

Conflict Management and Peace Science

<http://cmp.sagepub.com>

Explaining Interstate Conflict and War: What Should Be Controlled for?

James Lee Ray

Conflict Management and Peace Science 2003; 20; 1

DOI: 10.1177/073889420302000201

The online version of this article can be found at:
<http://cmp.sagepub.com/cgi/content/abstract/20/2/1>

Published by:



<http://www.sagepublications.com>

On behalf of:



Peace Science Society (International)

Additional services and information for *Conflict Management and Peace Science* can be found at:

Email Alerts: <http://cmp.sagepub.com/cgi/alerts>

Subscriptions: <http://cmp.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.co.uk/journalsPermissions.nav>

Citations <http://cmp.sagepub.com/cgi/content/refs/20/2/1>

EXPLAINING INTERSTATE CONFLICT AND WAR: WHAT SHOULD BE CONTROLLED FOR?*

James Lee Ray
Vanderbilt University

Abstract. Most multivariate models designed by analysts of international conflict focus on one key explanatory factor and include several control variables. There are prominent norms or customs in the subfield of international politics regarding the construction of multivariate models and the selection of control variables. Several of these norms or customs may make the results of multivariate analyses confusing and difficult to interpret. Analysts typically do not, for example, distinguish between confounding and intervening variables even though the implications and impacts of such variables are substantially different. Most researchers also fail to distinguish between confounding variables and variables that have an impact on interstate conflict that is complementary to that of the key explanatory factor. Commonly, control variables are included in a model for no other reason than that they also have an impact on interstate conflict or some other outcome variable. In some recent analyses, "independent" variables are included that are related by definition to the key explanatory variable, or to each other. This practice introduces into multivariate models artifactual, misleading degrees of statistical association between variables related to each other by definition with tautological relationships masquerading as empirical causal connections that complicate the interpretation of results. Finally, the construction of pooled cross-sectional, time series analyses is consistently based on the assumption that the key explanatory factor, as well as the control variables, have essentially identical impacts on interstate conflict across space, and over time. Substantial evidence, some of which is provided in this paper, suggests that this assumption is unwarranted. This paper provides five guidelines for the construction of multivariate models that address these issues in a manner aimed at making the results of multivariate analyses more intelligible and credible.

INTRODUCTION

If one can judge by the structure of such models in articles published in leading journals over the last ten years or so, there are norms or customs in the field of international politics in general, as well as in the subfield devoted to the systematic analysis of international conflict and

*Presidential Address to the Peace Science Society, University of Arizona, Tuscon, Arizona, November 2, 2002. I would like to thank Katherine Barbieri, Bruce Bueno de Mesquita, Douglas Lemke, Richard Stoll, Richard Tucker, and John Vasquez for insightful, careful and helpful comments on an earlier draft of this paper. Kwang Teo and Yijia Wang provided valuable assistance with data management and analysis.

war, regarding the construction and interpretation of multivariate models. My purpose here is to analyze the validity or utility of a number of these norms. This paper is evoked by my impression that there are no widely accepted, authoritative answers to at least some of the questions and issues raised by these norms. That impression has been confirmed by my review of all articles on topics in the subfield of international politics incorporating control variables in multivariate statistical models in the *American Political Science Review*, *International Interactions*, *International Studies Quarterly*, *Journal of Conflict Resolution* and the *Journal of Peace Research* in the years 1999, 2000, 2001, and the first half of 2002. If there were widely accepted, authoritative, logical and valid ways of dealing with the issues I intend to discuss here then multivariate models would not be constructed and interpreted, in my opinion, in the way they are on a consistent basis in major journals in the field. In fact, I am convinced that there are several important norms and customs regarding the construction of multivariate models in general and the addition or inclusion of control variables in particular that are counter-productive. They hinder, complicate, and decrease our understanding of the complex phenomena we are investigating. To anyone who might wonder by what authority or on what basis I take it upon myself to address these issues, I would point out that I am the president of the Peace Science Society (International). And you are not.

THE BASIC AIMS OF MULTIVARIATE MODELS

My review of many multivariate models in recent, prominent publications convinces me that there are two basic purposes to which such models are typically addressed. The first of these is a kind of ideal that has probably become less commonly pursued in recent years than it used to be. Some analysts are apparently, sometimes even explicitly aiming at constructing the best, most potent multivariate model they can devise. If the phenomenon to be explained is interstate war, for example, then an analyst might put together, say, eight independent or explanatory variables and try to demonstrate that these eight variables together constitute the best possible multivariate explanation of interstate war that exists. In such endeavors, the implicit argument needs to be that the explanatory power of this model cannot be improved significantly by the addition of any alternative variable. In other words, the argument would be that no really important variable has been left out of the model. And no variable included in the model would be superfluous. Removing any single variable from such an ideal model would

significantly detract from its ability to account for variation in the dependent, or outcome variable.

My impression is that analysts in the field of international politics, or peace science, were more likely to adopt this kind of ideal model as a goal twenty or even thirty years ago. Analysts would, for example, assemble data on a relatively large number of variables and then subject them to "path analysis." And the results would be presented as the putative "best possible" multivariate explanation of the phenomenon in question.

In recent decades, this kind of multivariate model appears less frequently in important books and journals in international politics, and even perhaps in political science in general. Certainly in research on the causes of war, general models aimed at the best fit for the model as a whole seem to have given way almost entirely to models whose basic purpose is to evaluate the impact of one key factor. Variables beyond that one key factor are added almost entirely for the purpose of providing a more sophisticated, thorough, and rigorous evaluation of a key hypothesis in question than would be possible with bivariate analyses. Most specifically, explicitly or implicitly, control variables are added to multivariate models in order to see whether the relationship of special interest persists. The implicit argument or assumption is that if a key relationship cannot survive the addition to the model of control variables then that relationship is exposed as less interesting. In short, it might be spurious.

Let me acknowledge here that most of the models and analyses I have in mind specifically here are those aimed in the last decade or so at the evaluation of the "democratic peace" hypothesis. But the issues I want to discuss are certainly not limited to that research on that topic. Typically, a lengthy list of control variables is added to multivariate models aimed at the evaluation of the hypothesis regarding the impact of democracy (or some other key factor) on interstate conflict. Authors who construct such models of this kind never, in my experience, assert that "this is the best possible multivariate model of interstate war that can be assembled..." Rather, the argument or conclusion is that "democracy and peace correlate with each other, and that correlation persists even when multiple control variables are added to the model. Therefore, the democratic peace hypothesis is deserving of more confidence than would be the case were only bivariate analyses provided."

Partly because most of what I have to say here about research on international conflict is critical, let me make a positive point here at the onset. I believe it usually makes more sense to construct multivariate models for the purpose of evaluating the impact of a key factor than it

does to try to put together the best possible overall multivariate model. My reasoning here is straightforward. We have no theory, or theoretical approach, that will tell us what are the best six, seven, or eight predictor variables to put into a multivariate model aimed at accounting for interstate war or conflict. That is, we have no sound theoretical basis for saying that "these are the seven most important variables on which to base any explanation of interstate wars. It is these seven and only these seven. No other variable is as important as these seven, and it necessary to include only these seven."

So, let me assert at the outset that I believe it is perfectly appropriate to utilize multivariate models in the way they are more commonly used in research in international politics, or peace science. That is, they are generally addressed to an evaluation of the relationship between one key explanatory factor, and the outcome of most interest.

The main point of this paper, however, is that analysts of international politics in general and interstate conflict in particular have in my view adopted a series of norms or customs regarding the construction and interpretation of multivariate models that have what seem to me to be unfortunate impacts on the quality of the evidence brought to bear on the evaluation of the relationships of special interest. What I will (immodestly) do here is to offer a series of guidelines that I believe could, if widely accepted and put into practice, improve the quality of evidence produced by multivariate analyses, as well as the credibility and intelligibility of that evidence.

