**Political Analysis Advance Access published May 18, 2005**

*Political Analysis* (2005) 13:1–20
doi:10.1093/pan/mpi014

# Understanding Interaction Models: Improving Empirical Analyses

**Thomas Brambor**
*New York University, Department of Politics,*
*726 Broadway, 7th Floor, New York, NY 10003*
*e-mail: thomas.brambor@nyu.edu*

**William Roberts Clark**
*University of Michigan, Center for Political Studies,*
*ISR 4202 Box 1248, 426 Thompson Street, Ann Arbor, MI 48106-1248*
*e-mail: wrclark@umich.edu*

**Matt Golder**
*Florida State University, Department of Political Science,*
*531 Bellamy Building, Tallahassee, FL 32306-2230*
*e-mail: matt.golder@nyu.edu (corresponding author)*

Multiplicative interaction models are common in the quantitative political science literature. This is so for good reason. Institutional arguments frequently imply that the relationship between political inputs and outcomes varies depending on the institutional context. Models of strategic interaction typically produce conditional hypotheses as well. Although conditional hypotheses are ubiquitous in political science and multiplicative interaction models have been found to capture their intuition quite well, a survey of the top three political science journals from 1998 to 2002 suggests that the execution of these models is often flawed and inferential errors are common. We believe that considerable progress in our understanding of the political world can occur if scholars follow the simple checklist of dos and don'ts for using multiplicative interaction models presented in this article. Only 10% of the articles in our survey followed the checklist.

*Authors' note:* Our thanks go to Nathaniel Beck, Fred Boehmke, Michael Gilligan, Sona Nadenichek Golder, Jonathan Nagler, and two anonymous reviewers for their extremely useful comments on this paper. We also thank the research assistants at *Political Analysis*—Jeronimo Cortina, Tse-hsin Chen, and Seung Jin Jang—for kindly double-checking the results from our literature survey. Finally, we are grateful to those authors who have provided us with their data. To accompany this paper, we have constructed a Web page at http://homepages.nyu.edu/~mrg217/interaction.html that is devoted to multiplicative interaction models. On this page, you will find (i) the data and computer code necessary to replicate the analyses conducted here, (ii) information relating to marginal effects and standard errors in interaction models, (iii) STATA code for producing figures illustrating marginal effects and confidence intervals for a variety of continuous and limited dependent variable models, and (iv) detailed results from our literature survey. STATA 8 was the statistical package used in this study.

## 1 Introduction

Multiplicative interaction models are common in the quantitative political science literature. This is so for good reason. Institutional arguments frequently imply that the relationship between political inputs and outcomes varies depending on the institutional context. Moreover, models of strategic interaction typically produce conditional, rather than unconditional, hypotheses. Indeed, it could be argued that any causal claim implies a set of conditions that need to be satisfied before a purported cause is sufficient to bring about its effect. It is little wonder, then, that conditional hypotheses such as "an increase in X is associated with an increase in Y when condition Z is met, but not otherwise" are ubiquitous in all fields of political science. It has been well established that the intuition behind conditional hypotheses is captured quite well by multiplicative interaction models (Wright 1976; Friedrich 1982; Aiken and West 1991). Given the ubiquity and centrality of conditional hypotheses in political science, it is somewhat surprising that the execution of multiplicative interaction models is often flawed and inferential errors are common. In a survey of the political science literature described below, we found that only 10% of all articles employing interaction models actually avoided all three of the commonly encountered problems that we address in this article.

Our goal is not to provide an exhaustive discussion of the rather vast literature on multiplicative interaction models.[1] Instead, it is to present several ways in which political scientists can dramatically improve empirical analyses employing these types of models. First, analysts should use interaction models whenever the hypothesis they want to test is conditional in nature. Second, scholars should include all constitutive terms in their interaction model specifications. Third, scholars should not interpret constitutive terms as if they are unconditional marginal effects. Finally, analysts should calculate substantively meaningful marginal effects and standard errors. While these points have been forcefully made in the statistics and political science literature for some time and may seem obvious to some, an examination of three leading political science journals from 1998 to 2002 indicates that they are rarely reflected in common practice. In fact, of the 156 articles that employed interaction models, only 16 actually included all constitutive terms, did not make mistakes interpreting these terms, and calculated substantively meaningful marginal effects and standard errors.

