Hypothesis Testing and Multiplicative Interaction Terms

Bear F. Braumoeller

Abstract When a statistical equation incorporates a multiplicative term in an attempt to model interaction effects, the statistical significance of the lower-order coefficients is largely useless for the typical purposes of hypothesis testing. This fact remains largely unappreciated in political science, however. This brief article explains this point, provides examples, and offers some suggestions for more meaningful interpretation.

Despite the remarkable successes of the subfield of political methodology during the past decade or more, a perusal of the applied political science literature gives the impression that the focus has been on running, or even flying, when the fundamentals of walking have yet to be made clear. Nowhere is this fact more apparent than in the case of the humble interaction term.

Political scientists are all familiar with research that tests models such as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon \tag{1}$$

that are used to assess whether or not an interactive relationship between X_1 and X_2 can be said to exist. The functional form may be something considerably more complex than basic multivariate regression, but the goal is the same. Researchers make claims of the following nature:

- β_1 is statistically significant; therefore, $H_1: \beta_1 \neq 0$ cannot be rejected, and the theory that relates X_1 to Y passes this test.
- β_2 is statistically significant; therefore, $H_2: \beta_2 \neq 0$ cannot be rejected, and the theory that relates X_2 to Y passes this test.
- β_{12} is statistically significant; therefore, $H_3: \beta_{12} \neq 0$ cannot be rejected, and the theory that relates the combination of X_1 and X_2 to Y passes this test.

Unfortunately, of these three, only one is a legitimate conclusion based on the results of such a test.

I am grateful to Tim McDaniel, Anne Sartori, and Beth Simmons for comments on a previous draft.

International Organization 58, Fall 2004, pp. 807–820
© 2004 by The IO Foundation.
DOI: 10.1017/S0020818304040251

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

TABLE 1. *The effect of recoding* X_2

Coefficient	X_2	X_2^*
β_0	−0.0217	−1.0086
	(0.1313)	(0.2147)
β_1	1.2983**	0.4155
	(0.2188)	(0.3587)
β_2	2.4674**	2.4674**
	(0.2250)	(0.2250)
β_{12}	2.2070**	2.2070**
	(0.3785)	(0.3785)

Note: Parameters are regression coefficients. Standard errors are in parentheses.
* significant at 0.05 level.
** significant at 0.01 level.

The most dramatic way to illustrate this point is to demonstrate that the coefficients β_1 and β_2, as well as their levels of significance, can be manipulated via simple additive transformations of the data. For the sake of illustration, I generated a data set of 1,000 observations using the data-generating process in equation (1).[1] The first column of Table 1 contains regression coefficients describing the data so generated; the second column contains regression coefficients for the same data when X_2 is recoded as $X_2^* = X_2 + 0.4$ (standard errors in parentheses). If one were to interpret these results as they are interpreted above, one would have to conclude that $H_1: \beta_1 \neq 0$ cannot be rejected in the first test but can be rejected in the second and, therefore, that the theory that relates X_1 to Y passes the first test but not the second.

Simple algebra suffices to provide the conditions under which this result holds.[2] Starting with the basic regression equation and adding an arbitrary constant c to X_2 to create X_2^*,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon \tag{2}$$

$$= \beta_0 + \beta_1 x_1 + \beta_2 (x_2^* - c) + \beta_{12} x_1 (x_2^* - c) + \varepsilon \tag{3}$$

$$= (\beta_0 - \beta_2 c) + (\beta_1 - \beta_{12} c) x_1 + \beta_2 x_2^* + \beta_{12} x_1 x_2^* + \varepsilon \tag{4}$$

1. The independent variables consist of 1,000 random draws from a uniform distribution on the unit interval, the error term consists of 1,000 random draws from a Normal (0,1) distribution, and $y = 0.2 + 1x_1 + 2x_2 + 3x_1 x_2 + \varepsilon$.
2. See, for example, Allison 1977.

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

Therefore, as long as $\beta_{12} \neq 0$, the value of β_1 will change if an arbitrary constant is added to X_2.[3] As long as there is an interaction effect, the values of the lower-order coefficients β_1 and β_2 can be manipulated in this fashion.

