Advance Access publication June 7, 2006 *Political Analysis* (2006) 14:311–331
doi:10.1093/pan/mpj018

A Simple Multivariate Test for Asymmetric Hypotheses

William Roberts Clark
Department of Political Science and the Center for Political Studies,
University of Michigan, ISR 4202 Box 1248,
426 Thompson Street, Ann Arbor, MI 48106-1248
e-mail: wrclark@umich.edu (corresponding author)

Michael J. Gilligan
Department of Politics, New York University,
7th Floor, 726 Broadway, New York, NY 10003
e-mail: michael.gilligan@nyu.edu

Matt Golder
Department of Political Science, Florida State University,
531 Bellamy Building, Tallahassee, FL 32306-2230
e-mail: mgolder@fsu.edu

In this paper, we argue that claims of necessity and sufficiency involve a type of asymmetric causal claim that is useful in many social scientific contexts. Contrary to some qualitative researchers, we maintain that there is nothing about such asymmetries that should lead scholars to depart from standard social science practice. We take as given that deterministic and monocausal tests are inappropriate in the social world and demonstrate that standard multiplicative interaction models are up to the task of handling asymmetric causal claims in a multivariate, probabilistic manner. We illustrate our argument with examples from the empirical literature linking electoral institutions and party system size.

1 Introduction

Political scientists often use notions of "necessity" and "sufficiency" in theorizing about politics. For example, the democratic peace conjecture states that joint democracy is a sufficient, but not a necessary, cause for peace within a dyad of countries. When neo-realists claim that anarchy "causes" war, they can only mean that anarchy is a necessary, but not sufficient, condition for inter-state war given that we observe many cases of "not war" under conditions of anarchy. Przeworski et al. (2000) claim that a level of gross domestic product (GDP) per capita equal to 6055 in 1985 PPP U.S. dollars is sufficient, but not necessary, for democratic consolidation. Mainwaring (1993) argues that democracy in

Authors' note: We would like to thank the participants at the University of Pittsburgh Methodology Seminar, Ernest Sergenti, and three anonymous reviewers for their useful comments. The data and computer code necessary to replicate the results and figures in this analysis will be made publicly available at http://www.fsu.edu/~polisci/people/faculty/mgolder.htm on publication. *Stata* 9 was the statistical package used in this study.

© The Author 2006. Published by Oxford University Press on behalf of the Society for Political Methodology.
All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org

presidential systems is more fragile when there are many parties than in parliamentary systems. In other words, having many parties is a necessary, but not sufficient, condition for increased democratic instability in presidential systems. A final example, which we will explore in greater detail below, is the claim, typically associated with Duverger ([1954] 1963), that small district magnitudes are sufficient to create two-party systems and that proportional representation (PR) is necessary for multiparty systems. The list could go on (Goertz 2002). For reasons that we will discuss below, we refer to these *necessary-like* and *sufficient-like* hypotheses as "asymmetric" hypotheses.

Many qualitative researchers have argued that asymmetric hypotheses cannot be evaluated with standard statistical methodologies (Ragin 1987, 2000, 2006; George and Bennett 2005; Mahoney and Goertz 2006). As a result, several scholars have recently proposed new techniques specifically designed for testing asymmetric hypotheses. Although these new techniques are welcome, we believe that they suffer from at least one of two possible shortcomings. On the one hand, some of the techniques require that the social process under examination be deterministic, such that a disconfirming case that arises due to purely stochastic factors would inappropriately lead to a rejection of our hypothesis. On the other hand, other tests, which do allow for stochastic variation in outcomes, are bivariate in nature and, as such, require that the process under examination be monocausal or, if not monocausal, that all other causes be uncorrelated with the causal variable in question (Braumoeller and Goertz 2000).

We argue here that standard linear models that include interaction terms offer a better way to test asymmetric hypotheses. We illustrate this by reexamining the much-studied empirical relationship between electoral institutions and party system size (Duverger [1954] 1963; Amorim Neto and Cox 1997; Clark and Golder 2006). Thus, our main argument, in direct contradiction to much of the qualitative methods literature, is that standard empirical techniques (like regression) are well suited to testing asymmetric claims so long as researchers are sufficiently sensitive to model specification and interpretation. Model specifications that do not take into account the asymmetric nature of their hypotheses may well lead researchers to inappropriately reject such hypotheses.

In the next section, we describe three binary characteristics of causal claims in order to first illustrate the difference between symmetric and asymmetric causal claims and to second point out the shortcomings of deterministic approaches to testing asymmetric causal claims. In Section 3, we briefly review existing probabilistic approaches to testing asymmetric causal claims and argue that they are effectively bivariate in nature; as such, they are inappropriate for multicausal settings. In the next two sections, we outline our technique for testing asymmetric causal claims through the use of multiplicative interaction models and demonstrate its usefulness with an application to Duverger's theory. Finally, we discuss how our technique relates to other familiar approaches to causality in Section 6.

2 Three Characteristics of Causal Claims

Approaches to testing causal claims can be thought of as differing along three binary characteristics: (1) determinism/probabilism, (2) monocausality/multicausality, and (3) symmetry/asymmetry. A factor X is a *deterministic* cause of Y if its presence assures the presence of Y.[1] In terms of continuous variables, a cause is deterministic if a given change in X produces a certain change in Y. In contrast, a causal effect is *probabilistic* if its

[1]Although this sentence uses the example of two variables that are positively related, there is no loss of generality since we can define X however we like.

presence changes the likelihood of, but does not assure, the presence of Y. In terms of continuous variables, a causal effect is probabilistic if a given change in X is expected to produce a certain change in Y plus or minus some error term. X is *monocausal* if it is the only cause that contributes to the occurrence (or, in the continuous case, the magnitude) of some effect Y. A hypothesis is multicausal if there are multiple factors that contribute to the occurrence (or magnitude) of Y.

We call a causal effect *symmetric* if the same change in the causally produced phenomenon is expected both when the cause is added and when it is taken away. For example, the claim that X is a symmetric cause of Y means that (1) if X is present, Y is more likely to occur (or in the continuous case, more of X causes more of Y) and (2) if X is absent, Y is less likely to occur (or in the continuous case, less of X causes less of Y).[2] As an example, we might say that warm weather in the spring "causes" the daffodils to bloom. The presence of the warm weather makes it more likely that the daffodils will bloom, and the absence of the warm weather makes it less likely they will do so. Similarly, we might say that the force applied to a BB in an air rifle causes the acceleration of the BB—more force causes more acceleration, less force causes less acceleration.[3]

Lieberson (1987, 174–8) refers to the situation where the change in the causally produced phenomenon is not of the same order of magnitude or direction when the cause is added as when it is taken away as an *asymmetric* cause. In the discrete case, the claim that X is an asymmetric cause of Y either means (i) that if X is present, Y will certainly occur, but if X is absent, Y may or may not occur or (ii) if X is present, Y may or may not occur, but if X is absent, Y will certainly not occur.[4] Although the language of "asymmetric" causes may be uncommon, these two possible meanings should look very familiar. After all, the statement in (i) is an example of a "sufficient" condition, whereas the statement in (ii) is an example of a "necessary" condition. An example of an asymmetric cause, in this case a necessary-like one, is oxygen as a cause for fire. Oxygen is a necessary condition for fire—take oxygen away and it guarantees that there will be no fire; but add oxygen to a given situation and, depending on other circumstances, there may or may not be fire. Although we shall use the phrase "asymmetric cause" in a slightly broader sense than Lieberson, his notion of an "asymmetric cause" is contained within our current concept. The boundaries of this broader concept will, hopefully, become clear as we proceed.

To put things more simply, "symmetry" refers to necessary and sufficient causes, whereas "asymmetry" refers to causes that are either necessary or sufficient but not both. Given that the language of necessary and sufficient conditions is well known and accepted in the literature, why do we introduce the terms symmetry and asymmetry at all? Our reason for doing this is that we will not actually be using the terms necessity and sufficiency in the standard way that they are used in formal logic. Specifically, formal logic typically uses these words in a deterministic fashion. In other words, if X is necessary for Y, then we will not observe Y in the absence of X—period. However, determinism is not particularly useful for testing theories in a probabilistic (i.e., the real) world. In what

[2]Note that the word "likely" in this sentence indicates that we are talking about a probabilistic symmetric causal claim. This example illustrates how fundamental these three characteristics—determinism/probabilism, monocausality/multicausality, and symmetry/asymmetry—are, in the sense that it is extremely difficult to write a sentence about one of them without having to refer to at least one of the others.