## 2 Include Interaction Terms

Analysts should include interaction terms whenever they have conditional hypotheses. A conditional hypothesis is simply one in which a relationship between two or more variables depends on the value of one or more other variables. Perhaps the simplest conditional hypothesis is:

> $H_1$: An increase in X is associated with an increase in Y when condition Z is met, but not when condition Z is absent.

[1]A good introduction to interaction models can be found in Friedrich (1982), Aiken and West (1991), as well as Kam and Franzese (2003). Other literature dealing with interaction models includes Wright (1976), Allison (1977), Bernhardt and Jung (1979), Griepentrog et al. (1982), Cleary and Kessler (1982), Busemeyer and Jones (1983), Cox (1984), Rosnow and Rosenthal (1989, 1991), Jaccard et al. (1990), Nagler (1991, 1994), Jaccard and Wan (1995), Gill (2001), Franzese (2003a), Ai and Norton (2003), Braumoeller (2004), and Norton et al. (2004).

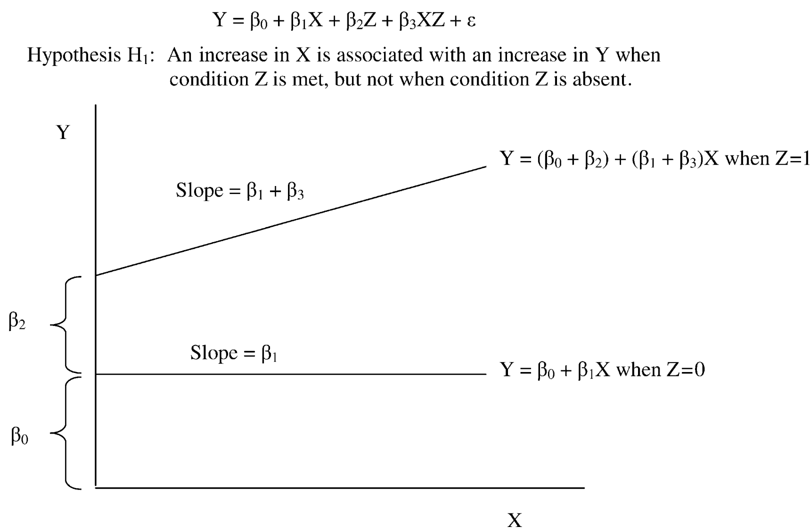


**Fig. 1** A graphical illustration of an interaction model consistent with hypothesis $H_1$.

Although these types of hypotheses are particularly prevalent in the field of comparative politics where the importance of context or "context conditionality" is strongly emphasized, they are increasingly common in the other subfields of political science (Franzese 2003b). Conditional hypotheses can easily be tested using multiplicative interaction models. To see this, assume that Y and X are continuous variables, while Z is a dichotomous variable that equals one when the required condition is met, and zero otherwise.

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ + \epsilon \tag{1}$$

It is relatively easy to see that the model presented in Eq. (1) captures the intuition behind our hypothesis.[2] This is because when condition Z is absent, i.e., Z = 0, Eq. (1) simplifies to:

$$Y = \beta_0 + \beta_1 X + \epsilon. \tag{2}$$

It should now be clear that $\beta_1$ in Eq. (1) captures the effect of a one-unit change in X on Y when condition Z is absent ($\frac{\partial Y}{\partial X}$ given Z = 0 is $\beta_1$). When condition Z is present, Eq. (1) can be simplified to:

$$Y = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)X + \epsilon. \tag{3}$$

This illustrates that the effect of a one-unit change in X on Y when condition Z is present is now $\beta_1 + \beta_3$ ($\frac{\partial Y}{\partial X}$ given Z = 1). Since our hypothesis is that Y is increasing in X *if and only if* Z is present, then we should expect to find that $\beta_1$ is zero and that $\beta_1 + \beta_3$ is positive. These two implications necessitate that $\beta_3$ should be positive. Figure 1 graphically