GUIDELINE #1: Do not control for intervening variables.

First, let me reiterate a point others, and myself, have made elsewhere. (See Ray 2000). By definition, a confounding variable is an *antecedent* third factor that brings about a statistical association or correlation between two other variables. In order to do that, it must be correlated with *both* of those two other variables. However, there is another type of variable, namely an *intervening* variable that is *also* statistically associated with two original variables of interest which, if controlled for, will eliminate the statistical correlation between those original independent and dependent variables. As Blalock pointed out decades ago,¹ "In the simple three-variable case, the models $X \rightarrow Z \rightarrow Y$ [Z here is a confounding variable], and $X \rightarrow Z \rightarrow Y$ [in this case, Z is an intervening variable] yield the same empirical prediction that $r_{xyz} = 0$ "

¹ At the time, I should probably acknowledge, when I received my formal training in statistics, methodology, and research design. Typically, in order to attend classes on those topics, I walked through snow drifts of two or three feet.

(Blalock 1964, 84). This is the main reason that King, Keohane and Verba (1994, 173) assert that “in general, *we should not control for an explanatory variable that is in part a consequence of our key causal variable*” (emphasis in the original). In other words, one should not control for a factor that is (1) a consequence of a key causal variable, and which then in turn (2) has an impact on the outcome variable.

The main reason that we should not do this is described succinctly by Blalock (1964, 85). He explains that if one adds an intervening variable to a multivariate model, and this modification eliminates the statistical association between the original key explanatory factor being evaluated and the outcome variable, then one has engaged in “interpretation” of that relationship. Such “interpretation” does not make the original relationship in question less interesting. On the contrary, “through interpretation one is putting the frosting on the cake....He is not discovering anything radically wrong with the notion that X causes Y. He is merely making it seem more plausible by finding the intermediate links.” This is a fundamentally different situation than that resulting from the addition of a potentially confounding variable to a model that eliminates the correlation between the original independent and dependent variables. In that case, one *is* discovering that there is something radically wrong with the notion that X causes Y, and any theory or hypothesis positing such a causal link is accordingly discredited.

This may seem a simple, even intuitively obvious, point. However, it is rather consistently overlooked. An article published in the *American Political Science Review* in 2002, for example, explains that “we also include five control variables that previous research has shown to have an impact on war and dispute involvement.”² The apparent implication is that this is the only characteristic that an appropriate control variable needs to have. But if the control variable in question is an *intervening* variable as opposed to a potentially confounding variable including that variable in the analysis can produce quite misleading results for reasons just pointed out above.

Another article published in the *American Political Science Review* in 2001 provides the following rationale for the control variables introduced into the model in question: “We include several control variables in our analysis of crisis outcomes...It is important that we include these variables because they are correlated with democracy. As such, they represent potentially confounding variables” (Appendix. See footnote 2.) But, being correlated with the key explanatory factor,

² The source for this quote, and many of those to follow, can be found in an Appendix to this paper available from the author by request to james.l.ray@vanderbilt.edu or from the editor of CMPS.

democracy in this case, these control variables *also* represent potentially *intervening* variables, and thus much less promising candidates for control variables.

Just to drive home the point that this problem is common in the recent research in the field, let me cite the rationale for control variables in two more recent articles. The first one was published in the *American Political Science Review* in 2000. It asserts that "it is also important to take account of variables that might be responsible for any observed relationship between regime type and bilateral trade" (Appendix). And an article published in the *International Studies Quarterly* in 1999 explained that "democracies may be more likely to align with one another than nondemocratic states, indicating the importance of including an alliance variable to clearly delineate the impact of democracy on conflict intensification" (Appendix). These are two more rationales that ignore, or certainly seem to, the guideline being proposed here. In response to the first rationale, it needs to be pointed out that both confounding *and* intervening variables are potentially "responsible for" a relationship between two other variables. This does not serve as a *convincing rationale for controlling for potentially intervening variables*. With respect to rationale in the 1999 *International Studies Quarterly* article, I would point out that alliance ties are almost certainly an *intervening* variable in the process leading from democracy to conflict levels. To argue that alliance ties are a potentially confounding variable would be to suggest that alliance ties between any two states should make them more likely to be democratic. This is a plainly implausible proposition.³

Note that these four examples are found in the discipline's "flagship" journal in three cases, and in the principle publication of the International Studies Association in the fourth case. They should suffice, I hope, to establish the point that the distinction between confounding and intervening variables is not routinely taken into account by analysts in the field, even in highly prestigious journals. It is, in fact, routinely ignored.

GUIDELINE #2: Distinguish between complementary and competing explanatory factors.

My review of recent articles relying on multivariate models makes it

³The proposition that any state allied to a democratic state is more likely to be (or become) democratic is not so implausible. But that is a proposition pertaining to the directed dyadic, or sub-dyadic level of analysis, and is not the proposition implied by the analysis or article in question.

quite apparent that authors rather consistently believe that any factor that *also* correlates with the outcome variable of major interest constitutes an explanatory factor that *competes* with explanations based on the key factor in which the study is most interested. The author of an article published in the *American Political Science Review* in 2000, for example, explains that “a serious test requires a battery of economic, institutional, and domestic political controls to minimize the possibility that any...correlation is spurious” (Appendix). The control variables included in this research effort are, it turns out, simply *alternative* causes of the outcome phenomenon of interest. They could not plausibly be categorized as antecedent correlates of both the independent and outcome variable in the study, i.e., confounding variables.

Another author of a paper published in the *American Political Science Review* in 2001 asks: “Do the results become spurious once exposed to competing explanations?” S/he goes on to explain that “to investigate whether the learning hypotheses remain valid under such conditions, I will introduce three control variables; the first two capture geopolitical determinants, and the last is an additional liberal factor” (Appendix). Once again, however, the control variables selected in this paper are not plausibly related to the original explanatory factor (as either confounding or intervening variables); they are instead simply alternative factors that might plausibly be argued to *also* have an impact on the outcome variable in question. The author apparently considers them *competing* simply because they are also correlated with the outcome factor in the study.

An article published in the *International Studies Quarterly* in 2000 asserts that “in order to show the heightened relevance of domestic politics, we must ensure that we control for rival explanations—particularly those on the international level” (Appendix). But these “rival” explanations are, once again, simply *alternative* causes of the use of force, in this case. They are not factors that are antecedent to and plausibly associated with both the domestic factors being considered and the use of force.

Finally, the authors of an article published in the *Journal of Conflict Resolution* in 2001 explain that “other factors besides political institutions must be controlled for that could influence the amount of trade between countries. For this reason, we incorporate a number of control variables into our empirical analysis” (Appendix). First, to hark back for a moment to the previous guideline, these “other factors” that also influence the amount of trade could be intervening variables, and therefore arguably should not be controlled for in an attempt to evaluate the impact of political institutions. But they might also be nothing more

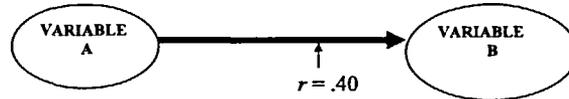
than alternative causes, rather than confounding variables.

Why does it make a difference whether a control variable that is introduced into a model is really a competing explanatory factor (i.e., a confounding variable), or simply a complementary explanatory factor that is *also* statistically associated with the amount of trade between countries, or any other outcome variable of interest? Because introducing competing explanatory factors, on the one hand, and complementary explanatory factors, on the other, can have precisely opposite impacts on the key relationship of interest within multivariate models.

