so at odds with both qualitative findings and theoretical logic? One possibility is that the theory is misspecified. This possibility is especially strong in studies of crises (or attempts at *immediate* deterrence), as opposed to those that examine the question of how states come to challenge deterrence commitments in the first place (the breakdown of *general* deterrence). The role of capabilities and willingness in the resolution of immediate deterrence situations is contested: while traditional rational deterrence theory suggests that both will be related to conflict initiation, game-theoretic formulations would support either a positive or null relationship between capabilities and escalation and a negative or null relationship between willingness and initiation (Fearon 1994a, b). By contrast, all accounts concur on the role of capabilities and willingness in the breakdown of general deterrence.[24]

These arguments render a valuable service by providing more explicit and detailed logic than had previously existed. Nevertheless, as the literature just cited demonstrates, these expectations are not realized in the small number of studies carried out to date. The positive relationship between capabilities and outcomes in immediate deterrence in Huth and Russett (1984) and Huth (1988) is inconsistent with Fearon (1994a), and the absence of a positive relationship between capabilities and the breakdown of general deterrence in Huth (1996) is inconsistent with any expectations. Even the four indicators of willingness examined in Fearon's reanalysis (1994b) of Huth (1988) provide ambiguous results: two (trade and arms transfers) are positively correlated with the success of immediate deterrence and two (contiguity and alliances) are negatively correlated with the success of immediate

[23] Expected utility theory, which does combine capabilities and willingness, is the exception; it merits two pages of text.

[24] Fearon (1994a, p. 586) argues that "rational states will 'select themselves' into crises on the basis of observable measures of relative capabilities and interests and will do so in a way that neutralizes any subsequent impact of these measures," and Fearon (1994b, pp. 244–247) asserts that capabilities and willingness should be associated with the success or failure of general deterrence in a manner consistent with traditional rational deterrence theory.

deterrence. In no case does the ratio of coefficient to standard error inspire high levels of confidence. In short, the data presented thus far are quite often inconsistent with either traditional or game-theoretic expectations.

A more plausible answer, I believe, is that deterrence theory posits a relationship of complex causation that is not captured by the standard statistical procedures used to test it to date. In all of the studies listed above, variables representing capabilities and resolve are put into a standard additive equation.[25] There are two reasons to find fault with a straightforward additive model. First, if either capabilities or willingness is absent, variations in the other should be irrelevant; only the conjuncture of the two should produce conflict. Second, as Huth (1996) points out, the willingness to engage in conflict can stem from multiple, substitutable sources. The logic of the theory therefore suggests both substitutability and conjunctural causation. The functional forms of the econometric tests used to evaluate it do not reflect either expectation.

4.1 *Resolving the Puzzle*

In attempting to resolve this discrepancy I will utilize data from Huth (1996), for three reasons. First, it is considerably more comprehensive than previous data sets and contains measures of a wider range of variables over a far larger number of cases. Second, it explicitly examines the initiation and maintenance of territorial disputes, as distinct from the escalation of such disputes to war. It should, therefore, be relatively uncontroversial to scholars in the rational deterrence school: traditional deterrence theory and the more recent game-theoretic literature agree on the sources of general deterrence failures (see footnote 24), though the same cannot be said of failures of immediate deterrence. Finally, it is an explicit attempt to combine realist and domestic theories of conflict, and for that reason it can illustrate the utility of the technique for the purpose of theory synthesis.

Huth examines territorial disputes in an attempt to determine why states initiate or maintain territorial claims. Faced with competing answers (realist vs. domestic politics) to this question, Huth incorporates insights from both schools into a larger explanation. In his words, he seeks to “rethink a realist approach by focusing specifically on the impact of domestic politics in the formation of the security policy of states” (p. 16).

In order to do so, Huth divides his independent variables into three categories: issues at stake, international context, and domestic context. He then runs a standard logit analysis with variables from all three categories as independent variables and bases his conclusions on the significance (or lack of it) of the individual variables. Briefly, the findings suggest four things: that the association of certain issues (e.g., strategic value) with a particular piece of territory increases the probability that a state will advance a claim to it; that such a claim is also made more likely in the presence of a host of domestic contextual variables; that factors defining the international context tend to make such claims less likely to occur; and finally, as an exception to the last generalization, that the balance of military forces does *not* exert an effect one way or the other on the probability that the state will forward or maintain a territorial claim. (The analysis is replicated in Table 3.)