The Issue

The reason that the relationships between the individual *X*s and *Y* can be manipulated apparently with such ease is simple: the results of the test are not being interpreted correctly. β_1 captures the impact of X_1 on Y when $X_2 = 0$ (and vice-versa),[4] not the impact of X_1 on Y in general. Because interactive relationships imply that the impact of X_1 on Y varies depending on the level of X_2, the idea of "the impact of X_1 on Y in general" is in fact a meaningless one. Nevertheless, even highly respected scholars continue to interpret lower-order interaction-term coefficients as if they were ordinary coefficients in a strictly additive model.

Such an interpretation is erroneous. If β_1 is statistically significant, it is only reasonable to conclude that $H_1: \beta_1 \neq 0$ cannot be rejected when $X_2 = 0$. The hypothesis may or may not be supported at other levels of X_2. Unless the hypothesis makes some claim of the (highly unusual) form

$$H_1: \beta_1 \neq 0 \text{ when } X_2 = 0,$$

this information is of little immediate use in hypothesis testing.

Moreover, the conclusion that $H_1: \beta_1 \neq 0$ cannot be rejected when $X_2 = 0$ is especially unhelpful if the range of the data, or of the relevant cases, does not include zero. If X_2 were gross national product (GNP), for example, β_1 would describe the estimated impact of X_1 on Y when GNP = 0. In short, it would tell one, literally, nothing. Concluding that a statistically significant relationship between X_1 and Y exists based on such information is simply incorrect.

For these reasons, statements about the statistical significance of β_1 and β_2, rather than being statements about the nature of the political world, at best represent statements about reality that only apply to a subset of the cases. That subset is typically quite small and there is no reason, *a priori*, to believe that it is representative of the rest—in fact, the presence of an interactive relationship guarantees that it will not be. These statements surely cannot help a researcher to evaluate hypotheses relating X_1 and X_2 to Y in general.[5]

3. As long as $\beta_2 \neq 0$, the value of β_0 will change as well.

4. Friedrich makes this point, as do numerous methods textbooks, but its implications for hypothesis testing remain underappreciated. Friedrich 1982, 804.

5. It should also be noted that nonlinearities in the relationship between X_1 (or X_2) and Y, if not modeled explicitly, could produce as an artifact a significant β_{12} if the two independent variables are correlated; I am grateful to an anonymous reviewer for pointing out this possibility.

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

Finally, it is worth noting that the interpretation of coefficients in models with higher-order interactions is more convoluted still. Given the equation

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{23} x_2 x_3 + \beta_{13} x_1 x_3 + \beta_{123} x_1 x_2 x_3,$$

the meaning of the coefficients is even more restricted, and their applicability to hypothesis tests is even more tenuous, than in the two-variable case:

- β_{123} describes the impact of a joint increase of X_1, X_2, and X_3 on Y.
- All other coefficients reflect the singular or joint impact of the independent variables to which their subscripts correspond on Y when all other independent variables are equal to zero.

So, for example, β_3 describes the impact of an increase in X_3 on Y when $X_1 = X_2 = 0$, and β_{13} describes the impact of a joint increase in X_1 and X_3 on Y when $X_2 = 0$.

Figure 1 illustrates these effects. In terms of the graph, which depicts Y at two different values of X_3, β_3 permits the surface to vary at the leftmost point in the graph, β_{13} permits variation of the slope along the edge at which $X_2 = 0$, and β_{123}