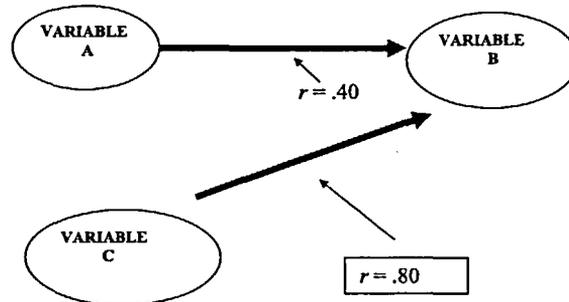
Consider an abstract example, depicted in Figure 1 (A, B, and C), where the bivariate Pearson's r correlation between Variable A and Variable B is .40. Any factor or Variable C that correlates with outcome B in such an example at the level of .80 might seem intuitively to be the basis for an explanation that *competes* with any explanation of Variable B based on Variable A. And in some ways it is, accounting by itself (as far as we can tell on the bivariate level at any rate) for 64% of the variance in Variable B.

FIGURE 1(A)

The bivariate Pearson's r correlation between Variable A and Variable B is .40.



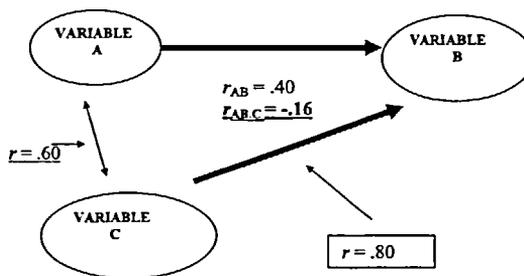
Add a control variable C, whose Pearson's r correlation with Variable B is .80:



WHAT IS THE IMPACT OF THIS CONTROL VARIABLE ON THE RELATIONSHIP BETWEEN VARIABLE A AND VARIABLE B?

FIGURE 1(B)

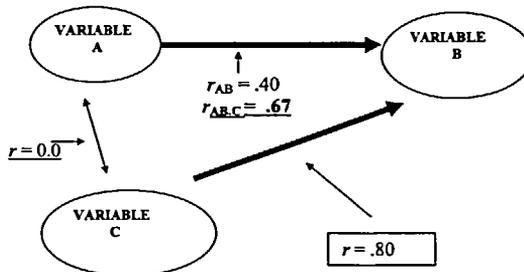
FIRST EXAMPLE: The Pearson's r correlation between Variables A and C is .6



The partial correlation of Variables A and B, controlling for C, or $r_{AB.C}$, is $-.16$. The original positive relationship between Variables A and B is revealed to be entirely spurious. Variable C in this case actually does provide the basis for a “competing” explanation to that based on Variable A.

FIGURE 1(C)

SECOND EXAMPLE: The Pearson's r correlation between Variables A and C is 0.



The partial correlation of Variables A and B, controlling for C, or $r_{AB.C}$ is $.67$. The partial correlation is stronger than the original correlation of $.40$. Variable C does not provide the basis for a “competing” explanation. Rather, it is an complementary explanatory factor, whose introduction into the model as a control serves to reveal a stronger relationship between Variables A and B than is apparent in bivariate analyses.

However, it would be a mistake to go from this intuitive notion that explanations based on Variable *C* *compete* with explanations based on Variable *A* to the conclusion that introducing control Variable *C* will necessarily reveal that the hypothesis positing an impact by Variable *A* on Variable *B* is invalid or fallacious. Whether or not this is the case depends on whether Variable *C* *also* correlates with Variable *A*. In Figure 1(B), Variables *A* and *C* correlate at the .60 level. And controlling for Variable *C* in this case wipes out entirely the original bivariate correlation between Variables *A* and *B*, reducing it from .40 to -.16. However, as depicted in Figure 1(C), if Variables *A* and *C* are not correlated at all, introducing Variable *C* into the model will not eliminate the correlation between Variables *A* and *B*. On the contrary, the control for Variable *C* in this case will relieve Variable *A*, so to speak, of the necessity to account for variation in Variable *B* with which it has no causal connection, because the variation in question in Variable *B* is in fact brought about by Variable *C*. Thus unburdened, as it were, Variable *A* can be shown in fact to correlate quite strongly with Variable *B*, once Variable *C* is controlled for. This is not to say that a control for this kind of *C*, a complementary cause, is uncalled for or undesirable. But it should be clearly distinguished theoretically, and in discussions or explanations of research findings from the process in which control variables are intended to expose the relationship of key interest as spurious.

Even in such cases, however, it can be argued that the hypothesis based on Variable *C* as a main explanatory factor *competes* with one based mainly on Variable *A*, despite the fact that controlling for Variable *C* increases the correlation of Variable *A* with Variable *B*. What the control for Variable *C* does, in this case, in other words, is to increase the goodness of fit between Variables *A* and *B*.

But Variable *C*, and the explanation based on that variable, might be much more important, as would be its relative impact on Variable *B* (as compared to that of Variable *A*). In the context of ordinary least squares techniques, such relative impacts can be evaluated in terms of a comparison of the regression coefficients. Even if the correlation between Variables *A* and *C* is very high, when Variable *B* is controlled for a change in Variable *C* might be associated with much larger changes in Variable *B*, and therefore, arguably, Variable *C* is a much more important explanatory factor than Variable *A*. For maximum likelihood techniques, the analogous gauges or measures of relative impacts focus on odds ratios, or percentage changes in the dependent variable associated with different explanatory variables with additional control variables held at (for example) their mean.

It is often stated that these odds ratios or percentage change figures are assessments of the *substantive* significance of the putative explanatory factors, as opposed to mere *statistical* significance, or “goodness of fit.” The clear implication is that statistical significance and goodness of fit are or should be of relatively minor interest, compared to the *substantive* significance of the various explanatory variables. Substantive significance is equated to the *amount* of change, or the extent of change in probabilities, brought about in the dependent variable apparently by a corresponding change in the independent variable.

Unfortunately, the measures of *substantive* significance in current use in association with maximum likelihood techniques are often quite misleading, tending to give exaggerated impressions of the impacts of the explanatory variables in question. As Gary King (2001, 501) points out with respect to odds ratios, or “relative risks,” “a relative risk of 2 could summarize a change in probability from 1 in a billion to 2 in a billion or from 0.4 to 0.8. In other words, the same relative risk could indicate a result that is substantively irrelevant or vitally important.” Typically even more misleading are reports of “percentage changes.” Oneal and Russett (2001) report an analysis of the impact of various factors on the probability of a militarized disputes within pairs of states for the time period from 1886 to 1992. They suggest (in a footnote) that “the baseline rate for a contiguous pair of states” in their analyses is “.043” (Oneal and Russett 2001, 474). When they report the “substantive effects” of factors such as democracy and trade, they claim that an increase in the former decreases the probability of a dispute by 64 percent, while an increase in the latter reduces the risk of a dispute by 56 percent. If the baseline is 4 percent, one might reasonably wonder, how can any factor reduce the probability of a dispute by 64 percent?

Similarly, the author of an article published in the *Journal of Conflict Resolution* in 2002 focused on developing a model to account for MIDs between 1950 and 1982. More specifically, he focused only on those MIDs resulting in the “use of force,” or the “initiation of interstate war.” MIDs are rare enough. MIDs that escalate at least to the ‘use force’ are even more rare.

Having run an initial analyses and reported which relationships were “significant,” this author declares in typical fashion under the subheading of “Substantive Effects” that “even though it is important to know whether the variables are significant and in what direction they point, of more interest is the actual effect of these variables on the prospects for international peace” (Appendix). By “actual effects,” the author means how much *change* there is in the probability of a MID reaching the “use of force” level in association with changes in the

predictor variables in his model. This article does note that in this data set the "baseline probability of conflict is .004." Since 4 out of every 1000 observations involve MIDs escalated to the "use of force" level, one must logically expect that any percentage reduction or increase in conflict brought about by changes in independent variables must be correspondingly small. However, instead this article reports, for example, that "a one-standard-deviation increase in the capability ratio variable increases the likelihood of conflict by 1,167%"! (emphasis, i.e., the exclamation point, added) (Appendix).