[3]These examples illustrate that the distinction between "discrete" and "continuous," which we have made for clarity, is actually superfluous because we can just think of changes in continuous X and Y as the relevant discrete quantities. If X is "more force" and Y is "more acceleration," then the statement "X causes Y" carries the precise meaning without requiring any further modification to accommodate the continuity of the variables.

[4]The point in footnote 2 applies here as well. In this case, the word "certainly" implies a deterministic asymmetric causal claim.

follows, we will be trying to export necessity-like and sufficiency-like ideas to a probabilistic setting. To avoid confusion with the more standard use of the terms "necessary" and "sufficient," we have adopted the term "asymmetric," which subsumes both of these concepts as limiting cases. Indeed, some scholars feel strongly that it is inappropriate to use the words necessity and sufficiency in a probabilistic setting for the reasons just mentioned. In introducing these new terms, we hope to avoid this semantic debate and focus, instead, on developing strategies for testing theoretical claims.

Table 1 illustrates the eight possible combinations of the three binary characteristics of causal claims and, where possible, identifies some representative scholarly works that advocate the use of each combination.

As we have already stated, the main purpose of our article is to discuss how best to test asymmetric hypotheses in a probabilistic world. However, we think that it would be informative to very briefly examine symmetric and asymmetric hypotheses in a deterministic world before we do this. The claim that X is necessary and sufficient for Y is a deterministic, symmetric, and monocausal claim (column 2 in Table 1). As illustrated by the Venn diagram in Fig. 1a, it is a claim that both $X \backslash Y = \emptyset$ and $Y \backslash X = \emptyset$. The claim that X is necessary (in the standard use of that term) for Y is a deterministic, asymmetric, and monocausal claim (column 8 in Table 1). As illustrated in Fig. 1b, a claim of necessity requires that $Y \backslash X = \emptyset$ but places no restrictions on $X \backslash Y$. The claim that X is sufficient (in the standard use of that term) for Y is also a deterministic, asymmetric, and monocausal claim. As illustrated in Fig. 1c, a claim of sufficiency requires that $X \backslash Y = \emptyset$ but places no restrictions on $Y \backslash X$.[5]

Note that it is possible for a single observation to cast doubt on both symmetric and asymmetric monocausal claims when they are deterministic.[6] In the case of symmetric claims, Fig. 2a illustrates that the occurrence of Y in the absence of X or the nonoccurrence of Y in the presence of X both demonstrate that X is not a necessary and sufficient condition for Y. In the case of asymmetric claims, Fig. 2b illustrates (i) that an occurrence of Y in the absence of X demonstrates that X is not necessary for Y and (ii) that the nonoccurrence of Y in the presence of X demonstrates that X is not sufficient for Y. It is worth pointing out that although there are two sets of possible observations that would falsify a symmetric causal claim (either an occurrence of $X \backslash Y$ or an occurrence of $Y \backslash X$), claims of necessity or sufficiency rule out only one set of observations. In this sense, asymmetric claims are more difficult to falsify—that is, they have less empirical content (Popper 1959). For example, the claim that X is necessary for Y rules out the possibility of observing not X and Y but is mute about how cases are distributed in the X column (see the left-hand side of Fig. 2b). Similarly, a claim of sufficiency is mute about how cases are distributed in the $\sim X$ column (see the right-hand side of Fig. 2b).

Our contention, which we regard as uncontroversial, is that the types of deterministic claims, whether symmetric or asymmetric, we have just discussed are inappropriate in the social science context given that a single occurrence in the relevant set would be sufficient to falsify a claim even if that occurrence was due to random error. A more appropriate set of techniques would take into account the inherently probabilistic nature of social science data. It is to such techniques that we now turn.

[5]The statement that X is necessary for Y amounts to claiming that cases of Y are a subset of the cases of X, i.e., $Y \subset X$, and the statement that X is sufficient for Y amounts to claiming that cases of X are a subset of the cases of Y, i.e., $X \subset Y$.

[6]Although this fact seems to be well received by some champions of small-N research, it epitomizes a dogmatic approach to falsificationism (Lakatos 1970). As a result, there are good reasons to resist making such claims.

Table 1 Three characteristics of causal claims and their possible combinations

	King, Keohane, and Verba	*Mill's method of constant conjunction*	*?*	*Bivariate regression*	*Present paper*	*Braumoeller and Goertz*	*Ragin's "comparative method"*	*Mill's methods of "agreement and difference"*
Symmetry	Y	Y	Y	Y	N	N	N	N
Probabilism	Y	N	N	Y	Y	Y	N	N
Multicausality	Y	N	Y	N	Y	N	Y	N

Note. "?" indicates that we know of no scholarly work that takes this position; "Y" indicates "Yes"; "N" indicates "No."

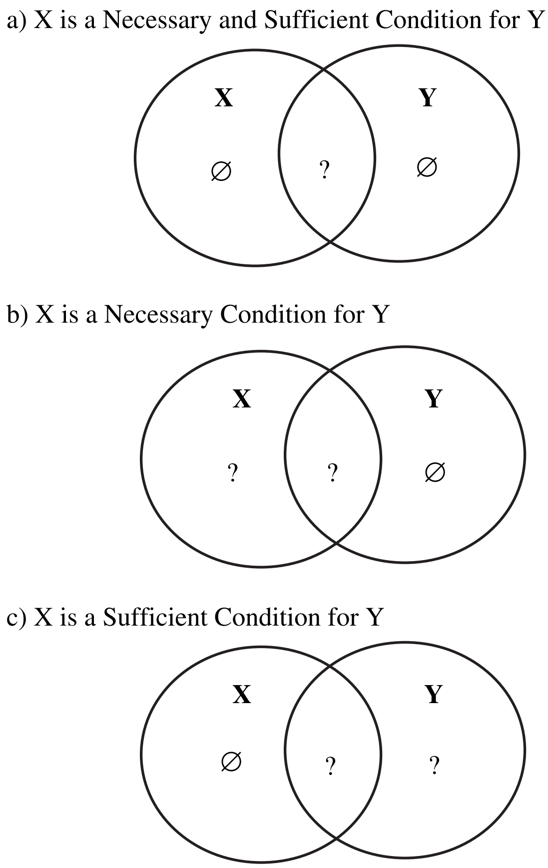

Fig. 1 Venn diagrams of various necessary and sufficient claims. "?" indicates possible observations that are consistent with the stated claims. Ø is the empty set.

3 Probabilistic Tests for Asymmetric Hypotheses

Consider the claim that multimember electoral districts are necessary to produce a multiparty system (a slightly weaker claim is made by Clark and Golder [2006]).[7] If this claim is true, then one would predict that the upper left cell of Table 2 would have zero observations and that the bottom right cell would have a nonzero number of cases. Now compare this set of predictions with the actual data on electoral and party systems from Amorim Neto and Cox (1997) in Table 2.[8] Although a single case in the upper left cell is enough to falsify a (deterministic) claim of necessity, we may, nonetheless, reasonably hold onto the claim that electoral systems *asymmetrically* cause party systems. This seems appropriate since Table 2 indicates that multimember districts are "nearly necessary" for multiparty

[7] In what follows, we illustrate our argument using only examples of necessary-like asymmetric causal statements. However, we should point out that the causal claim "*X* is necessary for *Y*" is equivalent to the claim that "not *X* is sufficient for not *Y*." In general, any statement of a necessary-like asymmetric cause can easily be translated into an equivalent statement of a sufficient-like asymmetric cause. As a result, we can dispense with a separate and redundant discussion of sufficient-like asymmetric causal claims without losing any generality.