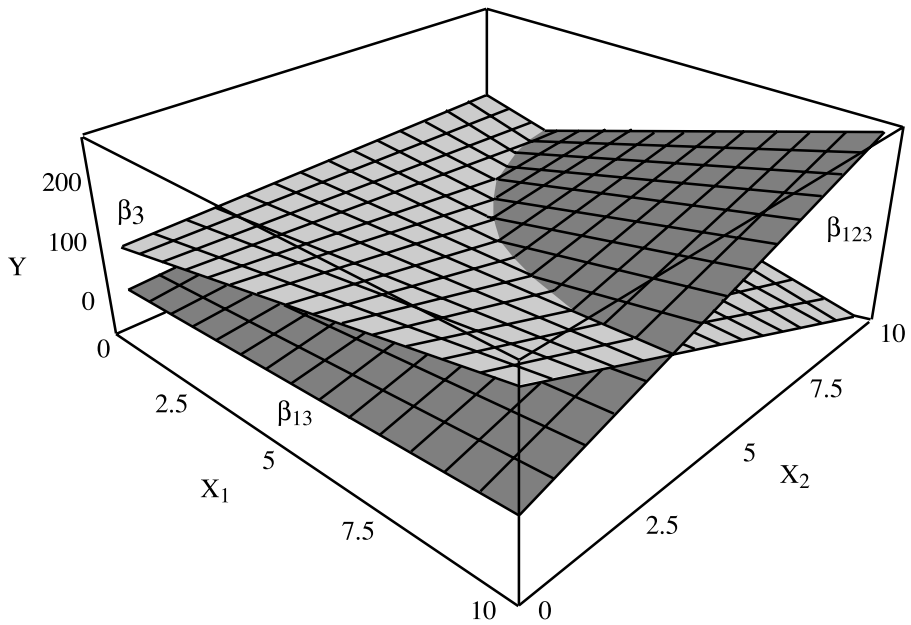


FIGURE 1. $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{23} x_2 x_3 + \beta_{13} x_1 x_3 + \beta_{123} x_1 x_2 x_3$ at two different levels of X_3, showing the impact of β_3, β_{13}, and β_{123}.

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

permits a change in the curvature of the surface. In this case, β_3 and β_{13} are positive, while β_{123} is negative: an increase in X_3, therefore, raises the point at $X_1 = X_2 = 0$ and increases the slope along $X_2 = 0$ while making the surface less convex. Accordingly, the consequences of misinterpretation of coefficients becomes more severe as the number of variables increases: far from saying something universal about the relationship between X_3 and Y, for example, a significant and positive β_3 says nothing when either $X_1 \neq 0$ or $X_2 \neq 0$. One might reasonably wonder whether theories about social phenomena really make predictions about the sign and magnitude of most of these coefficients.

Moreover, the illustration highlights the perils of omitting some or all of the lower-order terms. In any interaction of k independent variables, a full set of $\sum_{k=1}^{n} \binom{n}{k}$ coefficients must be estimated to avoid forcing the estimated hyperplane to assume a shape that may not conform to the general tendency of the pointcloud that it is intended to describe. If $\beta_{13}x_1x_3$ were omitted from the equation, for example, the slope of the line at $X_2 = 0$ would be held constant across all levels of X_3, by assumption, and the remaining coefficients most likely biased as a result.[6] The outcome is analogous to omitting the constant term from a simple bivariate regression, thereby forcing the regression line through the origin regardless of the pattern of the data: the consequences for inference may be negligible or severe.

Why Does It Matter?

Political science is a discipline in which an inordinate amount of importance is placed on *t*- or *z*-statistics.[7] Even those reviewers willing to overlook a significance level of 0.051 pay attention to the ratio of the coefficient to the standard error. Judging by the contents of even the top journals, few scholars realize that that ratio is both arbitrary and not representative of any general trend for the coefficients on the lower-order terms of a set of interactive variables. At the same time, a single article or book with a significant result on a prominent topic can be immensely persuasive, especially if no critics point to flaws in the data or methodology. In combination, these two facts imply that large numbers of scholars can be misled for long periods of time by the simple misinterpretation of a coefficient.

Without pointing fingers—the discipline, not any individual, is really culpable—let me illustrate this point with a reexamination of three articles, all of which were

6. Some of the subtler implications of this point are even more often missed. A scholar estimating an equation of the form $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \varepsilon$ for example, might think that X_2 and X_3, which are not multiplied together, do not interact. Because both are multiplied by X_1, however, they do interact: β_2 reflects the impact of X_2 on Y when X_1 and X_3 equal zero, and the omission of x_2x_3 and $x_1x_2x_3$ from the equation has the effects described above. The result of this "tacit interaction" is a set of coefficients that are both biased and misinterpreted.

7. Indeed, a significance level of $p \leq 0.05$ is mandatory in some major political science journals if the author is to use the coveted asterisk.