If we are talking probabilities here, then logically speaking an event cannot really be more than 100 percent likely. Or, in other words, strictly speaking, the probability of an event cannot exceed 1.0. What then, can it mean to declare that changes in the independent variable make a MID with use of force 1,167 percent more likely? The answer is that analysts come up with figures such as this by calculating "percentages of percentages," or presenting even microscopic changes in raw probabilities as proportions with those tiny probabilities in the denominator. If one finds that with other variables controlled, the probability of a dyad having a serious dispute is increased to .05068 from the baseline of .004, one can report an increase in the probability of such a dispute of .04688 (On a scale from 0 to 1.) But how much more attention grabbing it is to report instead that increases in the capability ratio augment the probability of serious disputes for dyads by .05068-.004/.004, an increase of 11.67, or 1,167 percent!

In short, by now a standard *accoutrement* of multivariate analyses consists of reports of "substantive significance" in terms of changes in the dependent variable apparently brought about by changes in a given independent variable, with some set of control variables. Unfortunately, these changes are hardly ever reported in terms of simple changes in probabilities on a scale from 0 to 1, or in percentage terms from 0 percent to 100 percent. Instead, they are usually provided in terms of odds ratios, percentages of percentages, and the like. Worse, they are invariably embedded within the context of sets of control variables that are, (in the absence of a theory specifying that this set of variables is uniquely identified as containing all important explanatory factors, and none that are superfluous) ultimately arbitrary. This context rather consistently makes these reports of "substantive significance" not only misleading but in principle of quite limited interest.

GUIDELINE #3: Do not introduce factors as control variables merely on the grounds that they have an impact on the dependent variable.

One cannot review a substantial number of articles dealing with international politics that include multivariate models with control variables without being struck by how brief and cryptic most rationales for the inclusion of control variables are. In many papers, there is no rationale at all. There is nothing more than a heading such as "Control Variables," followed by a list. That is it. An article published in *International Studies Quarterly* in 1999, for example, contains the following one sentence rationale for the control variables that are introduced: "Finally, we include three control variables suggested by the theoretical and empirical literatures" (Appendix). Another paper published in *International Studies Quarterly* in 2000 explains (in one and only one sentence) that "we also include several control variables to account for the effects of various other factors noted in the literature that may also influence the expected duration of a dispute" (Appendix). Finally, the author of an article published in *International Studies Quarterly* in 2001 explains that "Even the most vigorous proponents of an issue-based approach are likely to agree that issue salience and recent interaction over the issue do not tell the entire story; states' decisions are likely to be influenced by additional factors as well as by the issues themselves" (Appendix).

The assumption on which statements like these are based is clear. As long as the factors added *also* have an impact on the outcome phenomenon in question (such as international conflict) then they are legitimate candidates for inclusion as control variables. Furthermore, if it can plausibly be argued that the candidate control variables also have an impact on the outcome phenomenon under examination then nothing more need be said in defense of their inclusion (according to the rationale undergirding these one sentence, or no sentence explanations for the control variables that have been added to multivariate models).

Guideline #3, then, is basically a straightforward corollary of Guidelines #1 and #2. One problem with an implicit or explicit rationale for including control variables in models aimed at evaluating the impact of some key factor that asserts nothing more than that the factors in question are *also* a cause of whatever it is that is being accounted for is that it obscures the distinctions between confounding variables, intervening variables, and alternative causal factors. As we have already pointed out, one of the implications of the distinctions between these categories of control variables is that some factors *also* having an impact

on the dependent variable (namely, intervening variables) should typically *not* be controlled for, even though, and in fact precisely because they also have an impact on the outcome variable.

Furthermore, the reasons for including competing explanatory factors on the one hand, and alternative explanatory factors, on the other, are diametrically opposed to each other. Including competing explanatory factors, (i.e., confounding variables) subjects the hypothesis in question to a more rigorous test than is possible with bivariate analyses. In contrast, including alternative explanatory factors enhances the possibility that the key hypothesis will be provided sustaining evidence by the analysis in question.

There is at least one more reason for not including as control variables any factors that are also related to the outcome variable of interest. It is the reason providing the justification for this particular guideline. A rule of thumb indicating that any factor that is an important cause of, say, international conflict should be included as a control variable, or is at least a serious candidate for a control variable, is obviously far too un-discriminating. Applied with consistency, such a rule will virtually always lead analysts to violate what Christopher Achen (2002, 446) refers to as the "Rule of Three." This rule states that "*a statistical specification with more than three explanatory variables is meaningless*" (emphasis added).

Now, admittedly, Achen (2002, 446) stipulates that this rule applies only when "no formal theory structures the investigation." I would be inclined to be even more generous. It seems to me that it would be permissible to include more than three variables in a multivariate model even if the theory providing justification for each variable in the model were not *formal*, strictly speaking. But in the absence of a theory that stipulates clearly that the variables in the model are the complete set of factors necessary to explain the outcome phenomenon, then I would recommend that all analysts abide by the rule of three (as well as Guideline #3) religiously. (Or at least respectfully.)

But, the discriminating reader might wonder, are not Guideline #3 as well as Achen's rule of three regrettably oblivious to the possible ravages of "omitted variable bias?" King, Keohane and Verba (1994, 172) observe that "the ideal solution [to omitted variable bias] is not merely to collect information on all relevant variables, but explicitly and *simultaneously* to control for all relevant variables." However, in fact, it is in most studies of international conflict, at any rate, quite impossible even to think of, much less generate data on "all relevant variables." So, virtually all multivariate models of international conflict suffer from "omitted variable bias." It is in my view, better

to keep this in mind, and to be appropriately modest and circumspect about the validity of such multivariate models, than it is to add large numbers of control variables in a futile attempt to reduce significantly the huge number of potentially relevant variables that might be left out of any given multivariate analysis. Any set of control variables added to multivariate models merely to reduce the risk of omitted variable bias constitute, in the absence of a fully specified theory, a random and arbitrary portion of those variables whose omission might potentially bias the results. And they are likely to make the results of the multivariate analysis uninterpretable or confusing, as Achen's rule of three suggests.

GUIDELINE #4: Do not control for variables that are related to each other or the key explanatory factor by definition.

This guideline makes the apparently obvious argument that "independent" variables should not be related to each other by definition. Appearances in this case are deceiving. This is not a proposition that is universally recognized as legitimate or valid. At least as early as 1999, for example, Oneal and Russett control for both contiguity, i.e., the sharing of a border between two states and the distance between two states.⁴ A very general, wide-ranging article published in the *Journal of Conflict Resolution* in 2000 reviewing many issues and a multiplicity of analytical techniques utilizes controls for both contiguity and distance in a manner that may well be a significant step in a process leading such controls to become a standard operating procedure.⁵ My reading of several recent papers written for conferences and conventions suggest that it has in fact achieved this status.

An even more problematic example of such a practice can be found in an article published in the *Journal of Conflict Resolution* in 2002. In this research, political similarity is included as a control variable in a model devoted to evaluating the impact of joint democracy on conflict behavior (Appendix). Henderson (2002) includes political similarity (actually he labels it political dissimilarity) in an analysis also aimed at evaluating the impact of joint democracy on conflict behavior. He concludes from the impact of this control that joint democracy has no

⁴ Russett, Oneal and Davis (1998) explicitly avoid utilizing both contiguity and distance as control variables in the same model.