[8] We have dichotomized the continuous measures of party system size and electoral system permissiveness of Amorim Neto and Cox for the time being to make this example closer to the discussion at hand. However, we will generalize to continuous variables in a moment.

a) When we have a deterministic, *symmetric*, monocausal claim

Necessary and Sufficient Condition

	~X	X
Y	Disconfirming	
~Y		Disconfirming

b) When we have a deterministic, *asymmetric*, monocausal claim

Claim of Necessity

	~X	X
Y	Disconfirming	
~Y		

Claim of Sufficiency

	~X	X
Y		
~Y		Disconfirming

Fig. 2 Disconfirming observations.

systems. In other words, although a single observation in the upper left cell would be sufficient to demonstrate that *X* is not a necessary condition for *Y*, we would not want to immediately dismiss the possibility that *X* asymmetrically causes *Y* as a result of a small number of disconfirming observations. Recall that this is precisely why we have framed our debate in terms of asymmetry rather than necessary and sufficient conditions.

What causes apparently disconfirming observations? Disconfirming observations can be thought of as being generated by stochastic, systematic, or a combination of stochastic and systematic processes. Based on their intuition about which of these three processes produces the apparent anomalies, scholars have put forward different alternatives to the asymmetric, deterministic, and monocausal approaches to science outlined earlier. For the case where the disconfirming observations are thought to be generated by a purely stochastic process, researchers have developed statistical tests that allow a researcher to gauge the extent to which a relationship between two variables is consistent with a claim of necessity-like asymmetry (Hildebrand, Laing, and Rosenthal 1977; Braumoeller and Goertz 2000). The basic idea behind each of these approaches is that, unlike standard bivariate tests for positive and symmetrical relationships among categorical variables, observations in the bottom right cell are not treated as anomalies as they would be in a test of the symmetrical claim that single-member districts are necessary and sufficient to produce a two-party system (see Fig. 2a again). A benefit of this approach is that one

Table 2 Multimember electoral districts and multipartism

	Single member	*Multimember*
a) Predictions if multimember districts are a necessary cause of multipartism		
Multiparty	0	?
Two party	?	$n > 0$
b) Are multimember districts an asymmetric cause of multipartism?		
Multiparty	3	17
Two party	17	17

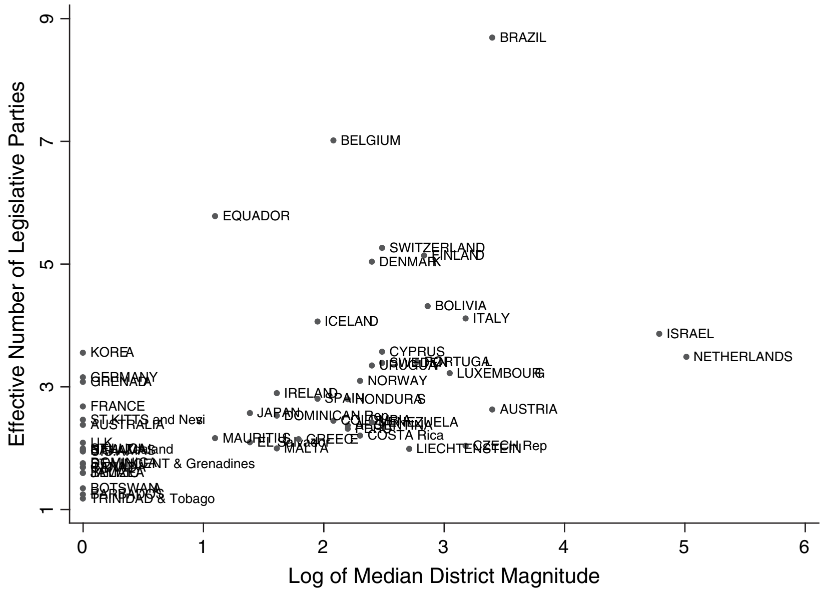

Fig. 3 The relationship between electoral laws and party system size.

no longer needs to dismiss an asymmetric causal claim because of a relatively small set of observations in the cell that should be empty (upper left) according to a claim of necessity. Although these statistical tests are clearly useful, one key weakness is that they can only accommodate categorical variables.

In contrast, our argument about necessity-like asymmetric relationships can easily be extended to handle continuous variables. The scatter plot of two continuous variables in a necessity-like asymmetric relationship would be "lower triangular." That is, the northwest corner of the matrix should be less densely populated than the rest of the graph. As an illustration of this, let us return to the question of whether electoral systems asymmetrically influence the number of parties. However, instead of using dichotomous variables of party system size and electoral system permissiveness, we now employ continuous measures of both variables.[9] The hypothesis is now that a sufficiently permissive electoral system is necessary for a multiparty system.[10] Figure 3 plots the most common measure of electoral system permissiveness—logged district magnitude—on the x axis and the effective number of legislative parties on the y axis.

As expected, cases in the northwest corner of Fig. 3 are relatively scarce. This is what we mean when we say that electoral systems with large district magnitudes are (probabilistically speaking) a necessary, but not sufficient, cause of a large party system. We will not observe a large number of parties unless there is a large district magnitude; however, a large district magnitude may or may not result in a large number of parties. This last point reveals an important aspect of asymmetric hypotheses—they imply conditional

[9] This is beneficial from the viewpoint that dichotomizing the data both "threw away" information and required us to choose an arbitrary cut point in our variables.

[10] We had previously dichotomized electoral system permissiveness into those systems characterized by multimember districts (permissive) and those characterized by single-member districts (nonpermissive). We now use the logged number of representatives elected from each district (district magnitude) as our continuous measure of electoral system permissiveness.

heteroskedasticity. In other words, there is more unexplained variance in the number of parties in those countries with permissive electoral systems than in those countries with restrictive ones. The dispersion (measured by the standard deviation around the mean) of the effective number of legislative parties in single-member district electoral systems is less than a quarter of that in multimember district ones.

If the phenomenon of interest (in this case, the number of parties) really is caused by a single explanatory factor (in this case, district magnitude), then this heteroskedasticity is unexplainable—it must be due to purely stochastic factors. This means that if we are certain that the causal process at work is monocausal, then a test for the presence of heteroskedasticity is equivalent to a test of our claim that the causal factor is necessary but not sufficient.[11] The greater dispersion of the number of parties in large district magnitude electoral systems is only possible because the necessary condition for many parties—a large district magnitude—is present. Countries with small district magnitudes do not permit large numbers of parties and, as a result, the dispersion around those data is small. As we noted earlier, another way to see this is to recognize that the statement "a large district magnitude is *necessary* for a high effective number of parties" is equivalent to the statement that "a small district magnitude is *sufficient* for a low number of parties." In short, a test for heteroskedasticity can serve as a test for an asymmetric causal relationship in the monocausal, continuous setting just as the methods proposed by Hildebrand, Laing, and Rosenthal (1977) and Braumoeller and Goertz (2000) serve as a test for an asymmetric causal relationship in the monocausal, categorical setting.

However, most social processes of interest are not monocausal. This means that our test for the presence of heteroskedasticity may not be appropriate to evaluate asymmetric hypotheses in many situations. One obvious reason for this is omitted variable bias. If the heteroskedasticity in our data is caused systematically by a different variable and that variable is correlated with the independent (and dependent) variable of interest, then treating the disconfirming cases as if they were generated by a purely stochastic process will lead to biased inferences about the relationship that the researcher is interested in. In these cases, a multicausal approach will be required.[12] It is precisely for these sorts of relationships—probabilistic, asymmetric, multicausal—that the approach we describe below is useful.

Consider again our example of electoral system permissiveness and party system size. If the heteroskedasticity in Fig. 3 is due to a purely stochastic process, then we must conclude that "Duverger's Law" can explain why two-party systems occur but cannot really explain why some systems with permissive electoral rules have many parties but others have few. In fact, Duverger believed that the heteroskedasticity that we observe in Fig. 3 is the result of a systematic (and stochastic) process. At one point in his classic book, Duverger ([1954] 1963, 205) asserts that social factors are the primary force behind the formation of parties. This assertion, together with his claim that single-member district plurality (SMDP) electoral systems "nearly always produce two-party systems," logically implies that social factors must be an important part of the reason why some countries with permissive electoral rules have many parties and others have few (Clark and Golder 2006). In other words, an asymmetric, probabilistic, multivariate causal claim is a claim about

[11]See Tsebelis (2003) for an interesting discussion of the connection between heteroskedasticity and necessary/sufficient conditions.