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

written by prominent and respected scholars and have played substantial roles in ongoing academic debates.[8] Schultz's (1999) article on whether democratic institutions lead to peace by constraining leaders or informing other states is an exemplary piece in many ways: a clean formal model leads to opposite predictions based on the two theories, and an empirical test favors the informational perspective. Similarly, Mansfield and Snyder's (2002) comprehensive test of the relationship between various regime transitions and war suggests quite strongly that incomplete democratization is hazardous: in twenty-nine of the thirty-five variants of the statistical model that are tested, there is a significant and positive relationship between incomplete democratic transition and war. Finally, Adserà and Boix's (2002) examination of government size (that is, the size of the public sector) makes a strong case for the importance of politics: whereas previous studies suggested that trade and the size of the public sector would be related for purely economic reasons, the authors derive a model in which both are the result of political decisions based ultimately on the distribution of domestic interests. As a result, the authors argue, government should be large in free-trading democracies but small elsewhere,[9] and the tendency toward large government should be exacerbated when state exports become less diversified, especially in democracies.[10] In all cases, an interaction term was included as part of the analysis. Schultz included the democracy of the initiator, the democracy of the target, and an interaction term, noting only that "[t]he hypotheses do not speak to the expected sign and significance of the coefficients on DEMTARG and DEMDEM, but they are nevertheless included as controls."[11] In Mansfield and Snyder's case, incomplete democratic transition was multiplied by a variable measuring the concentration of domestic authority in twenty-five of the thirty-five models; and in all twenty-five, the variable capturing incomplete democratic transition was statistically significant. In Adserà and Boix, the interaction of trade openness, democratic institutions, and export concentration was examined in the search for the sources of variation in government size.

Moreover, in all cases the main conclusions depended on the interpretation of coefficients from lower-order terms, and those coefficients are interpreted as if they applied across all cases. Schultz notes that

> the coefficient on DEMINIT is negative, meaning that the target was less likely to reciprocate a militarized action when the initiator was democratic than when it was not. . . . Moreover, the coefficient on DEMINIT is statistically significant at conventional levels whenever the world war MIDs are excluded. . . . Overall, these findings are consistent with hypothesis 3 and the informational perspective.[12]

8. I am grateful to the authors for being kind enough to provide the data and enough notes to permit the replications and extensions that follow.
9. Adserà and Boix 2002, 238.
10. Ibid., 247.
11. Schultz 1999, 251.
12. Ibid., 253.

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

Mansfield and Snyder, who find significant results for coefficients on both incomplete democratic transitions and the interaction of same with the concentration of domestic authority, discuss the implications of both but emphasize the former:

> We find that the heightened danger of war grows primarily out of the transition from an autocratic regime to one that is partially democratic. The specter of war during this phase of democratization looms especially large when governmental institutions . . . are especially weak.[13]

> incomplete democratic transitions . . . are especially likely to promote the outbreak of war. Furthermore, such transitions become an increasingly potent impetus to war as a state's institutional strength degrades.[14]

Adserà and Boix multiply three variables together (export concentration, trade openness, and democratic institutions), making the task of drawing inferences much more complex; moreover, they omit one term (export concentration $\times$ democratic institutions), forcing that coefficient to zero. Their conclusions based on lower-order terms are unconditional ones:

> The level of export concentration depresses public revenue significantly. The interactive variables of trade openness with export concentration and of these two measures and democracy have positive and statistically significant coefficients. As the tradable sector becomes less diversified and has a more central role in the domestic economy, the pressure for domestic compensation clearly goes up. Under democratic regimes, this pressure intensifies even more.[15]

Similar examples pervade the field, but more would simply belabor the point.

Does the general critique above imply that the authors' conclusions are wrong? No such simple assertion is possible, because—and this is the key point—the tests were never designed in such a way that the conclusions reached were meaningful ones for more than a subset of the data. Schultz's significant coefficient describes what happens to militarized interstate disputes when the initiators are democratic and the targets are not (DEMTARG$=0$), not what happens to militarized interstate disputes when the initiators are democratic in general. Mansfield and Snyder's coefficient describes what happens when there is an incomplete democratic transition and domestic concentration is at its lowest. Adserà and Boix's conclusion about the effects of export concentration apply, but only when trade openness and democratic institutions equal zero. Each conclusion is therefore correct, but only for a subset of the cases. That said, in the latter two instances that subset is actually empty: Mansfield and Snyder's data set contains no instances of incomplete dem-

13. Mansfield and Snyder 2002, 298.
14. Ibid., 318.
15. Adserà and Boix 2002, 247.

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

ocratic transitions when domestic concentration is at its lowest, and Adserà and Boix's contains no instances in which trade openness equals zero.

Reanalysis and Reinterpretation

How can one obtain more generalizable answers about the relationships of these key independent variables to the dependent variable of interest?