⁵ "We include a measure marking whether (1) or not (0) the two states in the dyad were contiguous on land. We also include the log of distance between two states" (Appendix).

important pacifying impact.⁶

The problem with the practice of including in the same model independent variables that are related to each other by definition is simple. The implicit, virtually explicit, argument made by authors who claim that political similarity is a variable creating a spurious correlation between democracy and peace implies in a straightforward fashion that political similarity is a cause of both democracy and peace. (Otherwise, it could not be claimed that political similarity is a confounding variable.) This in turn suggests that democracy has no independent, significant (in either the statistical or substantive sense) impact on peace.

But it is illogical to argue that political similarity has a causal impact on democracy, as one must argue in order to suggest that political similarity constitutes a confounding variable rendering the observed, oft-reported, and continuously commented on correlation between democracy and peace spurious. It is illogical because joint democracy *is*, by definition, a kind of political similarity. Being politically similar does not lead pairs of states to become jointly democratic. Political similarity and joint democracy correlate not as a result of a causal connection but because of a definitional connection. This by itself disqualifies political similarity as a candidate for a confounding variable with respect to the relationship between democracy and peace.⁷

In other words, the task confronting critics who wish to dispute a claim that A is a cause of B, when it is clearly established that A consistently correlates with B, is to find some factor, C, that has an impact on both variables A and B, thus bringing about the correlation that is pointed to as evidence of a causal connection between A and B. If a particular critic of the idea that A is a cause of B comes up instead with a factor that is statistically related to A because of definitional overlap, rather than by virtue of a causal connection, that critic has violated one of the clear (if, up to this point at least, largely implicit) rules of the game, as it were. The correlation in the model in that case does not correspond to some empirical connection out there in the "real world."

⁶ "The inclusion of the political dissimilarity variable reduces the impact of joint democracy on conflict, and the inclusion of trade interdependence reduces this impact to insignificance" (Henderson 2002: 44).

⁷ This is not to say that it is illegitimate to argue that it is political similarity that causes peace and that the correlation between joint democracy and peace is merely a reflection of that fact. The theoretical argument that political similarity causes peace is broader and more comprehensive than the argument that joint democracy causes peace. The former argument subsumes the latter argument, in other words. But including a control for political similarity in a model evaluating the impact of joint democracy on peace is not an appropriate way to establish the validity of the argument that the impact of political similarity is more comprehensive and fundamental than that of joint democracy.

It reflects, instead, a definitional connection in the artificial model constructed by the social scientist. Therefore, under such circumstances, the fact that the original relationship between variables A and B disappears if C is controlled for, is a statistical artifact brought about by the (confused) manipulation of concepts, rather than evidence of a spurious relationship in the "real world."

Such a criticism clearly applies if political similarity is suggested as a confounding variable with respect to the relationship between democracy and peace. But what is the harm in controlling both for contiguity, as well as distance between capital cities, if one is analyzing pairs of states in the international system?⁸ Contiguity and distance will be correlated with each other, but the correlation is not terribly strong. Contiguity is a dummy variable, usually, while distance is a ratio level variable with a wide range of values; this alone puts a rather severe limit on the extent to which the two variables will correlate. One can easily conjure up circumstances in which contiguity might have different causal impacts on relationships within a pair of states than distance between their capital cities. Furthermore, there are some states that are contiguous, even though their capital cities are relatively far apart (such as Russia and China, for example, or Mexico and the United States), while other pairs of states are not contiguous even though their capital cities are relatively close together. (Uruguay and Paraguay, or Hungary and the Czech Republic come to mind.) So, might not the inclusion in a model of both contiguity and distance as control variables justly be categorized as admirable thoroughness?

I think not, because of the cost one pays for such a strategy in the form of a decreased correspondence between causal processes in the real world and causal processes as they are reflected, depicted, or described in multivariate social scientific models. In order to come to conclusions as valid as possible about the "real world," the causal processes as depicted in a model should correspond as closely as possible to causal processes among the factors of interest as they interrelate in the "real world." Contiguity and distance between capital cities will correlate with each other in a social scientific model even though in the real world there is no causal connection between them. States that are contiguous will, on average, tend to have capital cities that are

⁸ Additional evidence of what may well be a trend toward increased reliance on such a strategy can be found in Reiter and Stam (2002, 51) who utilize both contiguity and political relevance as control variables in the same model. ("Political relevance" involves "contiguity" by definition.)

relatively close together.⁹ States with capital cities separated by great distances will, on average, be less likely to be contiguous. But this is because contiguity and distance between capital cities are both aspects of the geographic relationship between pairs of states. Contiguity does not cause capital cities to be close together, nor do widely separated capital cities prevent states from being contiguous. Contiguity and capital cities close together are both different aspects of the geographic relationship between the pair of states in question. They are in fact both related by definition to geographic proximity, and so related also by definition to each other. The correlation between contiguity and proximate capital cities is brought about by these definitional connections, rather than by empirical, causal connections. The model containing both factors as control variables creates a background for the examination of empirical connections between other variables that is artifactually different from the "real world" background in which the causal processes in question take place. The resulting relationships between the key independent and dependent variables in question are correspondingly distorted, bearing an even more uncertain relationship than is customary to their relationships in the "real world."

Running two different analyses, one with contiguity as the control for geographic proximity and another with distance between capital cities as a control variable, is a perfectly legitimate strategy. Selecting the model with the "best" results from the point of view of the analyst is also justified, especially if the analyst can come up with some plausible theoretical conjecture that might account for the differences in the two analyses. This strategy is clearly preferable, in my view, to one involving the inclusion of both contiguity and distance as control variables in the same model or analysis. That strategy, to repeat, in effect provides misleading information to the analytical technique in question, suggesting that there is an empirical connection between the variables, when in fact the connection is definitional instead.

There is at least one circumstance, however, when I believe that including control variables that are related to each other by definition, may be justified. If an analyst is interested in the possibility of interaction effects between independent variables X_1 and X_2 , she will typically

⁹ The argument that this is a definitional connection is, admittedly, not intuitively so obvious. Lemke and Reed (2001, 138) assert that "as the distance between capitals decreases, the two states are increasingly likely to be contiguous" in a fashion making it apparent that they see this as at least to some degree as a falsifiable proposition or hypothesis. Their doubts about including contiguity as a control variable in a model of conflict among "politically relevant" states (many of whom are contiguous by definition) is more in accord with the spirit of Guideline #4.

include in a multivariate model not only the multiplicative term $X_1 * X_2$, but also the individual variables X_1 and X_2 separately. In such a model, obviously X_1 and $X_1 * X_2$, as well as X_2 and $X_1 * X_2$, are independent variables in the same equation that are related to each other by definition.

In fact, in the past, such models have evoked some suspicion (Althausser 1971; Smith and Sasaki 1979), in part because the definitional connection typically creates such high levels of multicollinearity between the interaction term and its separate components (Blalock 1964; Lewis-Beck 1980). There is also a certain amount of awkwardness purely on the conceptual (as opposed to the statistical) level involved in the idea of "holding constant" X_1 , while allowing $X_1 * X_2$ to vary. Nevertheless, in *this* case it does seem to me that it is legitimate to include independent variables in the same multivariate model that are related to each other by definition. In this situation, all three coefficients together can serve as the basis for interpreting the relationship between X_1 and X_2 , on the one hand, and the dependent variable, on the other, as well as the relationship between the dependent variable and the interaction of X_1 and X_2 . There is no implicit argument in such a model or implied in this selection of independent and control variables that the correlation between X_1 and X_2 , and $X_1 * X_2$ is the result of an empirical, causal connection between the separate components, on the one hand, and the interaction term, on the other. Rather, it is understood that all three terms are included in the model for the purpose of creating three separate coefficients on which interpretations of the relationships between just two variables (in their independent and their interactive forms), and the dependent variable can be based. In fact, what this means is that this practice, specifically and technically speaking does not really violate Guideline #4. It does not involve including different independent variables that are related to each other by definition. It involves instead the inclusion of the *same* independent variables in different functional forms for the purpose of evaluating their individual and combined impacts (Friedrich 1982).