Huth’s hypotheses are linear and additive, so the test is appropriate given the hypotheses. At the same time, however, the argument contains subtleties that are not reflected in the hypotheses. First of all, balance of forces is unlikely to play a role similar to that played by the other variables. Systemic or domestic factors provide the willingness to forward a claim,

[25]The exception is Maoz, who bases his findings on both an additive multivariate analysis and separate bivariate analyses. Neither, however, is exempt from the criticism that follows.

Table 3 Initiation or maintenance of claim implying territorial dispute: replication of original analysis

Variable	*Coefficient*	*SE*	*p*
Issues at stake			
Strategic location of territory	2.637	0.147	.000
Ties to bordering minority	−0.073	0.103	.477
Political unification	1.085	0.100	.000
Economic value of territory	0.563	0.083	.000
International context			
Balance of military forces	0.083	0.170	.624
Prior gain of territory	−0.969	0.136	.000
Common alliance	−2.342	0.110	.000
Previous settlement	−3.339	0.115	.000
Domestic context			
Prior unresolved dispute	3.761	0.124	.000
Prior loss of territory	2.231	0.102	.000
Decolonization norm	0.740	0.112	.000
Constant	−2.152	0.117	.000

whereas greater capabilities provide the ability to do so.[26] Huth (1996, p. 39) expresses this distinction directly when he notes that modified realism portrays leaders as

> limiting the threat or use of military power to international disputes where politically salient issues are at stake *and* the costs of military conflict are not substantial (emphasis added).

The fact that claims are likely to be forwarded only when states have both the capabilities and willingness to do so suggests an overarching process of conjunctural causation. If we use p_c to denote the probability that the leadership believes that it is capable of making a difference and p_w to denote the probability that the state is willing to begin a dispute, the first point suggests that the probability of a dispute will be high when leaders both perceive an incentive to act and possess the ability to do so:

$$p_{disp} = p_c \times p_w. \tag{10}$$

Second, willingness to engage in conflict can arise from either domestic or international sources. The heart of Huth's "modified realism" is the insight that states can be prompted either by domestic incentives or by systemic incentives to become involved in a dispute:

> [S]tate leaders must juggle two critical political roles: (1) they are held accountable for preserving the national security of their country, and (2) they are politicians who seek to remain in power and thus are concerned with current or potential political opposition from counterelites (Huth 1996, p. 42).

Therefore, system- and domestic-level incentives should exhibit substitutability: either should, independently of the other, be able to provide the impetus for conflict. If we use p_{di} to denote the probability that domestic audiences will perceive an incentive to initiate

[26]On willingness, e.g., "foreign policy makers are also domestic political leaders who seek to remain in power, which should influence their assessment of what salient issues are at stake in a territorial dispute" (p. 50); on ability, "The stronger the challenger, the higher the probability that the challenger can overturn the territorial status quo" (p. 53).

a dispute and p_{si} to denote the probability that leaders will perceive such an incentive for systemic reasons, the second point suggests that either domestic or systemic factors can provide the willingness to act:

$$p_w = 1 - [(1 - p_{si}) \times (1 - p_{di})]. \tag{11}$$

Substituting the right-hand side of Eq. (11) into Eq. (10) yields the overall calculation:

$$p_{disp} = p_c \times \{1 - [(1 - p_{si}) \times (1 - p_{di})]\}. \tag{12}$$

That is, a dispute is likely to be initiated and maintained if the state is able to effect a change in the status quo *and* the state desires (either for traditional *realpolitik* reasons *or* for the purpose of pleasing domestic audiences) to do so.

This logic in turn suggests a recategorization of the relevant independent variables. The variable most relevant to the state's ability to effect a change in the status quo, obviously, is the balance of military forces. As to the issues at stake, it is difficult to offer a perfectly clear delineation between issues that would prompt a response for systemic reasons and those that would be salient to a domestic audience. Nevertheless, we can start by thinking about which variables should matter most from the point of view of traditional *realpolitik*. Obviously, the strategic location of a particular piece of territory qualifies, as does the existence of a common alliance. Perhaps not as obviously, the existence of a previous settlement qualifies as well because of the state's desire to avoid the reputation costs involved in going back on its word.[27]

On the other hand, the remaining issues seem most likely to be important because of their salience to a domestic audience. Ties to bordering minorities are obviously considerably more important for their domestic implications than for their effect on the strategic calculus of leaders, as are calls for political unification. Prior gains and losses of territory and prior unresolved disputes would be viewed by strict *realpolitik*ers as sunk costs (or benefits); their primary importance would be to disgruntled populations (e.g., Alsace-Lorraine). Finally, whether or not the issue is seen as relevant to the issue of decolonization in light of the substantial post-World War II decolonization norm may make domestic audiences more sympathetic to intervention, but again, its relevance from a purely systemic point of view is negligible.