In the case of Schultz, the interaction term was added to the equation without much theoretical justification, and the coefficient suggests strongly that it adds little to the results. The simple remedy is to drop it. In Table 2, therefore, I have reanalyzed the main model from Schultz's article, dropping first the interaction term and then the DEMTARG variable. As intuition might suggest, the omission of

TABLE 2. *Reanalyses of Schultz (1999) Model 2*

Variable	*Replication*	*Reanalysis 1*	*Reanalysis 2*
Constant	0.081	0.070	0.040
	(0.112)	(0.111)	(0.104)
DEMOCRATIC INITIATOR	−0.240*	−0.212*	−0.176*
	(0.099)	(0.090)	(0.085)
DEMOCRATIC TARGET	−0.100	−0.078	—
	(0.090)	(0.084)	
BOTH DEMOCRACIES	0.148	—	—
	(0.227)		
CONTIGUOUS	0.451**	0.451**	0.427**
	(0.087)	(0.087)	(0.086)
ALLIANCE	−0.032*	−0.023	−0.060
	(0.102)	(0.101)	(0.096)
MAJOR-MAJOR	−0.226	−0.228	−0.283*
	(0.125)	(0.125)	(0.123)
MAJOR-MINOR	−0.245*	−0.246*	−0.282**
	(0.100)	(0.100)	(0.095)
MINOR-MAJOR	−0.033	−0.033	−0.080
	(0.128)	(0.128)	(0.125)
TERRITORY	0.235*	0.240*	0.284**
	(0.102)	(0.102)	(0.099)
POLICY	−0.697**	−0.694**	−0.618**
	(0.094)	(0.094)	(0.093)
GOVERNMENT/REGIME	0.375	0.378	0.357
	(0.206)	(0.205)	(0.185)
OTHER	−0.552	−0.550	−0.446
	(0.302)	(0.302)	(0.294)

Note: Parameters are probit coefficients. Standard errors are in parentheses and are adjusted for clustering on dyad to ensure consistency with original analysis.
* significant at 0.05 level.
** significant at 0.01 level.

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

the relatively insignificant interaction term makes little difference for the effects of democratic initiation. The coefficient decreases in magnitude by roughly 25 percent but remains statistically significant ($p = 0.037$, versus $p = 0.015$ in the replication). Happily, the main result stands, though its substantive and statistical significance are somewhat attenuated.

In the Mansfield and Snyder piece, on the other hand, omitting the interaction term is not a legitimate option: both theory and empirical results suggest that it belongs in the equation. At the same time, one would like to know more about the relationship between incomplete democratic transitions and war in cases in which domestic concentration is greater than zero.

One simple solution[16] is to take advantage of the fact that β_1 measures the impact of X_1 on Y when $X_2 = 0$ by recoding X_2 in steps and describing how β_1 changes. In this case, that means subtracting n ($n = 1, 2, \ldots, 9$) from domestic concentration and reestimating the logit equation to see how the coefficient and standard errors on incomplete democratic transition vary as a function of domestic concentration.