[12]Incidentally, we should note that this point amounts to a serious criticism of the dichotomous, monocausal approaches mentioned above (Hildebrand, Laing, and Rosenthal 1977; Braumoeller and Goertz 2000) since they are incapable of including other causal factors. Hildebrand, Laing, and Rosenthal do give some indication as to how their approach could be extended to a multivariate setting; however, their suggestions are practical for only a relatively small number of ordinal variables.

a conditional relationship between variables.[13] In the next section, we show that linear multiplicative interaction models are adequate for the task of capturing the asymmetry behind necessity-like causal claims.

4 A Multiplicative Interaction Test of Asymmetric Hypotheses

The claim that X is necessary, but not sufficient, for Y raises the question of why some cases where X is present result in Y while others do not. In effect, a claim of necessity shifts the question from asking why Y sometimes occurs and sometimes does not to asking why, given that the necessary factor X is present, Y sometimes occurs and sometimes does not. In other words, it shifts the "causal field" (Mackie 1965) from the universe of cases to the subset of cases where X is present.[14] As noted earlier, answers to this question can take on two forms. Either Y occurs in some cases of X and not others for purely stochastic reasons or there are some other factors that determine why, given X, Y sometimes occurs and sometimes does not. The first case (to the extent that it actually exists) can be handled by testing for conditional heteroskedasticity as described above. Stasavage (2002) provides a nice example of this. He tests the claim that constraints on government action (checks and balances) are sufficient, but not necessary, to induce greater private investment because they assure investors that the state will not act opportunistically. Stasavage provides evidence for his claim by showing that the conditional variance of investment as a percent of GDP is lower when greater levels of "checks and balances" are present than when they are absent.

We now turn to the second case and propose an approach based on multiplicative interactive models. Consider the following linear multiplicative interaction model in equation (1), where X_1 and X_2 are thought to be alternative causes of Y.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + \varepsilon. \tag{1}$$

This model's relative simplicity belies the fact that it enables us to determine whether X_1 and/or X_2 is necessary, sufficient, or necessary and sufficient for Y. For example, imagine a claim that both X_1 and X_2 are individually necessary, but not sufficient, for Y. If this claim were true and assuming that both variables have a positive effect on Y, then it should be the case that $\beta_1 = \beta_2 = 0$ and $\beta_3 > 0$. Why? First, note that the marginal effect of X_1 on Y is $\beta_1 + \beta_3 X_2$ and that the marginal effect of X_2 on Y is $\beta_2 + \beta_3 X_1$. Second, note that the marginal effect of X_1 and X_2 on Y is β_3, i.e., $\partial Y/\partial X_1 \partial X_2 = \beta_3$. The claim that X_2 is necessary for Y implies that X_1 should have no effect in the absence of X_2, i.e., β_1 should be zero. Similarly, the claim that X_1 is necessary for Y implies that X_2 should have no effect in the absence of X_1, i.e., β_2 should be zero. Given the standard assumption that we have the correct and fully specified model, the claim that X_1 and X_2 are necessary for Y implies that both X_1 and X_2 need to be present for Y to occur, i.e., β_3 should be positive. Thus, we can conclude that each of the explanatory variables is necessary, but not sufficient, for Y whenever the coefficients on both constitutive terms are zero and the coefficient on the interaction term is different from zero.[15]

[13]This means that omitted variable bias is not the only problem for probabilistic, asymmetric, and monocausal claims. Even if the researcher were to include all the necessary variables, he would still miss the asymmetric nature of his hypothesis unless they were added in an interactive fashion.

[14]For this reason, it is odd that some scholars have suggested that claims of necessity warrant selecting on the dependent variable (Dion 1998; Braumoeller and Goertz 2000; Ragin 2000; Collier, Mahoney, and Seawright 2004).

[15]It should also be the case that the marginal effect of X_1 on Y, i.e., $\beta_1 + \beta_3 X_2$ and the marginal effect of X_2 on Y, i.e., $\beta_2 + \beta_3 X_1$ are both positive and significant when X_2 and X_1 are present, respectively (in the dichotomous case), or when X_2 and X_1 are sufficiently high, respectively (in the continuous case).

Now imagine a claim that X_1 and X_2 are both individually sufficient, but not necessary, for Y. If this claim were true and still assuming that both variables have a positive effect on Y, then it should be the case that $\beta_1 > 0$ and that $\beta_2 > 0$. The claim that X_1 is sufficient for Y implies that it should have an effect in the absence of X_2, i.e., β_1 should be positive. The claim that X_2 is sufficient for Y implies that it should have an effect in the absence of X_1, i.e., β_2 should be positive. There are, however, three ways in which two variables can each be sufficient to bring about an outcome. If the magnitudes of the effects of X_1 and X_2 on Y are the same regardless of the presence of the other sufficient condition, then β_3 should be zero.[16] However, if the effects of X_1 and X_2 are "reinforcing" in some way, then β_3 should be positive. In this latter case, we can think of X_1 and X_2 as being individually sufficient, but complementary, causes of Y. In contrast, if the magnitude of the effect of each sufficient condition is smaller when the other sufficient condition is present, then the coefficient on β_3 should be negative; if this is the case, we can think of X_1 and X_2 as "substitutes." This suggests how the concept of "equifinality" discussed by qualitative scholars such as Mahoney and Goertz (2006) and George and Bennett (2005) can be captured in a statistical model.[17] Note also that in the special case where the right-hand side variables in question are dummy variables and $\beta_1, \beta_2 > 0$ and $\beta_1 = \beta_2 = -\beta_3$, then we can say that X_1 and X_2 are "perfect substitutes" in the sense that in the absence of the other, each phenomenon has exactly the same expected effect on Y and in the presence of the other, adding the second phenomenon has no net effect on Y.[18] Space does not permit a discussion of all the logically possible combinations of coefficients in our simple interaction model. As a summary, we list all the logically possible combinations of coefficients (assuming, for simplicity's sake, only nonnegative coefficients) along with their interpretation in Table 3. Each combination is relevant for evaluating a particular claim about sufficiency or necessity.[19]

We should note that our interaction model allows us to talk about degrees of "necessity" or degrees of "sufficiency." For example, the magnitude of the coefficient on the cause that is purported to be necessary, but not sufficient, is a measure of "sufficiency." Conversely, the magnitude of the coefficient on the interaction term is a measure of the extent to which the purported cause is necessary. If β_1 is large relative to β_3 (and they have the same sign), the more "sufficient" is X_1 and the less "necessary" is X_2 for Y. Similarly, if β_2 is large relative to β_3, the less "necessary" is X_1 and the more "sufficient" is X_2 for Y. As you can see, the rather simple looking multiplicative interaction model outlined in equation (1) can throw considerable light on our hypotheses about necessity and sufficiency if they are interpreted thoroughly and carefully (Brambor, Clark, and Golder 2006). We now

[16]This suggests that a typical linear-additive specification amounts to an assumption that each variable is independently sufficient to bring about Y.

[17]See Braumoeller (2003) and Gordon and Smith (2004) for a more sophisticated approach to estimating substitutes, complements, and complex causation more generally.

[18]Note also that, in this case, the sufficient conditions (X_1 and X_2) have both "floor" and "ceiling" effects on Y. Consequently, we conjecture that the incipient debate between Goertz and Ragin on the presence of "floor" effects induced by sufficiency (Ragin 2000; Goertz 2001) can be resolved by arguing that they are both correct—there is a sufficiently large variety of types of sufficiency that cases exist to support each of their positions. This kind of logic can be generalized to the case where the right-hand side variables are continuous variables; however, the precise values of the coefficients that indicate perfect substitutability will depend on the scales upon which the variables are measured, and our confidence that we have identified perfect substitutes will depend on whether or not our variables are measured with "natural zeros" (Braumoeller 2004).