One issue, the economic value of the territory in question, could reasonably provide incentives for conflict for either systemic or domestic reasons: economic strength is one dimension of a state's overall capabilities, and economically valuable territory is desirable to domestic audiences capable of capitalizing on its resources.[28] Here we see another advantage of Boolean techniques: because independent variables from one dimension are not simply added to independent variables from another dimension, the same variable can be included in more than one dimension without inducing perfect multicollinearity. Therefore, it is possible to ascertain whether a territory's economic value matters to decisionmakers for systemic or domestic reasons (or both) by including the variable in both dimensions. These insights point to the recategorization scheme laid out in Table 4.

Now that the revised theory has been laid out, a likelihood function must be derived from the basic logic of the theory. Two initial difficulties must be addressed. First, utilizing the

[27] Huth makes precisely this argument on p. 59.
[28] The author is explicitly agnostic on this point; see Huth (1996, p. 52).

Table 4 Original categorization of variables (Huth 1996) and recategorization for Boolean logit analysis

Original model	*Boolean logit*	*Symbols*
Issues at stake	Capabilities (p_c)	βX_{1k}
Strategic location of territory	Balance of military forces	X_{11}
Ties to bordering minority		
Political unification	Systemic incentives (p_{si})	βX_{2l}
Economic value of territory	Strategic location of territory	X_{21}
	Economic value of territory	X_{22}
International context	Prior gain of territory	X_{23}
Balance of military forces	Shared alliance	X_{24}
Prior gain of territory	Previous settlement	X_{25}
Common alliance		
Previous settlement	Domestic incentives (p_{di})	βX_{3m}
	Ties to bordering minority	X_{31}
Domestic context	Political unification	X_{32}
Prior unresolved dispute	Economic value of territory	X_{33}
Prior loss of territory	Prior unresolved dispute	X_{34}
Decolonization norm	Prior loss of territory	X_{35}
	Decolonization norm	X_{36}

notation used so far would produce a monstrous function. In the hopes that doing so will aid intuition more than hinder it, I will denote the constant as β_{j0} rather than α_j and introduce an abbreviated form βX_{jk},

$$\beta X_{jk} \equiv \beta_{j0} + \sum_{k=1}^{K} \beta_{jk} x_{ijk}, \tag{13}$$

which should make the overall logic of the likelihood function considerably easier to see.

Second, Huth utilizes logit rather than probit analysis in his hypothesis test because his sampling rate varies depending on the value of the dependent variable and, while logit coefficients remain consistent under those circumstances, probit coefficients do not.[29] Therefore, Boolean logit rather than probit should be used here. The derivation is precisely the same, save that $\Phi(\beta X)$ is replaced by $\frac{e^{\beta X}}{1+e^{\beta X}}$.

Generalizing from the logic in Eq. (12),

$$\begin{array}{ccccccc} p_{\text{disp}} & = & p_{\text{c}} & \times & [1-(1-p_{\text{si}}) & \times & (1-p_{\text{di}})] \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \end{array}$$

$$\Pr(y_i = 1 \mid \beta, x_i) = \frac{e^{\beta X_{1k}}}{1+e^{\beta X_{1k}}} \times \left[1 - \left(1 - \frac{e^{\beta X_{2l}}}{1+e^{\beta X_{2l}}}\right) \times \left(1 - \frac{e^{\beta X_{3m}}}{1+e^{\beta X_{3m}}}\right)\right].$$

[29] See Huth (1996, Appendix D, footnote 2); the interested reader will find Achen (1999) to be a clear discussion of the underlying statistical logic. Here, as in Huth's analysis and the replication, ln(.289) has been subtracted from each of the βX_j prior to estimation to force an adjustment to the constant that will compensate for undersampling of zeros.

Table 5 Boolean logit model—results

Variable	*Coefficient*	*SE*	*p*
Capabilities			
Balance of military forces	1.310	0.490	.007
Constant	1.269	0.163	.000
Systemic incentives			
Strategic location of territory	3.826	0.232	.000
Economic value of territory	0.702	0.115	.000
Prior gain of territory	−1.791	0.164	.000
Common alliance	−1.173	0.120	.000
Previous settlement	−18.87	176.9	.915
Constant	−1.590	0.068	.000
Domestic incentives			
Ties to bordering minority	−0.920	0.209	.000
Political unification	−0.570	0.212	.007
Economic value of territory	−0.022	0.196	.909
Prior unresolved dispute	8.346	0.416	.000
Prior loss of territory	5.039	0.369	.000
Decolonization norm	3.500	0.283	.000
Constant	−7.399	0.374	.000