[2]There is no requirement that the form of the interaction term in interaction models be the product of the constitutive terms X and Z as is the case here. However, we focus on these multiplicative interaction models because they are the most common in political science.

illustrates a multiplicative interaction model that is consistent with hypothesis $H_1$. Note that Fig. 1 assumes that $\beta_0$ and $\beta_2$ are both positive. While this simple example involves a dichotomous modifying variable—condition Z modifies the effect of X on Y—this basic interaction model can be extended to take account of continuous modifying variables or arguments with even greater causal complexity. We use a dichotomous modifying variable in this example and elsewhere purely for ease of presentation. However, the points that we make do not depend on whether the modifying variable Z is dichotomous or continuous.

## 3 Include All Constitutive Terms

Analysts should include all constitutive terms when specifying multiplicative interaction models except in very rare circumstances. By constitutive terms, we mean each of the elements that constitute the interaction term. Thus, X and Z are the constitutive terms in Eq. (1). Throughout this article we focus on multiplicative interaction models that include interaction terms such as XZ. Again, we do this for purely presentational purposes. The reader should note, though, that multiplicative interaction models can take a variety of forms and may involve quadratic terms such as $X^2$ or higher-order interaction terms such as XZJ. No matter what form the interaction term takes, all constitutive terms should be included. Thus, X should be included when the interaction term is $X^2$ and X, Z, J, XZ, XJ, and ZJ should be included when the interaction term is XZJ.

Although the statistics (and political science) literature is clear that all constitutive terms should be included, scholars often fall prey to the temptation to exclude one or more of them. To understand the consequences of doing this, it is useful once more to examine the simple interaction model specified in Eq. (1). Below, we investigate the consequences of omitting the constitutive term Z. Note that excluding Z is equivalent to assuming that $\beta_2$ in Eq. (1) is zero. The resulting model is now:

$$Y = \gamma_0 + \gamma_1 X + \gamma_3 XZ + \nu. \tag{4}$$

In those cases in which analysts actually say anything about why they omit a constitutive term, they typically provide one of the following justifications for why they specify a model similar to Eq. (4) rather than the fully-specified model in Eq. (1). First, some claim that they do not believe that Z has any effect on Y on average and that, as a result, they do not need to include it as a separate term in the model. Second, others claim that they do not believe that Z has an effect on Y when X is zero and this means they can exclude it as a separate variable from their model. Below, we show that the first of these claims is never justified and that the second claim is rarely defensible.

Note that both claims used to justify the omission of the constitutive term Z are based on the expectation that $\beta_2$ in Eq. (1) is zero. The first claim is relatively easy to refute. As we show in more detail in the next section, $\beta_2$ does not represent the average effect of Z on Y; it only indicates the effect of Z when X is zero.[3] Thus, even if the average effect of Z on Y is zero, it need not be the case that $\beta_2$ is zero. As a result, the assertion that the average effect of Z on Y is zero is never a justification for omitting the constitutive term.

The second claim is based on a theoretical expectation that Z has no effect on Y when X is zero ($\beta_2 = 0$). Unfortunately, there is reason to believe that the omission of a constitutive term may still lead to inferential errors even when the analyst is armed with a strong

[3] It is true that $\beta_2$ is the average effect of Z on Y when there is no interaction effect present, i.e., $\beta_3 = 0$. However, if this is the case, then neither model (1) nor model (4) is appropriate.