[19]Note that the definitions of "necessity" and "sufficiency" have implications for the conditional nature of all the other variables thought to influence Y. In other words, if there is a vector of "control" variables $\{X_3 \ldots X_k\}$ included in the model specification, then every implication that a claim of necessity or sufficiency has for β_1, β_2, and β_3 discussed above would also be true for the βs on each of these additional variables and their interactions with X_1 and/or X_2.

Table 3 All logically possible combinations of coefficients from a simple interaction model and their interpretation

Combination	β_1	β_2	β_3	*Valid conclusion about X*	*Valid conclusion about Z*
1	=0	=0	>0	Necessary for *Y*	Necessary for *Y*
2	>0	=0	>0	Necessary and sufficient for *Y*	Conditionally independent from *Y*
3	=0	>0	>0	Conditionally independent from *Y*	Necessary and sufficient for *Y*
4	>0	>0	>0	Sufficient for *Y* and reinforcing	Sufficient for *Y* and reinforcing
5	=0	=0	=0	Unconditionally independent from *Y*	Unconditionally independent from *Y*
6	>0	=0	=0	Necessary and sufficient for *Y*	Unconditionally independent from *Y*
7	=0	>0	=0	Unconditionally independent from *Y*	Necessary and sufficient for *Y*
8	>0	>0	=0	Sufficient for *Y*	Sufficient for *Y*

Note. We are assuming only nonnegative coefficients (for ease of presentation).

turn to an illustration of our proposed technique by returning, once again, to our running example of why some polities have many parties but others have few.

5 The Effect of Electoral Rules on Party System Size—An Application

Duverger ([1954] 1963) proposed three separate, but related, hypotheses in his theory of party system size:

Hypothesis 1: SMDP electoral systems almost always produce a two-party system—Duverger's Law.

Hypothesis 2: PR electoral systems encourage multiple parties—Duverger's Hypothesis.

Hypothesis 3: Sociological factors drive the formation of parties.[20]

The first hypothesis amounts to a claim of near sufficiency. In other words, SMDP systems are nearly sufficient for two-party systems. If having a single-member district is sufficient to produce a two-party system, then it must be the case that SMDP systems are absent if multipartism is to arise. It logically follows from this that multimember districts must be necessary for multipartism. But did Duverger also think that multimember districts were sufficient for multipartism? The word "encourage" in the second hypothesis suggests that he did not. Duverger thought that PR electoral systems "encourage" multiple parties in the sense that their presence is sometimes insufficient to produce multiple parties. Thus, the first and second hypotheses together imply that multimember electoral districts are necessary, but not sufficient, for multipartism.

[20]Duverger ([1954] 1963, 205) writes that "the multiplication of parties, which arises as a result of other factors, is facilitated by one type of electoral system and hindered by another. Ballot procedure, however, has no real driving power. The most decisive influences in this respect are aspects of the life of the nation such as ideologies and particularly the socio-economic structure." Although Duverger is often portrayed as the father of the so-called institutionalist approach to electoral and party systems, he clearly believed that social forces, not electoral rules, were the driving force behind the formation of parties. For a good overview of the ways in which Duverger has been misunderstood and misrepresented over the years, see Afonso Da Silva (2006).

A closer look at all three hypotheses reveals that the second hypothesis is not actually required to establish the claim that multimember districts are necessary, but not sufficient, for multipartsim. Note that if the first hypothesis is correct, then the third hypothesis must either be false or in need of qualification. Why? Well, the first hypothesis claims that a country with an SMDP system will only have two parties. If this is the case, then it cannot always be true that sociological factors increase the number of parties as the third hypothesis claims. Thus, the first and third hypotheses are inconsistent as they stand. However, they can be reconciled by qualifying the third hypothesis in the following way:

Hypothesis 3a: Sociological factors drive the formation of parties *in multimember districts*.

Thus, Duverger's three hypotheses indicate that multimember districts are necessary, but not sufficient, for multipartism since they imply that multipartism arises in some countries with multimember districts as a result of sociological factors but not in others. It is precisely because he recognized that multipartism may or may not occur in multimember districts that Duverger referred to his claim about the effect of PR electoral systems as a "Hypothesis," whereas he referred to his claim about the effect of SMDP systems as a "Law" (Clark and Golder 2006).[21]

Because it makes statements about the conditions under which multipartism will not occur, Hypothesis 1 raises the question of when multipartism will occur. In line with our earlier discussion, it begs the question as to why PR electoral systems sometimes produce multipartism and sometimes do not. Whereas Hypothesis 2 provides little help in answering this question, Hypothesis 3 provides a potential answer. Without Hypothesis 3, or something like it, we would be forced to conclude that the process determining whether multimember districts produce multipartism or not is purely stochastic. If, however, we have reason to believe that there are other causes of multipartism, as we do here, then an asymmetric claim about multimember districts amounts to the claim that those other factors have a different effect on multipartism when multimember districts are used than when they are not. In our specific example, our claim is that social forces are more likely to produce additional parties when countries employ multimember districts than when they do not.

We can easily evaluate this asymmetric causal claim with the following linear multiplicative interaction model:

$$\begin{aligned}\text{Legislative Parties} = {} & \beta_0 + \beta_1 \text{ Multimember District} + \beta_2 \text{ Social Heterogeneity} \\ & + \beta_3 \text{ Multimember District} \times \text{Social Heterogeneity} + \varepsilon, \qquad (2)\end{aligned}$$

where Legislative Parties is the effective number of legislative parties, Multimember District is a dichotomous variable indicating whether a country has single- or multimember districts, and Social Heterogeneity is a dichotomous variable indicating whether a country is ethnically heterogeneous (more ethnic groups than the median country) or ethnically homogenous (less ethnic groups than the median country). Note that Duverger's theory leads us to believe that both multimember districts and social heterogeneity are necessary, but not sufficient, to produce more legislative parties. As we can see from Table 3, a claim like this implies that $\beta_1 = \beta_2 = 0$ and that $\beta_3 > 0$.

The results from the model outlined in equation (2) are shown in column 1 of Table 4. As predicted, the coefficients on Multimember District and Social Heterogeneity are

[21]Riker (1982) placed great emphasis on the difference in the degree of determinism between Duverger's "Law" and his "Hypothesis." However, he did not realize that this difference in determinism actually follows from the conditional effects of societal influences on the party system.

Table 4 The determinants of multipartism with dichotomous and continuous predictor variables

Predictor variables	*Model 1* *Dichotomous variables*	*Model 2* *Continuous variables*
Electoral system permissiveness	0.55 (0.50)	−0.19 (0.30)
Social heterogeneity	−0.83 (0.55)	−0.36 (0.35)
Electoral system permissiveness × Social heterogeneity	1.65** (0.70)	0.48** (0.18)
Constant	2.52*** (0.41)	2.67*** (0.61)
N	54	54
Adjusted R^2	0.26	0.32

Note. Electoral system permissiveness: single- or multimember district (dichotomous); logged median district magnitude (continuous). Social heterogeneity: ethnically heterogeneous (above median) or homogenous (dichotomous); effective number of ethnic groups (continuous). Data are from Amorim Neto and Cox (1997). Effective number of legislative parties is the dependent variable. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ (two tailed).

indistinguishable from zero and the coefficient on the interaction term is both positive and statistically significant. If our asymmetric claim is correct, we should also expect the marginal effect of Social Heterogeneity ($\beta_1 + \beta_3$) to be positive and statistically significant when there are multimember districts.[15] This is indeed the case. The marginal effect of Social Heterogeneity in a country with multimember districts is 0.82. This effect is statistically significant at the 94% level.

Substantively, the results in column 1 of Table 4 indicate that an increase in the heterogeneity of a country is not expected to increase the size of the party system in countries with single-member districts. In fact, the estimated marginal effect of social heterogeneity is negative (though far from statistically significant). In contrast, Duverger's claim that social forces are the engines that drive party formation in multimember districts appears to be consistent with the evidence presented here. Thus, multimember districts are necessary, but not sufficient, for multipartism in the sense that (1) in their absence, the other known cause of multipartism (social heterogeneity) is not expected to bring it about and (2) in their presence, the other known cause of multipartism is needed to bring it about.