However, in general, any analyst who includes in a statistical model variables on both sides of the equation that are related to each other by *definition*, and then portrays the resulting statistical association between those variables as evidence in favor of an hypothesized causal connection would be justifiably be accused of presenting an illogical argument. According to Guideline #4, definitional connections between variables on the same side of the equation are equally objectionable, and for reasons that are analogous.

GUIDELINE #5: Control for possible differences between across space and over time relationships.

In one of the more visible and authoritative statements on an issue to which little attention has been paid recently, Green, Kim and Yoon (2001, 458) declare in a straightforward manner that:

Cross-sectional inference is not inherently invalid, but it cannot be considered reliable if contradicted by time-series analysis...If...we are interested in estimating the structural parameters that govern cause and effect...time series analysis and cross-sectional analysis should, in principle, give the same answers.

Intriguingly, one of the main targets of the critique mounted by Green, Kim and Yoon (2001) explicitly agree with their critics on this particular point. According to Oneal and Russett (2001, 481), "The analyses of time-series and cross-sectional data should give us the same answers, as Green, Kim, and Yoon note."

The idea that across space analyses should produce the same results as over time analyses, and that if they do not it is the across space results that are somehow suspect and inferior, is quite longstanding in the field. For example, Stimson (1985, 917) asserted almost two decades ago that "analysts of cross-sections, insofar as they are concerned with causation at all, typically observe covariation presumed to be produced by unobserved causal processes operating at some time before the data were gathered. Time series analysts typically wish to model a causal process captured by the longitudinal data." The presumed inferiority of cross-sectional analyses here is quite clear. If one is lucky, Stimson seems to be saying, cross-sectional analysis will reveal the impact or character of causal processes operating *overtime*. But longitudinal analyses provide a much more direct approach to causal linkages, in Stimson's opinion. Similarly, William Berry (1993, 22-23) asserts that "regression models can be *cross-sectional...or time series....*In most cases, regression coefficients having dynamic interpretations tend to be more interesting.... This is not to say that cross-sectional relationships are completely uninteresting to social scientists." Not completely uninteresting, according to Berry, but of value only as a possible reinforcement to longitudinal results that might not be obtainable because of the absence of data.

The origins of such ideas about the relationship between and the relative worth of across space and over time analyses are not entirely obvious. One of them seems to involve the fact that temporal order is an essential part of what social scientists (and people in general) mean by the notion of "cause." In other words, any argument that A is a cause of

B must establish that A came first, and was followed by B, rather than vice versa. It just seems natural and logical to assume that over time analyses will be better suited to "get at" this fundamentally important temporal ordering issue when hypothesized causal linkages are being investigated.

But this seemingly natural and logical assumption is in fact misleading, if not downright erroneous. One can establish temporal order within the context of across-space analyses. In fact, in at least some circumstances it is easier to establish temporal order between variables in across space analyses than it is in over time analyses. In most cases, one can observe, in cross-sectional analyses, the independent variable at t_{-1} , and observe the dependent variable at t_0 . In other cases, there can in fact be no doubt about the temporal ordering of variables involved in cross-sectional analyses. If one correlates gender or ethnic identity with vote, for example, or contiguity with conflict propensity for pairs of states, is there any doubt about which comes first? (Voting behavior, in other words, does not occur prior to, or have an impact on, ethnic identity or gender, and conflict behavior could not precede nor have an impact on the geographic locations of states.) Rarely, in fact, can one rule out an endogeneity problem with such confidence as in those examples, in the context of a longitudinal, or over time analysis.

Perhaps another basis for prejudice against cross sectional analyses has its origin in "modernization" theory. Lipset (1959, 75) presented an elaborate argument to the effect that "democracy is related to the state of economic development." He buttressed that argument with a series of straightforward, cross sectional comparisons of relatively wealthy with relatively poor societies, finding, of course, that the wealthier societies were by almost any measure more democratic. In ensuing years, the useful observation was made that cross-sectional comparisons of poor societies with rich societies do not necessarily provide a sound basis for inferring what relationships between economic growth and regime transitions over time will be like within poor states. In other words, there is no logical guarantee, at least, that because rich societies are currently more democratic on average than poor societies, currently poor societies will become more democratic when they get wealthier. This sequence of events in the field seems to have confirmed for some analysts in the field the basic inferiority of cross-sectional analyses, and the assumption that for the most part they are relatively poor substitutes for more relevant, and more theoretically interesting over time analyses.

On one point those who discuss the relationship between over time and across space analyses are clearly cogent. Over time analyses and across space analyses of the same variables will not necessarily lead to

the same, or even similar results. Therefore, it is risky at best to infer what over time relationship might be on the basis of across space analyses, and vice versa. But to go from this valid point regarding the relative independence of the results produced by over time and across space analyses to the conclusion that across space analyses are somehow inherently inferior to or less interesting than over time analyses is unwarranted. Causal relationships can operate over time, and across space. Furthermore, they may operate in a different ways in both dimensions. In other words, the relationship between Variables A and B may be positive over time but negative across space. There is no reason in principle why contrasting relationships could not be posited within logically consistent and plausible theoretical frameworks.

Furthermore, it so happens that some of the most robust, and theoretically important relationships within international politics and within political science, generally speaking, happen to be cross-sectional rather than longitudinal or over time relationships. This is not the place to speculate in detail about why this might be the case. But one obvious reason is that many important variables are "time invariant." In other words, they vary not at all, or only rarely, over time. Think about studies of voting behavior in the United States (or elsewhere, for that matter). Some of the most important predictors of, say, votes in presidential elections are such factors as ethnic identity or gender. One will obviously not find that across space analyses will produce the same results as over time analyses involving these variables, because with *very* rare exceptions the ethnic identity or the gender of voters will not vary over time. Even such independent variables, in the case of voters, as social class or party identification, tend not to vary at all over time, or at least not with sufficient intensity or regularity to produce significant relationships with variations in voting behavior over time. But across space, such factors as ethnic identity, gender, social class, and certainly party identification, have very important relationships with voting behavior.

In international politics, certainly one of the most consistent and powerful predictors of conflict (MIDs or wars) within pairs of states is contiguity, or distance. Pairs of states that are close together are much more likely to become involved in international conflict than pairs of states that are far apart. And this is not a trivial correlation. It is important to the investigation, for example, of the role of territorial issues in the escalation of interstate conflict (Vasquez 1993, 1995; Senese 1997) Contiguity and/or distance have been included as predictor and control variables in countless pooled time series, cross-sectional analyses of the incidence of interstate conflict in the last twenty years or so. Every

analyst who has presented such a model over that time period has assumed, at least implicitly, that the relationship between contiguity and conflict across space is the same as that between contiguity and conflict over time. "Typically, authors do not distinguish between- and within-cluster effects...and so implicitly assume that these effects are the same" (Neuhaus and Kalbfleisch 1998, 639). But obviously, the across space relationship between contiguity or distance, on the one hand, and the conflict propensity of pairs of states, on the other, is not the same as the relationship between those two factors over time, for the simple but profoundly important reason that there is virtually no variation over time in their geographic relationship for the vast majority of pairs of states included in any empirical analysis. And, according to Neuhaus and Kalbfleisch (1998, 639), "models that incorrectly assume common effects lead to very misleading assessments."

So, when in the process of constructing multivariate models, one should not assume that that across space and over time relationships are the same. They may well be strikingly different. Zorn (2001), for example, provides intriguing evidence on this issue. He reanalyzes data utilized by Beck, Katz, and Tucker (1998), who themselves focused on data originally relied on by Oneal and Russett (1997). These data focus on "politically relevant" dyads from the years 1950 to 1985. The dependent variable is the occurrence of a militarized interstate dispute. The model includes as predictor variables economic growth, alliance ties, contiguity, capability ratios, regime type, and trade ties. In other words, it is a multivariate model addressing international conflict of a type and construction published dozens of times in the last twenty years or so.