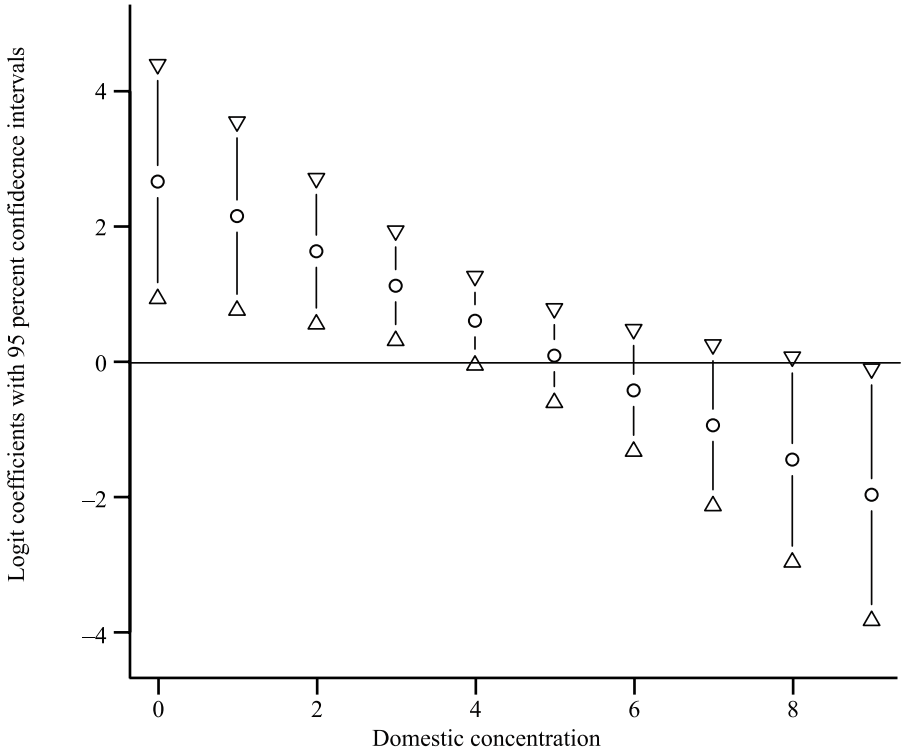
Figure 2 illustrates the results of such an analysis. The figure demonstrates that the positive relationship between incomplete democratization and war, far from being a general result, is limited only to cases in which democratic concentration is relatively low—say, from zero to 4 on the 10-point concentration scale. Only about 27 percent of the cases of incomplete democratization fall into this range. In the bulk of the cases—those in which concentration ranges from 5 to 7, which constitute 67 percent of all cases of incomplete transition—there is enough uncertainty about the relationship that it cannot reliably be distinguished from zero and would fail conventional tests of statistical significance. Interestingly, in cases of high concentration (an 8 or a 9, which constitute about 6 percent of the transition cases), the coefficients are negative and significant, suggesting that incomplete democratization produces peace, not war.

Rather than concluding, then, that incomplete democratic transitions "are a potent impetus to war, especially when the level of domestic concentration is low,"[17] the authors could more reasonably have concluded that such transitions are an impetus to war only in those few cases in which the level of domestic concentration is low, and that they might even be conducive to peace if levels of domestic concentration are high. In more than 70 percent of the cases, an unconditional assertion

16. There are a multitude of alternative techniques for modeling interaction effects—such as Cobb-Douglas production functions (Cobb and Douglas 1928) for continuous, nonnegative dependent variables or Boolean logit and probit (Braumoeller 2003) for dichotomous dependent variables—or for increasing the flexibility of the functional form: anything from simply breaking the concentration variable into nine dummies to using variable-parameter models (Kennedy 1985, 74–76) to using generalized additive models (GAMs; Beck and Jackman 1998) would permit researchers to do so. The solution advocated here is designed for the researcher who wishes to use simple multiplicative interaction terms but who nevertheless desires a more thorough and meaningful interpretation of the relationships involved.

17. Mansfield and Snyder 2002, 322.

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press



Note: Circles are coefficients, triangles are one- and two-standard error confidence intervals.

FIGURE 2. *Logit coefficients relating incomplete democratic transition to war, at different levels of domestic concentration (base model)*

that incomplete democratic transitions increase the probability of war cannot be supported.

Finally, the Adserà and Boix analysis provides an illustration of some of the intricacies involved in analyzing interactions with more than two variables. Although the interaction of product concentration of exports and democratic institutions is not thought to play a prominent theoretical role in their story, the interaction term must nevertheless be included because the product of those two variables is multiplied by trade openness. Omitting concentration × institutions (or EC × DI) implies that the impact of a joint increase in those two variables when trade openness is zero must equal zero, by assumption. In terms of Figure 1, the slope of the surface along the X_2 (democracy) axis where X_1 (trade openness) equals zero cannot vary.

The results of a reanalysis including (concentration × institutions), summarized in Table 3, demonstrate this point. When EC × DI is included in the model, the