The results from this model can also be used to predict the number of parties under alternative institutional and sociological conditions (see Table 5). For example, the model predicts that there will be 2.52 legislative parties in homogenous societies employing single-member districts compared to 1.68 legislative parties in heterogeneous societies with single-member districts. Both of these predictions are consistent with Hypothesis 1—single-member districts almost never produce multiparty systems. Although the predicted number of parties differs between homogenous and heterogeneous societies employing single-member districts, this difference is not statistically significant (this difference is captured by the coefficient on Social Heterogeneity). Similarly, although the predicted number of parties differs between homogenous countries employing single-member and multimember districts, this difference is also not statistically significant (this difference is captured by the coefficient on Multimember District). In contrast, the difference in the expected number of parties between heterogeneous societies that employ multimember districts and those that employ single-member districts is statistically significant (this difference is captured by the coefficient on the interaction term).[22] This is

[22]The coefficient on the interaction term also captures the difference between homogenous and heterogeneous societies that employ multimember districts.

Table 5 The predicted number of legislative parties

	Single-member districts	*Multimember districts*
Socially heterogeneous	1.68 (0.94–2.43)	3.88 (3.26–4.50)
Socially homogenous	2.52 (1.69–3.34)	3.06 (2.48–3.65)

Note. The 95% confidence intervals are in parentheses.

exactly as predicted. In sum, the results in column 1 of Table 4 provide strong evidence that both multimember districts and social heterogeneity are necessary, but not sufficient, for more legislative parties.

Although the connection between multiplicative interaction models and testing for necessary and/or sufficient conditions is clearest when we have dichotomous predictor variables as in this example, the asymmetry behind necessary or sufficient conditions is easily retained when we have continuous explanatory variables. To illustrate this, let us now restate Duverger's hypotheses in terms of a continuous measure of social heterogeneity (effective number of ethnic groups) and electoral system permissiveness (natural log of median district magnitude).

Hypothesis 1^C: Small district magnitudes almost always produce a two-party system.
Hypothesis 2^C: Large district magnitudes encourage multiple parties.
Hypothesis 3^C: Social heterogeneity encourages the proliferation of parties when the district magnitude is sufficiently large.

These hypotheses can be captured with the following linear multiplicative interaction model:

$$\begin{aligned}\text{Legislative Parties} = \gamma_0 + \gamma_1 \ln(\text{Magnitude}) &+ \gamma_2 \text{ Social Heterogeneity} \\ &+ \gamma_3 \ln(\text{Magnitude}) \times \text{Social Heterogeneity} + \varepsilon. \qquad (3)\end{aligned}$$

Larger district magnitudes can be thought of as being necessary, but not sufficient, for more parties when (1) $\gamma_1 = 0$ and $\gamma_3 > 0$ and (2) $\gamma_2 + \gamma_3 \ln(\text{Magnitude})$ is positive and statistically significant for a nontrivial range of ln(Magnitude). This second condition insures that there is a district magnitude large enough to help the identified other cause (social heterogeneity) to have its hypothesized effect on party system size. That is, if $\gamma_2 + \gamma_3 \ln(\text{Magnitude})$ is never different from zero, then having large district magnitudes may be trivially necessary for more parties in the sense that while social heterogeneity does not bring about larger party systems in their absence, it also does not bring them about in their presence.

The results from the model in equation (2) are shown in column 2 of Table 4. As predicted, the coefficients on both constitutive terms are statistically indistinguishable from zero, and the coefficient on the interaction term is relatively large and positive. This suggests that social heterogeneity has no distinguishable causal effect on party system size when district magnitude is one, i.e., when $\ln(\text{Magnitude}) = 0$. Similarly, an increase in district magnitude also has no distinguishable causal effect on party system size when a society is entirely homogenous. However, the positive and relatively large coefficient on the interaction term means that social heterogeneity will increase party system size when the district magnitude is sufficiently large.

To determine exactly how large the district magnitude needs to be for social heterogeneity to have its hypothesized positive effect on party system size, we need to calculate the

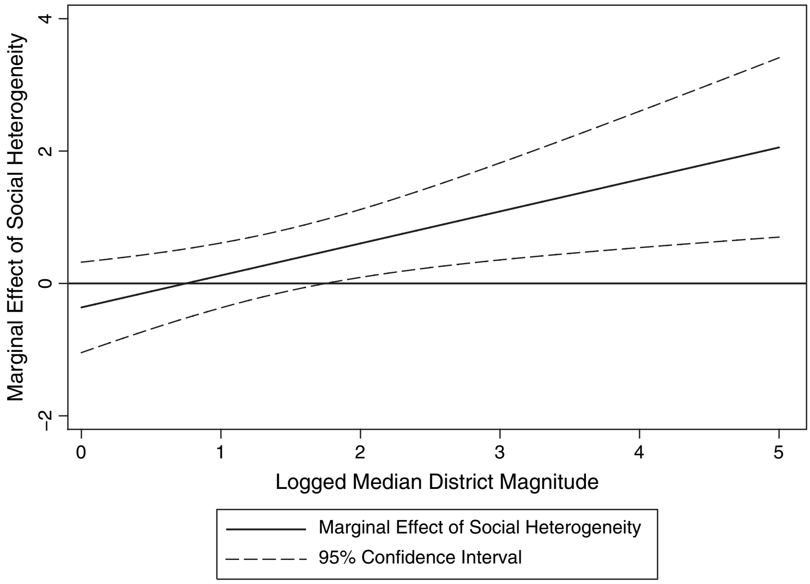

Fig. 4 Marginal effect of social heterogeneity. Effective number of legislative parties is the dependent variable.

marginal effect of social heterogeneity ($\gamma_2 + \gamma_3$ ln(Magnitude)) and its associated confidence intervals across the observed range of district magnitudes. These are plotted in Fig. 4. As you can see, social heterogeneity has no discernible effect on the number of legislative parties when logged average district magnitude is close to zero. However, as predicted, the effect of social heterogeneity becomes both clear and pronounced once the district magnitude becomes sufficiently large. Specifically, social heterogeneity has a positive and significant effect on party system size whenever the district magnitude is greater than six (or whenever ln(Magnitude) is greater than 1.8).

6 Discussion

The approach that we advocate above can accommodate other well-known approaches to causality. One such approach, often promoted by qualitative researchers, is Mackie's (1965) classic "INUS" approach, in which a cause is defined as the Insufficient but Necessary part of an Unnecessary but Sufficient relationship. Mackie illustrates his concept of a causal field with the example of trying to explain why some people contract a certain strain of the flu and others do not. In order to contract the flu, it is first necessary to be exposed to the flu virus. Within the class of people who have been exposed to the flu virus, only those whose immune systems are sufficiently weak will actually contract the disease. Although a particular case of the disease is caused by exposure to the virus according to the INUS approach, this is only a necessary condition since some people who are exposed to the virus will never contract the disease. Mackie's notion of a causal field implies that we can think of our causal explanations as the process of defining ever narrower subsets of cases on which the causal process works. Figure 5 illustrates this approach using Mackie's flu example.[23] We begin by

[23]The Venn diagram in Fig. 5 automatically dichotomizes the data—a case is either in a set or it is not. As described above, however, our approach does not require dichotomous data. We use the Venn diagram only for expositional convenience.