Except that Zorn (2001) does not assume that each of those variables relate to interstate conflict in the same fashion over time, and across space. He focuses in particular on possible differences in the over time and across space impacts on interstate conflict of regime type and interstate trading ties. In order to capture across space variation in isolation, he substitutes for each observation of a dyad's annual regime type and trading ties its mean values on those variables for the entire time period. (In other words, he controls for over time variation by eliminating it altogether.) The variable that captures only within dyad, over time variation consists of the raw scores each dyad receives annually, minus its mean score for the entire time period. What Zorn finds is that the relationship between trade and conflict, as well as that between joint democracy and conflict, differ over time, and across space. "The between- and within-dyad effects of trade on conflict run in opposite directions....[D]yads which trade more are less likely to enter

into conflicts than those which trade less. At the same time, exceptionally high (low) levels of within-dyad trade are associated with an increased (decreased) probability of conflict" (Zorn 2001, 439-440). Zorn also reports that there is no significant relationship between variations in joint democracy within dyads and the incidence of conflict they experience over time. "The results here support the proposition that 'If dyad A is more democratic than dyad B, then dyad A is less likely to engage in conflict.' Only very marginally supported, however, is the hypothesis that within-dyad increases in democracy levels lead to more peaceful behavior..." (Zorn 2001, 439).

However, Zorn's findings regarding different relationships between trade and democracy, and interstate conflict over time and across space are embedded in a typically lengthy, complex set of control variables. That list is sufficiently complex that one might wonder whether the findings could somehow be a result of that potentially confusing multivariate background. Tables 1 and 2 below report the initial results of an attempt to address that possibility and to replicate and extend Zorn's analyses in certain other respects. These analyses focus on all dyads, rather than just politically relevant dyads, and for a somewhat longer period of time, from 1950 to 1992. Across space variation in the independent variables is captured in manner somewhat different than that utilized by Zorn. He, recall, simply replaces raw scores for each annual dyadic observation with the mean score on each variable in question for that dyad for the entire time period. This might obscure important variation, and variation that could have an important across space impact. In these analyses across space variation is indexed in a manner somewhat more closely analogous to the manner in which over time variation is measured. As in Zorn's analyses, over time variation in Tables 1 and 2 is measured by subtracting from each dyad's annual raw scores the mean score for that dyad for the entire time period. In short, only variations around that mean remain. In our analyses, across space variation is captured by subtracting from each raw score the mean score for the international system at that given point in time.

Table 1 focuses on a maximum-likelihood logit estimation (with Stata®) of the relationship between trade and conflict (or MID involvement), with only contiguity (as well as peace years and splines) introduced as control variables. What it shows is, as was reported by Zorn (2001), that the relationship between trade and conflict is different depending on whether one focuses on the across space or the over time relationship. Across space, the relationship between trade and interstate conflict is negative, and significant. Over time, increased trade is associated with increased conflict, and at a significant level. Whether or

TABLE 1

RELATIONSHIP BETWEEN LEVELS OF INTRADYADIC TRADE ACROSS SPACE AND OVER TIME, AND THE PROBABILITY OF MILITARIZED DISPUTES, 1950-1992#

(AS= ACROSS SPACE)
(OT=OVER TIME)

VARIABLES	COEFFICIENTS	Z-SCORES	SIGNIFICANCE
CONTIGUITY	3.53	47.69	0.000
TRADE(AS)	-34.62	-4.15	0.000
TRADE(OT)	48.85	4.90	0.000

N=278,967

#Peace Years and Splines not shown. Those utilized in this and subsequent analyses were generated with the help of Tucker (1999). The data set relied on to produce this and subsequent tables was assembled with EUGene (Bennett and Stam 2000).

TABLE 2

RELATIONSHIP BETWEEN JOINT DEMOCRACY ACROSS SPACE AND OVER TIME, AND THE PROBABILITY OF MILITARIZED DISPUTES, 1950-1992#

(AS= ACROSS SPACE)
(OT=OVER TIME)

VARIABLES	COEFFICIENTS	Z-SCORES	SIGNIFICANCE
CONTIGUITY	3.48	49.11	0.000
DEMOC. (AS)	-.043	-5.48	0.000
DEMOC. (OT)	.01	.70	0.482

N=278,967

#Peace Years and Splines not shown.

not these contrasting relationships over time and across space might help to account for substantial inconsistencies in findings regarding the trade-conflict relationship reported in prior research (as in, for example, Barbieri 2002; Russett and Oneal 2001), is an intriguing possibility that may deserve further investigation.

Table 2 deals with a similar, simple analysis focusing on the over time and across space relationships between democracy and conflict for all dyads from 1950 to 1992, with only contiguity, peace years and splines as control variables. Again, in results mirroring Zorn's, these analyses suggest that while there is a significant negative relationship

between democracy and conflict across space, there appears to be no pacifying impact of democracy over time. (In fact, the relationship is positive, though not significant). Table 2, in conjunction with results reported by Zorn (2001), imply that the democratic peace proposition to which so much attention has been paid in recent decades (e.g., Rummel 1979, Bueno de Mesquita and Lalman 1992; Russett 1993; Ray 1995; Russett and Oneal 2001), may only be valid within a cross-sectional context. If joint democracy does exert an important pacifying impact over time, regime type and interstate conflict within pairs of states may not vary over time with sufficient intensity and regularity to produce consistent empirical evidence of that impact.

TABLE 3
RELATIONSHIPS BETWEEN SEVERAL "STANDARD" PREDICTORS
ACROSS SPACE AND OVER TIME, AND THE PROBABILITY OF
MILITARIZED DISPUTES, 1950-1992#

(AS= ACROSS SPACE)
 (OT=OVER TIME)

VARIABLES	COEFFICIENTS	Z-SCORES	SIGNIFICANCE
CONTIGUITY	2.82	26.19	0.000
DISTANCE	-.0002	-9.83	0.000
TRADE(AS)	-77.40	-4.80	0.000
TRADE(OT)	89.24	5.07	0.000
CAP. RATIO (AS)	.430	2.99	0.003
CAP. RATIO (OT)	-.088	-0.21	0.833
MAJMIN	2.03	19.00	0.000
MAJMAJ	4.00	11.70	0.000
ALL.TIES (AS)	-.218	-2.26	0.024
ALL.TIES (OT)	.733	3.10	0.002
DEMOC. (AS)	-.041	-4.51	0.000
DEMOC. (OT)	.011	0.64	0.522

N=278,967

#Peace Years and Splines not shown.

Table 3 subjects the relationships reported in Tables 1 and 2 to an evaluation within the context of a set of control variables whose only virtue (in my opinion) is that it is quite customary in character. The predictor variables in Table 3 are contiguity, distance between capital cities, trade ties, capability ratios, a dummy variable equal to one if the dyad contains a major power and a minor power, and 0 otherwise, another dummy variable equaling 1 if both states in a dyad are major powers, and 0 otherwise, alliance ties, and regime type. The dependent variable is again involvement in militarized disputes. It is perhaps (I rather hope) unnecessary to point out that the model relied on to gener-

ate the results in Table 3 violates most of the guidelines for constructing multivariate models discussed in this paper. One point of the analyses reported in Table 3 is to demonstrate that the contrasting and significant across space and over time relationships between trade and democracy, on the one hand, and involvement in MIDs, on the other, can withstand the onslaught, as it were, of a typically lengthy (and arbitrary) list of control variables. I am adhering here, in other words, to a custom I would in the long run hope to discourage. But another point supported by Table 3 is that trade and democracy are not the only variables that may relate differently to interstate conflict over time, and across space. That table fails to show different spatial and temporal relationships between contiguity and distance, on the one hand, and dispute involvement on the other, only because neither contiguity nor distance vary over time; therefore, it is impossible to estimate the over time relationship between either measure of geographic proximity, and conflict. Table 3 does show that capability ratios (which in this case equal 1 when two states in dyad are equal in power, and less otherwise) are positively related to conflict propensity across space (i.e., more equal states are more likely to become involved in interstate disputes), but the over time relationship between capability ratios and interstate conflict is not even close to being significant. And pairs of states that are allied to each other are significantly less likely to become involved in disputes than states that are not allied to each other, at any given point in time. But over time (for some reason, within the context, at least, of this particular, arbitrary and peculiar, even if typical set of control variables), states are more likely to become involved in disputes with each other when they are allied than when they are not allied.