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

TABLE 3. *Reanalyses of Adserà and Boix (2002) Table 1, Model 5*

Variable	*Replication*	*Reanalysis*
Constant	−23.917***	−15.896
	(8.886)	(9.934)
PER CAPITA INCOME	6.892***	6.994***
	(0.815)	(0.811)
TRADE OPENNESS (TO)	−0.780	−2.974*
	(1.402)	(1.809)
DEMOCRATIC INSTITUTIONS (DI)	−8.713*	−23.953***
	(4.708)	(7.884)
DI × TO	3.248**	7.026***
	(1.286)	(2.027)
AREA	0.943**	0.945***
	(0.354)	(0.352)
DISTANCE	−0.841**	−0.817**
	(0.352)	(0.348)
SUBSAHARAN AFRICA	−0.908	−0.625
	(2.783)	(2.795)
EAST ASIA	−5.895**	−5.493**
	(2.516)	(2.516)
LATIN AMERICA	−8.973***	−9.235***
	(2.724)	(2.731)
OECD	−4.597*	−4.752**
	(2.368)	(2.388)
EXPORT CONCENTRATION (EC)	−27.269**	−47.014***
	(12.447)	(18.109)
EC × TO	7.336**	12.024***
	(3.062)	(4.373)
EC × TO × DI	0.353	−10.107*
	(0.858)	(5.606)
EC × DI	—	42.928*
		(22.565)

Note: Parameters are ordinary least squares (OLS) coefficients. Standard errors are in parentheses. Procedure assumes heteroskedastic panels and common AR(1) coefficient for all panels to ensure consistency with original analysis.
* significant at 0.10 level.
** significant at 0.05 level.
*** significant at 0.01 level.

estimated coefficient is far from zero. Accordingly, while the coefficients on the variables not included in the interaction remain roughly the same, those associated with variables in the interaction term (and the constant) vary wildly. The lower-order coefficients indicate some changes that appear more alarming—and more meaningful—than they really are, given the correct interpretation of the coefficients: trade openness appears to have a negative impact on size of government (when institutions are nondemocratic and concentration, which never reaches zero, is zero); democratic institutions have a much larger and more significant negative impact on size of government (when trade openness and concentration, which never

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

reach zero, are zero). The most striking change is the finding that joint increases in trade openness, concentration, and democratic institutions, which were previously thought to have no impact on size of government, are now shown to have a negative and significant impact. Figure 3 illustrates the impact of changes in product concentration on the relationship between trade, democratic institutions, and size of government.[18]

Substantively, the results tell a rather complex story, one that only partly agrees with the original conclusions. The authors argued that a large public sector is the product of the combination of high trade openness and political democracy. When product concentration is at its lowest (top illustration in figure), that is, when the country exports a diverse array of products, the generalization holds—but the positive interaction term and the negative marginal terms produce a saddle effect. The results suggest that size of government is nearly as large in autarkic autocracies as it is in free-trading democracies, a finding unanticipated by the theory.

When product concentration is at its highest (bottom illustration in figure), meaning that the country only exports a single product, the marginal relationships are reversed, and the joint effect of trade and democracy becomes negligible—in fact, ever so slightly negative. The effects of decreasing diversification, therefore, have more nuanced effects than the authors suggest: it drives the pressure for domestic compensation (as measured by size of government) down rather than up for a wide range of states, and it actually increases that pressure more in free-trading autocracies than it does in free-trading democracies.

Conclusion

When independent variables are multiplied together to model interaction, a set of coefficients jointly describes the behavior of the variables. By virtue of their interactive nature, no statistical wizardry can "centrifuge out" a coefficient that corresponds to what most hypothesis tests take to be a standard regression coefficient—one that allows researchers to test the theory that a unit increase in X_1 is associated with a fixed change in Y at all levels of X_2. For that reason, although their estimation is typically necessary to avoid introducing artificial constraints into the analysis, lower-order coefficients are not quantities of direct interest for most hypothesis tests. Indeed, these coefficients often describe relationships that exist only outside of the range of the actual data. Failure to appreciate this relatively straightforward methodological point is widespread, even among the most respected scholars, and can have profound substantive repercussions.