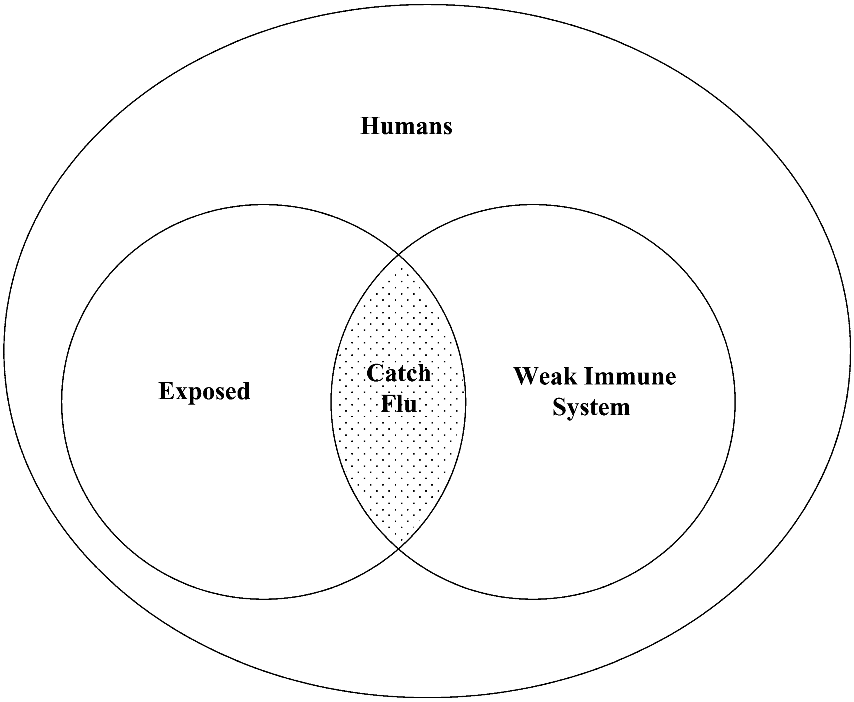

Fig. 5 Mackie's causal fields and flu acquisition.

defining the broadest class of cases that may contract the disease. In this case, it is the class of all human beings. Within this class, there is a subclass of people who have been exposed to the flu virus. Within this subclass, there is a further subclass of people whose immune systems are not sufficiently robust to resist the disease. It is this last subclass that contains the group of people who contract the disease. Thus, causal explanation according to this approach amounts to defining ever narrower subsamples of cases until we are left with a class that possesses the needed combination of necessary and sufficient conditions to generate the outcome to be explained.

Notice, however, that the process of defining narrower and narrower subsamples is methodologically equivalent to interacting the variable that defines the subsample with the other causal variables in our model. In this particular example, the INUS approach suggests estimating a model of flu acquisition on the subsample of cases in which the subject has been exposed to the virus and using the robustness of a subject's immune system as the only independent variable. The coefficient on this independent variable would provide a useful summary of the joint causal effects of being exposed to the virus and having a weak immune system. If we ran the same model on the sample of people not exposed to the virus, we would presumably find that a subject's immune system has no effect on flu acquisition since people with robust immune systems are no less likely to contract the disease than people with weak immune systems.[24]

[24]Of course, this second model is not, in fact, possible because there would be no variation on the dependent variable—none of the nonexposed people would have contracted the disease. This relationship may not be so strong in social science data, and there may be variation on the dependent variable in this subsample. If so, we could conduct the test just described.

In contrast to this, our approach would be to run a regression of flu acquisition on the full sample (all human beings) and include three independent variables: "exposed to virus," "weakness of immune system," and an interactive term between the two. As you can see, the two approaches—our interactive approach and the INUS approach described in the previous paragraph—are nearly equivalent methodologically. Both give us the piece of information we want—the causal impact of the joint condition "exposed to virus and weakness of immune system." Our argument here runs directly counter to those qualitative researchers who argue that INUS approaches (George and Bennett 2005) and "subsetting" (Ragin 2000) are quite different from standard statistical methodologies. In this sense, our argument is similar in spirit to Lieberson's (1991) argument that the requirements for valid inference from Mill's (1986 [1874]) methods are essentially the same as those for valid inference from applied regression models, as well as Seawright's (2005) argument that the requirements for valid inference using Ragin's "Qualitative Comparative Analysis" are also essentially the same as those required for valid inference from applied regression models.

There are some important differences with the INUS approach, though. The most important is that the interactive approach we describe in this paper more readily supplies information as to whether the independent variables of interest constitute a symmetric or an asymmetric causal relationship. Our approach produces a regression coefficient on the variable "weakness of immune system." This coefficient allows us to test systematically if the impact of weakness of immune system given no exposure to the virus is actually zero as is required by the asymmetric causal claim that exposure to the virus is necessary but not sufficient. The INUS approach described earlier does not automatically perform this test. A second advantage of the interactive approach is that, in most cases, it will preserve degrees of freedom (df) in a way that sample splitting does not.[25] Imagine we have observations on 20 people, half of whom have been exposed to the virus. Sample splitting would result in two regressions with 8 df each. Our approach would require one regression with 16 df. As a result, the coefficients from our interactive approach are likely to be estimated with greater accuracy.[26]

A second important approach to causality, which is becoming increasingly prevalent in political science, is the Rubin causality model.[27] Unfortunately, our discussion of this model is more speculative at this point than our discussion of the INUS causality model. However, we conjecture that the Rubin model can be amended to incorporate the kinds of asymmetric hypotheses we discuss in this article. Those who employ the Rubin causality model speak in the language of experiments, and their goal is to determine the effect of a given treatment T on some outcome variable Y. The key assumptions of the Rubin causality model are (1) unit homogeneity, (2) the stable unit treatment values assumption (SUTVA), and (3) conditional mean independence. The first assumption implies that two units with identical covariates will have the same values of Y. Assuming SUTVA simply means assuming that the outcome of one case does not affect the performance of another case. For example, in terms of our discussion of Duverger's theory, SUTVA would be

[25]We say "in most cases" because one can imagine cases where this statement is not true. These would be cases with a huge number of independent variables and in which we need to interact all the independent variables with each other in order to test our causal story. We believe that such cases are going to be exceedingly rare in practice.

[26]For a more detailed discussion of the advantages of the interactive approach over sample splitting, see Kam and Franzese (forthcoming).

[27]Our discussion of the Rubin causality model is based on Przeworski (forthcoming) and Rosenbaum (2002). The approach is often called the Rubin causality model in political science because of Donald Rubin's many important contributions to the topic since his seminal piece in 1974.

violated if voters in one country learned of the maladies (or benefits) of multipartism in a neighboring country and changed their voting behavior accordingly.

The assumption of conditional mean independence requires a somewhat lengthier discussion. Consider the case where a researcher wants to test the hypothesis that some treatment T is a sufficient-like asymmetrical cause for a particular outcome Y. As indicated by our discussion earlier, if such a claim is correct, then the values of Y in the treated cases should be more alike (less dispersed) than they are in the untreated cases. Put differently, removing the treatment should make the Y values more dispersed. The reason for this is that the sufficient-like effect of the treatment causes similar outcomes of Y regardless of the effect of the other independent variables (both observed and unobserved); the values of Y in the untreated cases will take on a variety of values depending on the effects of the other covariates. Similarly, in the case where treatment T is hypothesized to have a necessary-like causal effect, the control (i.e., the untreated) cases should exhibit less dispersion in the values of Y than the treated cases; adding the treatment should cause this dispersion to increase. Our discussion of Duverger's theory offers an example of this latter case. The treatment in our example can be thought of as multimember districts. As Table 2b and Fig. 3 illustrate, cases that were treated with multimember districts exhibited greater dispersion in the size of their party systems than the control cases (i.e., singe-member districts).

The assumption of conditional mean independence in the Rubin model of causality means that, *conditional on observed covariates*, the control units would respond to the treatment in the same way as those units that were actually exposed to the treatment. The italicized portion of the previous sentence is important because, as we just discussed, a sufficient-like treatment will cause the values of Y to cluster around a similar level. This means that control cases whose values of Y are farther away from this level would respond to treatment to a greater extent (that is their values of Y will change by more) than those control cases whose values of Y are closer to the level of Y exhibited by the treated cases. To address this issue, it is important to condition on the covariates that cause the dispersion in the control state (or in the treated state for a necessary-like claim). In practice, this is done by matching treated and control cases on those covariates. As an aside, one of the main shortcomings of such techniques is that the researcher can only match on observed covariates, which amounts to the implicit assumption that by matching on observed covariates, one is also matching on the unobserved ones.

Our purpose in dwelling on these assumptions is to point out that none of them are violated by attempting to test for asymmetric causality in general, although obviously any of them (particularly SUTVA and the assumption that the unobserved covariates are matched) may be violated for specific research questions. Thus, although the application of the Rubin causality model to asymmetric hypotheses must remain a topic for future research, we see no reason why such an approach could not be developed.