In short, analyses by Zorn (2001) and those reported in Tables 1, 2 and 3 above, provide reasonably substantial evidence that over time and across space relationships between many "standard" predictors and interstate conflict are consistently inconsistent. As Neuhaus and Kalbfleisch (1998, 644) conclude, "When between- and within-cluster covariate effects are different, models that assume that these effects are the same do not provide estimates of any substantive interest." I do not believe that all those results based on multivariate pooled cross-sectional, time series analyses of interstate conflict published in the last twenty years or so based on the dubious assumption that over time and across space relationships are the same are totally without substantive interest. I would suggest that such analyses could be improved upon by taking into account possible differences in across space and over time relationships.

CONCLUSION

TABLE 4

WHAT SHOULD BE CONTROLLED FOR?:
FIVE GUIDELINES

1. *Do not control for intervening variables.*
2. *Distinguish between complementary and competing explanatory factors.*
3. *Do not introduce factors as control variables merely on the grounds that they have an impact on the dependent variable.*
4. *Do not control for variables that are related to each other or the key explanatory factor by definition.*
5. *Control for possible differences between across space and over time relationships.*

For various reasons, systematic empirical research into the causes of interstate conflict has come to rely on multivariate models constructed in such a fashion that the results of statistical evaluations of their validity are commonly quite difficult to interpret. Many of these problems originate in norms or customary practices that have come to determine quite consistently the selection of control variables that are added to models whose main purpose is to increase the understanding of the impact of one key factor, such as joint democracy, power transitions, or international trade. In addition, many analyses utilizing pooled cross-section, time series analytical techniques have implicitly accepted the assumption that across space and over time relationships are essentially identical to each other. This paper calls for a modification or elimination of many of these norms or practices along the lines suggested in Table 4. Were these guidelines to be adopted widely in future research efforts, the results of empirical evaluations of multivariate models might be more readily interpretable and understood. Ideally, research on international conflict would then be more cumulative, as well as more productive of valuable insights into those processes leading to interstate conflict and war.

REFERENCES

- Achen, Christopher H. 2002. "Toward a New Political Methodology: Microfoundations and ART." *Annual Review of Political Science* 5: 423-450.
- Althaus, Paul D. 1977. "Multicollinearity and Non-additive Regression Models." in Hubert M. Blalock (ed.) *Causal Models in the Social Sciences*. Chicago: Aldine-Atherton.
- Barbieri, Katherine. 2002. *The Liberal Illusion: Does Trade Promote Peace?* Ann Arbor, MI: University of Michigan Press.
- Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42: 1260-1288.
- Bennett, D. Scott, and Allan C. Stam. 2000. "EUGene: A Conceptual Manual." *International Interactions* 26: 179-204.
- Berry, William. 1993. *Understanding Regression Assumptions*. Beverly Hills, CA: Sage Publications.
- Blalock, Hubert M. 1964. *Causal Inferences in Nonexperimental Research*. Chapel Hill, NC: University of North Carolina Press.
- Bueno de Mesquita, Bruce, and David Lalman. 1992. *War and Reason*. New Haven, CT: Yale University Press.
- Friedrich, Robert J. 1982. "In Defense of Multiplicative Terms in Multiple Regression Equations." *American Journal of Political Science* 26: 797-833.
- Green, Donald P., Soo Yeon Kim, and David H. Yoon. 2001. "Dirty Pool." *International Organization* 55: 441-468.
- Henderson, Errol A. 2002. *Democracy and War: The End of an Illusion?* Boulder, CO: Lynne Rienner Publishers.
- King, Gary. 2001. "Proper Nouns and Methodological Propriety: Pooling Dyads in International Relations Data." *International Organization* 55: 497-507.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, N.J.: Princeton University Press.
- Lemke, Douglas, and William Reed. 2001. "The Relevance of Politically Relevant Dyads." *Journal of Conflict Resolution* 45: 126-144.
- Lewis-Beck, Michael S. 1980. *Applied Regression: An Introduction*. Beverly Hills, CA: Sage Publications.
- Lipset, Seymour Martin. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53: 69-105.
- Neuhaus, J.M., and J.D. Kalbfleisch. 1998. "Between- and Within-

Cluster Covariate Effects in the Analysis of Clustered Data.”

Biometrics 54: 638-645.

- Oneal, John R., and Bruce Russett. 1997. “The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950-1985.” *International Studies Quarterly* 41: 267-294.
- Oneal, John R., and Bruce Russett. 1999. “The Kantian Peace: The Pacific Benefits of Democracy, Interdependence, and International Organizations, 1885-1992.” *World Politics* 52: 1-37.
- Oneal, John R., and Bruce Russett. 2001. “Clear and Clean: The Fixed Effects of Liberal Peace.” *International Organization* 55: 469-485.
- Ray, James Lee. 1995. *Democracy and International Conflict*. Columbia, SC: University of South Carolina Press.
- Ray, James Lee. 2000. “On the Level(s): Does Democracy Correlate with Peace?” In John A. Vasquez (ed.) *What Do We Know About War?* Lanham, MD: Rowman and Littlefield Publishers.
- Reiter, Dan, and Allan C. Stam. 2002. *Democracies at War*. Princeton: Princeton University Press.
- Rummel, R. J. 1979. *Understanding Conflict and War: Volume 4, War, Power, and Peace*. Beverly Hills: Sage Publications.
- Russett, Bruce. 1993. *Grasping the Democratic Peace*. Princeton: Princeton University Press.
- Russett, Bruce, and John R. Oneal. 2001. *Triangulating Peace: Democracy, Interdependence, and International Organizations*. New York: Norton.
- Russett, Bruce, John R. Oneal, and David R. Davis. 1998. “The Third Leg of the Kantian Tripod for Peace: International Organizations and Militarized Disputes, 1950-1985.” *International Organization* 52: 441-468.
- Senese, Paul D. 1997. “Dispute to War: The Conditional Importance of Territorial Issue Stakes and Geographic Proximity.” Paper presented to the International Studies Association, Toronto.
- Smith, Kent W. and M.S. Sasaki. 1979. “Decreasing Multicollinearity: A Method for Models with Multiplicative Functions.” *Sociological Methods and Research* 8: 35-56.
- Stimson, James A. 1985. “Regression in Space and Time: A Statistical Essay.” *American Journal of Political Science* 29: 914-947.
- Tucker, Richard 1999. BTSCS: A Binary Time-Series—Cross-Section Data Analysis Utility. Version 4.0.4. <http://www.vanderbilt.edu/~rtucker/programs/btscs>
- Vasquez, John A. 1993. *The War Puzzle*. Cambridge: Cambridge University Press.
- Vasquez, John A. 1995. “Why do Neighbors Fight: Proximity,

Interaction, or Territoriality?" *Journal of Peace Research* 32: 277-293.

Zorn, Christopher. 2001. "Estimating Between- and Within-Cluster Covariate Effects, With an Application to Models of International Disputes." *International Interactions* 27: 433-445.