Accordingly,

$$\Pr(y_i = 0 \mid \beta, x_i) = 1 - \frac{e^{\beta X_{1k}}}{1+e^{\beta X_{1k}}} \times \left[1 - \left(1 - \frac{e^{\beta X_{2l}}}{1+e^{\beta X_{2l}}}\right) \times \left(1 - \frac{e^{\beta X_{3m}}}{1+e^{\beta X_{3m}}}\right)\right].$$

The likelihood function then becomes

$$L(Y \mid \beta, X) = \prod_{i=1}^{N} \left(\frac{e^{\beta X_{1k}}}{1+e^{\beta X_{1k}}} \times \left[1 - \left(1 - \frac{e^{\beta X_{2l}}}{1+e^{\beta X_{2l}}}\right) \times \left(1 - \frac{e^{\beta X_{3m}}}{1+e^{\beta X_{3m}}}\right)\right]\right)^{y_i} \times \left(1 - \frac{e^{\beta X_{1k}}}{1+e^{\beta X_{1k}}} \times \left[1 - \left(1 - \frac{e^{\beta X_{2l}}}{1+e^{\beta X_{2l}}}\right) \times \left(1 - \frac{e^{\beta X_{3m}}}{1+e^{\beta X_{3m}}}\right)\right]\right)^{1-y_i}.$$

Maximizing this likelihood function yields the results given in Table 5.

4.2 *Results*

The results of the Boolean logit analysis call into question some of the main findings from the simple additive model. Balance of forces, originally deemed irrelevant based on the simple logit model, proves to have a highly significant effect on initiation or maintenance of a territorial claim *in conjunction with* either systemic or domestic incentives to make such a claim. At the same time, ties to a bordering minority and the presence of the question of political unification both exert a negative influence on the domestic audience's willingness to press a territorial claim; the first variable was found to be insignificant in the original study and the second was found to exert a significant and *positive* effect on the probability of conflict. The latter finding suggests that domestic audiences are on average more willing to press a claim for territory when they feel no affinity for those from whom it is to be taken (and who might have to be shot if that claim is to be realized).

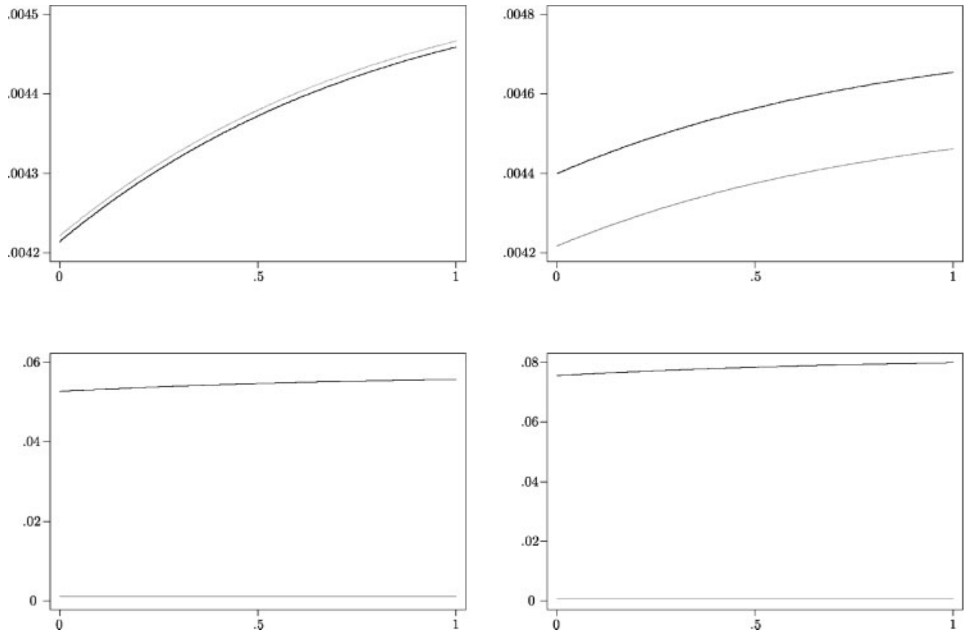

Fig. 3 Predicted probabilities of dispute (y-axis) given balance of forces (x-axis) and changes in (from upper left to lower right) prior gain of territory, strategic location of territory, prior loss of territory, and prior unresolved dispute; changes are from absence of causal factor (lighter lines) to presence (darker lines).