18. All other variables are held at mean values, and all geographical dummies are set to zero.

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

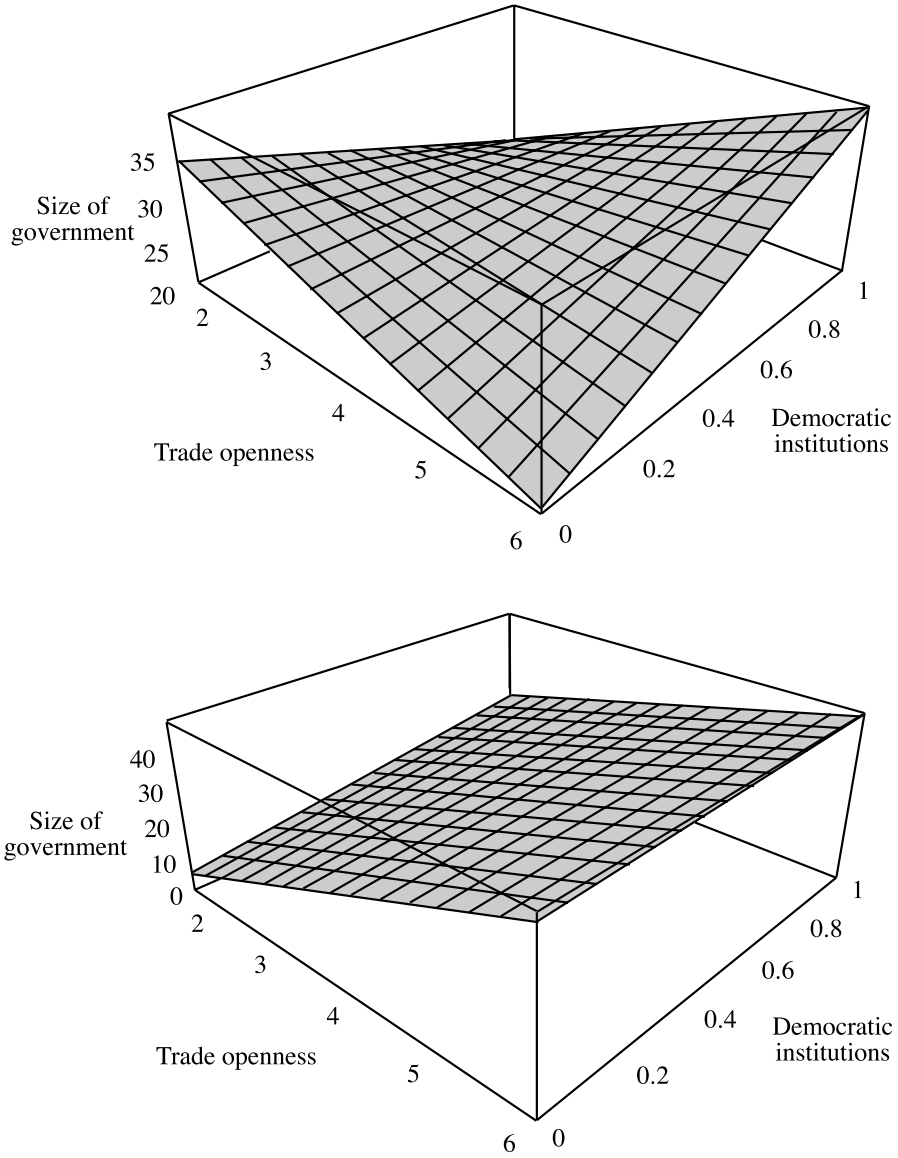


FIGURE 3. *Estimated relationship between trade openness, democracy, and size of government, at low (top) and high (bottom) levels of export product concentration*

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press

References

Adserà, Alícia, and Carles Boix. 2002. Trade, Democracy, and the Size of the Public Sector: The Political Underpinnings of Openness. *International Organization* 56 (2):229–62.

Allison, Paul D. 1977. Testing for Interaction in Multiple Regression. *American Journal of Sociology* 83 (1):144–53.

Beck, Nathaniel, and Simon Jackman. 1998. Beyond Linearity by Default: Generalized Additive Models. *American Journal of Political Science* 42 (2):596–627.

Braumoeller, Bear F. 2003. Causal Complexity and the Study of Politics. *Political Analysis* 11 (3): 209–33.

Cobb, Charles W., and Paul H. Douglas. 1928. A Theory of Production. *The American Economic Review* 18 (1):139–65.

Friedrich, Robert J. 1982. In Defense of Multiplicative Terms in Multiple Regression Equations. *American Journal of Political Science* 26 (4):797–833.

Kennedy, Paul. 1985. *A Guide to Econometrics*. 2d ed. Cambridge, Mass.: MIT Press.

Mansfield, Edward D., and Jack Snyder. 2002. Democratic Transitions, Institutional Strength, and War. *International Organization* 56 (2):297–337.

Schultz, Kenneth. 1999. Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War. *International Organization* 53 (2):233–66.

https://doi.org/10.1017/S0020818304040251 Published online by Cambridge University Press