The reason that we believe the development of such a technique is worthwhile and the problem with using unadulterated matching techniques to test asymmetric causal claims is that the most-often-used matching techniques may lead to the inappropriate rejection of asymmetric hypotheses in much the same way that regular regression analysis might. To illustrate the potential problem, consider the case of Liechtenstein in Fig. 3. Although it has a relatively high district magnitude, Liechtenstein's effective number of parties is on par with that of the United Kingdom that employs single-member districts. Although Liechtenstein has been "exposed" to the necessary-like "treatment" of a relatively large district magnitude, the "treatment" has seemingly caused little increase in the size of its party system. According to Duverger's theory, this is presumably because Liechtenstein has a relatively low level of social heterogeneity. The most commonly used matching techniques will simply average the relatively muted causal effect of units like Liechtenstein (muted in ways that are consistent with the asymmetric causal claim) with the more robust causal effects of other units. On average, the causal effect may appear to be statistically insignificant because of cases like Liechtenstein even though such cases are completely consistent with the asymmetric nature of the causal claim. Developing a solution to this problem is the subject of ongoing research.[28]

Finally, no discussion of the use of linear techniques to assess causality would be complete without the well-known qualification that although causation implies correlation, correlation does not imply causation. All of the myriad problems that have long been familiar to those who study symmetric causes (endogeneity, selection bias, specification error, and so on) are equally problematic in our approach to asymmetric causes. Our conjecture is that the approach that we discuss in this article could be amended to deal with such issues and, indeed, our earlier discussion of the Rubin causality model may suggest one avenue for how scholars can deal with asymmetric causation when the treatment in question is not randomly assigned. We leave these problems as a topic for future research.

7 Conclusion

In this paper, we have argued that claims of necessity and sufficiency involve a type of asymmetric causal claim that is useful in many social scientific contexts. In contrast to many qualitative scholars, though, we argue that there is nothing about such asymmetries that should lead scholars to depart from standard social science practice. We accept as given that deterministic and monocausal tests are inappropriate in the social world and argue that standard linear multiplicative interaction models are up to the task of handling asymmetric causal claims in a multivariate, probabilistic manner. Linear regression and its nonlinear counterparts are the bread and butter of scholars who seek to test symmetric, probabilistic, and multicausal hypotheses. Such scholars are deeply familiar with the notion that treating such hypotheses as deterministic or monocausal could lead to the inappropriate rejection of valid hypotheses. We are simply making the same claim for asymmetric hypotheses; namely that testing such hypotheses with techniques designed for symmetric claims may lead to the inappropriate rejection of valid causal claims. It is in response to this that we propose including simple multiplicative interaction terms in our linear models as a convenient way to give such claims a more appropriate test.

References

Afonso Da Silva, Virgílio. 2006. Duverger's Laws: Between social and institutional determinism. *European Journal of Political Research* 45:31–41.

Amorim Neto, Octavio, and Gary Cox. 1997. Electoral institutions: Cleavage structures and the number of parties. *American Journal of Political Science* 41:149–74.

Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. Understanding interaction models: Improving empirical analyses. *Political Analysis* 14:63–82.

Braumoeller, Bear. 2003. Causal complexity and the study of politics. *Political Analysis* 11:209–33.

———. 2004. Hypothesis testing and multiplicative interaction terms. *International Organization* 58:807–20.

Braumoeller, Bear F., and Gary Goertz. 2000. The methodology of necessary conditions. *American Journal of Political Science* 44:844–58.

[28]Chapter 5 of Rosenbaum (2002) offers some intriguing possibilities for how such a solution might be accomplished.

Clark, William Roberts, and Matt Golder. 2006. Rehabilitating Duverger's theory: Testing the mechanical and strategic modifying effects of electoral laws. *Comparative Political Studies* 39.

Collier, David, James Mahoney, and Jason Seawright. 2004. Claiming too much: Warnings about selection bias. In *Rethinking social inquiry: Diverse tools, shared standards*, eds. Henry E. Brady and David Collier, 85–102. Lanham, MD: Rowman and Littlefield.

Dion, Douglas. 1998. Evidence and inference in the comparative case study. *Comparative Politics* 30:127–45.

Duverger, Maurice. [1954] 1963. *Political parties: Their organization and activity in the modern state.* New York: John Wiley and Sons, Inc.

George, Alexander L., and Andrew Bennett. 2005. *Case studies and theory development in the social sciences.* Cambridge, MA: MIT Press.

Goertz, Gary. 2001. Necessary condition theory: What Kingdon, Ostrom, and Skocpol have in common. AZ: Department of Political Science, University of Arizona.

———. 2002. The substantive importance of necessary condition hypotheses. In *Necessary conditions: Theory, methodology, and applications*, ed. Gary Goertz and Harvey Starr. New York: Rowman and Littlefield.

Gordon, Sandford C., and Alastair Smith. 2004. Quantitative leverage through qualitative knowledge: Augmenting the statistical analysis of complex causes. *Political Analysis* 12:233–55.

Hildebrand, David K., James D. Laing, and Howard Rosenthal. 1977. *Analysis of ordinal data: Quantiative applications in the social sciences.* Beverley Hills, CA: Sage Publications.

Kam, Cindy D., and Robert J. Franzese, Jr. Forthcoming. *Modeling and interpreting interactive hypotheses in regression analysis: A brief refresher and some practical advice.* Ann Arbor: University of Michigan Press.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing social inquiry: Scientific inference in qualitative research.* Princeton, NJ: Princeton University Press.

Lakatos, Imre. 1970. Falsification and the methodology of scientific research programmes. In *Criticism and the growth of knowledge*, ed. Imre Lakatos and Alan Musgrave, 91–195. Cambridge: Cambridge University Press.

Lieberson, Stanley. 1987. *Making it count: The improvement of social research and theory.* Berkeley, CA: University of California Press.

———. 1991. Small N's and big conclusions: An examination of the reasoning in comparative studies based on a small number of cases. *Social Forces* 70:307–20.

Mackie, John L. 1965. Causes and conditions. *American Philosophical Quarterly* 2:245–64.

Mahoney, James, and Gary Goertz. 2006. A tale of two cultures: Contrasting quantitative and qualitative research. *Political Analysis* 10.1093/pan/mpj017.

Mainwaring, Scott. 1993. Presidentialism, multipartism, and democracy: The difficult combination. *Comparative Political Studies* 26:198–228.

Mill, John Stuart. 1986 [1874]. *A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation.* 8th ed. New York: Harper and Brothers.

Popper, Karl. 1959. *The logic of scientific discovery.* New York: Basic Books.

Przeworski, Adam. 2006. Forthcoming. Is the science of comparative politics possible? In *Oxford handbook of comparative politics*, ed. Carles Boix and Susan Stokes. New York: Oxford University Press.

Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and development: Political institutions and well-being in the world, 1950–1990.* New York: Cambridge University Press.

Ragin, Charles C. 1987. *The comparative method: Moving beyond qualitative and quantitative strategies.* Berkeley, CA: University of California Press.

———. 2000. *Fuzzy-set social science.* Chicago, IL: University of Chicago Press.

———. 2006. Set relations in social research: Evaluating their consistency and coverage. *Political Analysis* 10.1093/pan/mpj019.

Riker, William H. 1982. The two-party system and Duverger's Law: An essay on the history of political science. *American Political Science Review* 76:753–66.

Rosenbaum, Paul R. 2002. *Observational studies.* 2nd ed. New York: Springer.

Rubin, Donald. 1974. Estimating the causal effects of treatments in randomized and nonrandomized studies. *Journal of Education Psychology* 66:688–701.

Seawright, Jason. 2005. Qualitative comparative analysis vis-à-vis regression. *Studies in Comparative International Development* 40:3–26.

Stasavage, David. 2002. Private investment and political institutions. *Economics and Politics* 14:41–63.

Tsebelis, George. 2003. Veto players and law production in parliamentary democracies: An empirical analysis. In *Necessary conditions: Theory, methodology, and applications*, ed. Gary Goertz and Harvey Starr, 249–76. Lanham, MD: Rowman and Littlefield.