Discussing the substantive impact of changes in independent variables on the dependent variable can be a bit more complex when Boolean probit and logit techniques are used, but it is hardly impossible. Figure 3 provides four examples. Each demonstrates how the probability of initiation or maintenance of a dispute varies as a function of balance of forces when a particular independent variable is either absent or present and the remainder are held at their mean values. Moving from upper-left to lower-right demonstrates that some variables have a considerably larger substantive impact on dispute behavior than do others under these circumstances: while prior gain of territory is virtually irrelevant, the presence of a prior unresolved dispute can increase the probability of a dispute by nearly 8%.

The ability to model the economic value of the territory as both a systemic and domestic incentive has a very clear payoff: while its effects as a systemic incentive are quite significant statistically, the ratio of coefficient to standard error suggests quite strongly that its effects as a domestic incentive are virtually nil. Put more precisely, the results suggest that the economic value of a piece of territory contributes substantially to an additive measure of systemic willingness but contributes little, if anything, to an additive measure of domestic willingness. This sort of conclusion was simply not possible in the original study because of the methodology employed.

The results also highlight the effects of the so-called curse of dimensionality. One variable, prior settlement of dispute, exhibits the classic symptom of underidentification: a grossly inflated standard error. The evidence of underidentification can be found by holding all coefficients but one constant at their values at convergence, varying the coefficient for one of the variables, plotting its value against the model's overall log-likelihood, and looking for plateaus in the likelihood function. Figure 4 demonstrates that such a plateau is not

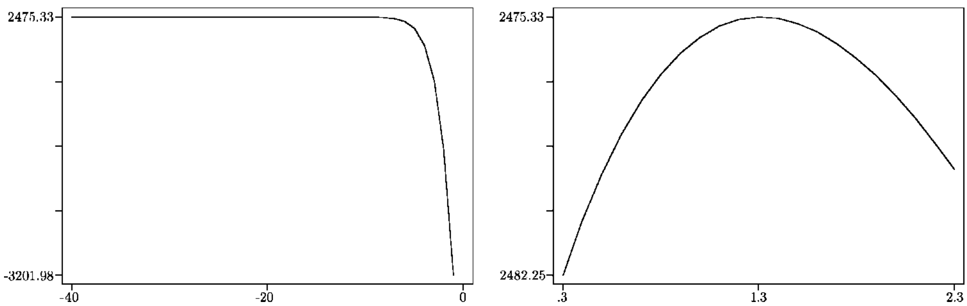

Fig. 4 Log-likelihoods (y-axis) vs. coefficient values (x-axis) for prior settlement of dispute (L) and balance of forces (R).

only present but pronounced for the prior settlement variable. (An example of a variable for which underidentification is not a problem, balance of forces, is provided for comparison.)[30]

This problem occurs here because evaluating theories that posit causal complexity requires what for lack of a better term I will call *complex covariation*: there must be enough cases in the critical regions of a multiply partitioned dataspace to permit inference. In this example, despite the fact that there are 8328 overall observations, the high balance of forces/high domestic willingness category[31] contains too few cases of prior settlement (to be exact, four) to permit reliable inferences to be drawn regarding its effects on conflict initiation.[32] Evaluating multiple causal path theories hinges critically on complex covariation, regardless of sample size. This is a feature of reality, not of the model—it holds regardless of the particular methodology used. Boolean probit and logit techniques correctly alert the user to the fact that the number of cases in a particular category is too low to permit inference.

Finally, one of Huth's main objectives was to combine realism and theories of domestic politics. As Moravcsik (1997, p. 542) writes,

> If foreign policymaking is a process of constrained choice by purposive states, a view shared by realist, institutionalist, and liberal theory, there may well be cases in which a *combination* of preferences and constraints shapes state behavior. In such cases, a multicausal synthesis, one that treats these theories not as substitutes but as complements, is required. (emphasis in original)

One of the main theoretical advantages of modeling complex causation is that it facilitates this sort of synthesis. Implicit in the original additive formulation is the notion that a unit decrease in x_1 can be counteracted by a $\frac{\beta_1}{\beta_2}$-unit increase in x_2: any two variables are functionally interchangeable. In the revised formulation, no such assumption is made across variables derived from different theories. The nature of the variables' interaction is clearly—and theoretically—specified. The level of constraint is determined by relative capabilities; preference stems either from traditional *realpolitik* goals or from domestic audience incentives; and only the combination (conjuncture) of constraint and preference determines outcomes.

---

[30] See King (1989, pp. 192–193) for a discussion of plateaus in likelihood functions. Note that the only real "fix" for this problem is to gather more data.

[31] That is, balance of forces $> \frac{2}{3}$ and $\frac{e^{\beta X_{3m}}}{1+e^{\beta X_{3m}}} > 0.5$.

[32] All variables except balance of forces are binary—a fact that aggravates identification problems considerably.

5 Conclusion

Causal complexity is an ideal problem for statistical methodology to tackle because of its broad applicability and utility. The solution provided here is designed specifically to deal with the peculiar empirical implications of complexity, and as such it adds a prominent new weapon to our collective methodological arsenal.

The details of the procedure itself, however, should not obscure the broader message: the full potential of quantitative methodology—or of any methodology, for that matter—in the study of politics can only be realized if theoretical logic rather than convention drives the selection, and if necessary the creation, of appropriate methods. Separatism, or even antagonism, among methodological subfields ill-serves this goal, especially when genuine accommodation is possible. Indeed, by combining their contributions we may find that the whole is greater than the sum of the parts.

References

Abowd, J. M., and H. S. Farber. 1982. "Job Queues and the Union Status of Workers." *Industrial and Labor Relations Review* 35:354–367.

Achen, C. H. 1999. "Retrospective Sampling in International Relations." Presented at the Annual Meeting of the Midwestern Political Science Association, Chicago, IL.

Achen, C. H. 2002. "Toward a New Political Methodology: Microfoundations and ART." *Annual Review of Political Science* 5:423–450.

Aldrich, J. H., and F. D. Nelson. 1984. *Linear Probability, Logit, and Probit Models* (Sage University Paper Series on Qualitative Applications in the Social Sciences, series no. 07–045). Beverly Hills, CA: Sage.

Alker, Jr., H. R. 1966. "The Long Road to International Relations Theory: Problems of Statistical Nonadditivity." *World Politics* 18:623–655.

Banaszak, L. A. 1996. *Why Movements Succeed or Fail: Opportunity, Culture, and the Struggle for Woman Suffrage* (Princeton Studies in American Politics: Historical, International, and Comparative Perspectives). Princeton, NJ: Princeton University Press.

Bartolini, S., and P. Mair. 1990. *Identity, Competition, and Electoral Availability: The Stabilization of European Electorates 1885–1985*. Cambridge: Cambridge University Press.

Bellman, R. E. 1961. *Adaptive Control Processes: A Guided Tour.* Princeton, NJ: Princeton University Press.

Bennett, A., and A. George. 1997. "Research Design Tasks in Case Study Methods." In *MacArthur Foundation Workshop on Case Study Methods*. Belfer Center for Science and International Affairs (BCSIA), Harvard University.

Blainey, G. 1973. *The Causes of War*. New York: Free Press.

Braumoeller, B., and G. Goertz. 2000. "The Methodology of Necessary Conditions." *American Journal of Political Science* 44:844–858.

Bueno de Mesquita, B. 1981. *The War Trap.* New Haven, CT: Yale University Press.

Bueno de Mesquita, B. 1985. "The War Trap Revisited: A Revised Expected Utility Model." *American Political Science Review* 79:156–177.

Cioffi-Revilla, C., and H. Starr. 1995. "Opportunity, Willingness and Political Uncertainty: Theoretical Foundations of Politics." *Journal of Theoretical Politics* 7:447–476.

Colomer, J. M. 1991. "Transitions by Agreement: Modeling the Spanish Way." *American Political Science Review* 85:1283–1302.

Converse, P. 1964. The Nature of Belief Systems in Mass Publics. In *Ideology and Discontent*, ed. D. Apter. New York: Free Press, pp. 206–261.

Cowhey, P. F. 1990. "The International Telecommunications Regime: The Political Roots of Regimes for High Technology." *International Organization* 44:169–199.

Dassel, K., and E. Reinhardt. 1999. "Domestic Strife and the Initiation of Violence at Home and Abroad." *American Journal of Political Science* 43:56–85.

Dion, D. 1998. "Evidence and Inference in the Comparative Case Study." *Comparative Politics* 30:127–145.

Doyle, M. W. 1983a. "Kant, Liberal Legacies, and Foreign Affairs." *Philosophy and Public Affairs* 12:205–235.

Doyle, M. W. 1983b. "Kant, Liberal Legacies, and Foreign Affairs, Part 2." *Philosophy and Public Affairs* 12:323–353.

Evangelista, M. 1995. “The Paradox of State Strength: Transnational Relations, Domestic Structures, and Security Policy in Russia and the Soviet Union.” *International Organization* 49:1–38.

Fearon, J. D. 1994a. “Domestic Political Audiences and the Escalation of International Disputes.” *American Political Science Review* 88:577–592.

Fearon, J. D. 1994b. “Signaling versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model.” *Journal of Conflict Resolution* 38:236–269.

Fearon, J. D. 1995. “Rationalist Explanations for War.” *International Organization* 49:379–414.

Geller, D. S., and J. D. Singer. 1998. *Nations at War: A Scientific Study of International Conflict.* Cambridge: Cambridge University Press.

George, A. L., and R. Smoke. 1974. *Deterrence in American Foreign Policy: Theory and Practice.* New York: Columbia University Press.

Goertz, G. 1994. *Contexts of International Politics.* Cambridge: Cambridge University Press.

Goldmark, J. C. 1904. “Street Labor and Juvenile Delinquency.” *Political Science Quarterly* 19:417–438.

Good, I. J. 1985. “Causal Propensity: A Review.” In *PSA 1984,* Vol. 2, eds. P. D. Asquith and P. Kitcher. East Lansing, MI: Philosophy of Science Association, pp. 829–850.

Greene, W. H. 2000. *Econometric Analysis,* 4th ed. Upper Saddle River, NJ: Prentice Hall.

Hirlinger, M. W. 1992. “Citizen-Initiated Contacting of Local Government Officials: A Multivariate Explanation.” *Journal of Politics* 54:553–564.

Hirschman, A. O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States.* Cambridge, MA: Harvard University Press.

Hobbes, T. 1651 (1997). *Leviathan,* New York: W. W. Norton.

Hoffmann, S. H. 1959. “International Relations: The Long Road to Theory.” *World Politics* 11:346–377.

Huntington, S. P. 1958. “Arms Races: Prerequisites and Results.” *Public Policy* 8:41–83.

Huntington, S. P. 1991. *The Third Wave: Democratization in the Late Twentieth Century.* Norman: University of Oklahoma Press.

Huth, P. 1996. *Standing Your Ground: Territorial Disputes and International Conflict.* Ann Arbor: University of Michigan Press.

Huth, P., and B. Russett. 1984. “What Makes Deterrence Work? Cases from 1900–1980.” *World Politics* 36:496–526.

Huth, P. K. 1988. *Extended Deterrence and the Prevention of War.* New Haven, CT: Yale University Press.

Jennings, M. K., and E. A. Andersen. 1996. “Support for Confrontational Tactics among AIDS Activists: A Study of Intra-Movement Divisions.” *American Journal of Political Science* 40:311–334.

Jervis, R. 1978. “Cooperation Under the Security Dilemma.” *World Politics* 30:167–214.

Jervis, R. 1997. *System Effects: Complexity in Political and Social Life.* Princeton, NJ: Princeton University Press.

Kinder, D. R., and T. Mendelberg. 1995. “Cracks in American Apartheid: The Political Impact of Prejudice Among Desegregated Whites.” *Journal of Politics* 57:402–424.

King, G. 1989. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference.* Cambridge: Cambridge University Press.

Kiser, E., and M. Levi. 1996. Using Counterfactuals in Historical Analysis: Theories of Revolution. In *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives,* eds. P. E. Tetlock and A. Belkin. Princeton, NJ: Princeton University Press, pp. 187–207.

Krutz, G. S., R. Flesher, and J. R. Bond. 1998. “From Abe Fortas to Zoe Baird: Why Some Presidential Nominations Fail in the Senate.” *American Political Science Review* 92:871–881.

Levy, J. 1983. “Misperception and the Causes of War: Theoretical Linkages and Analytical Problems.” *World Politics* 36:76–99.

Lijphart, A. 1963. “The Analysis of Bloc Voting in the General Assembly: A Critique and a Proposal.” *American Political Science Review* 57:902–917.

Lupia, A., and K. Strom. 1995. “Coalition Termination and the Strategic Timing of Parliamentary Elections.” *American Political Science Review* 89:648–665.

Mackie, J. L. 1965. “Causes and Conditions.” *American Philosophical Quarterly* 2:245–264.

Maddala, G. S. 1994. *Limited-Dependent and Qualitative Variables in Economics* (Econometric Society Monographs). Cambridge: Cambridge University Press.

Maltzman, F., and P. J. Wahlbeck. 1996. “Strategic Policy Considerations and Voting Fluidity on the Burger Court.” *American Political Science Review* 90:581–592.

Manton, K. G., E. Stallard, and H. D. Tolley. 1991. “Limits to Human Life Expectancy: Evidence, Prospects, and Implications.” *Population and Development Review* 17:603–637.

Manton, K. G., H. D. Tolley, and S. S. Poss. 1976. “Life Table Techniques for Multiple-Cause Mortality.” *Demography* 13:541–564.

Maoz, Z. 1983. “Resolve, Capabilities, and the Outcomes of Interstate Disputes, 1816–1976.” *Journal of Conflict Resolution* 27:195–229.

McDermott, M. 1997. “Voting Cues in Low-Information Elections: Candidate Gender as a Social Information Variable in Contemporary United States Elections.” *American Journal of Political Science* 41:270–283.

McKenzie, R. T., and A. Silver. 1967. The Delicate Experiment: Industrialism, Conservatism, and Working-Class Tories in England. In *Party Systems and Voter Alignments: Cross-National Perspectives,* eds. S. M. Lipset and S. Rokkan. New York: The Free Press, pp. 115–125.

Moore, Jr., B. 1966. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World.* Boston: Beacon Press.

Moravcsik, A. 1997. “Taking Preferences Seriously: A Liberal Theory of International Politics.” *International Organization* 51:513–553.

Morrow, J. 1999. The Strategic Setting of Choices: Signaling, Commitment, and Negotiation in International Politics. In *Strategic Choice and International Relations,* eds. D. A. Lake and R. Powell. Princeton, NJ: Princeton University Press, pp. 77–114.

Morrow, J. D. 1991. “Alliances and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances.” *American Journal of Political Science* 35:904–933.

Most, B., and H. Starr. 1989. *Inquiry, Logic, and International Politics.* Columbia: University of South Carolina Press.

Most, B. A., and H. Starr. 1984. “International Relations Theory, Foreign Policy Substitutability, and ‘Nice’ Laws.” *World Politics* 36:383–406.

Mueller, J. 1988. “The Essential Irrelevance of Nuclear Weapons: Stability in the Postwar World.” *International Security* 13:45–69.

Nagler, J. 1994. “Scobit: An Alternative Estimator to Logit and Probit.” *American Journal of Political Science* 38:230–255.

Organski, A. F. K., and J. Kugler. 1980. *The War Ledger.* Chicago: University of Chicago Press.

Peterson, M. A. 1990. *Legislating Together: The White House and Capitol Hill from Eisenhower to Reagan.* Cambridge, MA: Harvard University Press.

Poirier, D. J. 1980. “Partial Observability in Bivariate Probit Models.” *Journal of Econometrics* 12:209–217.

Posen, B. 1984. *The Sources of Military Doctrine:France, Britain and Germany Between the World Wars.* Ithaca, NY: Cornell University Press.

Prentice, R. L. 1976. “A Generalization of the Probit and Logit Methods for Dose Response Curves.” *Biometrics* 32:761–768.

Przeworski, A., and J. R. Vreeland. 2002. “A Statistical Model of Bilateral Cooperation.” *Political Analysis* 10:101–112.

Ragin, C. C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies.* Berkeley, University of California Press.

Ragsdale, L., and J. G. Rusk. 1993. “Who Are Nonvoters? Profiles from the 1990 Senate Elections.” *American Journal of Political Science* 37:721–746.

Reed, W. 2000. “A Unified Statistical Model of Conflict Onset and Escalation.” *American Journal of Political Science* 44:84–93.

Riker, W. H., and I. Sened. 1991. “A Political Theory of the Origins of Property Rights: Airport Slots.” *American Journal of Political Science* 35:951–969.

Salmon, W. 1998. *Causality and Explanation.* Oxford: Oxford University Press.

Schelling, T. C. 1966. *Arms and Influence.* New Haven, CT: Yale University Press.

Schmidt, P. 1984. “Limited-Dependent and Qualitative Variables in Econometrics (Review).” *Journal of the American Statistical Association* 79:738.

Shain, Y., and J. J. Linz, eds. 1995. *Between States: Interim Governments and Democratic Transitions* (Cambridge Studies in Comparative Politics). Cambridge: Cambridge University Press.

Sikkink, K. 1993. “Human Rights, Principled Issue-Networks, and Sovereignty in Latin America.” *International Organization* 47:411–441.

Skocpol, T. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China.* Cambridge: Cambridge University Press.

Smith, A. 1999. “Testing Theories of Strategic Choice: The Example of Crisis Escalation.” *American Journal of Political Science* 43:1254–1283.

Sniderman, P. M., E. G. Carmines, G. C. Layman, and M. Carter. 1996. “Beyond Race: Social Justice as a Race Neutral Ideal.” *American Journal of Political Science* 40:33–55.

Stein, J. G. 1987. “Extended Deterrence in the Middle East: American Strategy Reconsidered.” *World Politics* 39:326–352.

Tetlock, P. E., and A. Belkin, eds. 1996. *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives.* Princeton, NJ: Princeton University Press.

Venn, J. 1866. *The Logic of Chance.* London: Macmillan.

Western, B. 1998. "Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Approach." *American Journal of Political Science* 42:1233–1259.

Zaller, J. R. 1992. *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.