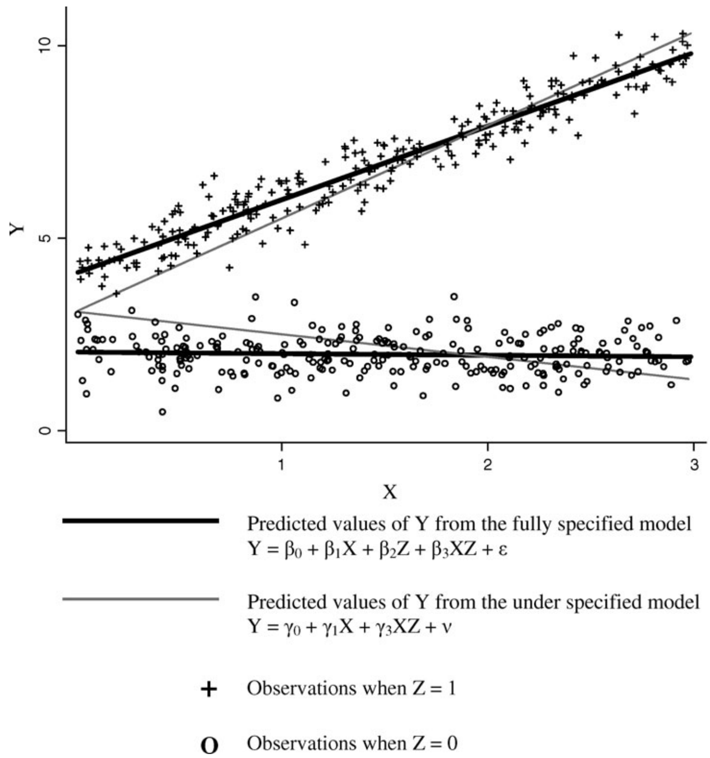


**Fig. 2** An illustration of the consequences of omitting a constitutive term.

conditional theory such as this. The basic point is that the analyst's theory may be wrong and $\beta_2$ may in fact not be zero. If this is the case and Z is correlated with either XZ (or X) as will occur in virtually any social science circumstance, then omitting the constitutive term Z will result in biased (and inconsistent) estimates of $\beta_0$, $\beta_1$, and $\beta_3$. Although not always recognized as such, this is a straightforward case of omitted variable bias (Greene 2003, pp. 148–149). It is worthwhile spending a moment to examine exactly how and why the omission of a constitutive term may lead to biased estimates in multiplicative interaction models.

Figure 2 illustrates a scatter plot of 500 observations generated by the process implied in Eq. (1), where $\beta_0 = 2$, $\beta_1 = 0$, $\beta_2 = 2$, and $\beta_3 = 2$, and Z is a dichotomous modifying variable.[4] The crosses indicate observations when condition Z is present (Z = 1) and the circles indicate observations when condition Z is absent (Z = 0). The thick dark lines capture the predicted values of Y derived from a fully specified model that includes Z; the thin light lines capture the predicted values of Y from model (4), which omits Z.

As we have already stated, omitting the constitutive term Z is the same as assuming that $\beta_2 = 0$. Since Fig. 1 indicated that $\beta_2$ captures the difference in the intercepts between the regression lines for the case in which condition Z is present and the case in which

[4] X was generated as a uniform variable on the unit interval and then multiplied by three. Z was originally drawn from a uniform distribution on the unit interval; it was then recoded as 1 if $Z \geq 0.5$ and 0 if $Z < 0.5$. The error term was randomly drawn from a normal distribution with mean 0 and variance 0.5. The true parameters of this model indicate that X has no effect on Y when Z is absent ($\beta_1 = 0$) but that X increases Y when Z is present ($\beta_1 + \beta_3(1) = 2$).

condition Z is absent, it should now be clear that omitting Z amounts to constraining the two regression lines to meet on the Y axis. Note that in some sense this is equivalent to specifying a model without a constant term and forcing the regression line to go through the origin. As one would expect, forcing the two lines to meet on the Y axis can happen only if the slopes of the regression lines (and the angle between them) change. In effect, instead of estimating the slopes as $\beta_1$ and $\beta_1 + \beta_3$, the model omitting the constitutive term estimates them as $\gamma_1$ and $\gamma_1 + \gamma_3$. Instead of estimating two intercepts ($\beta_0$ and $\beta_0 + \beta_2$), the underspecified model estimates only one ($\gamma_0$). In other words, the estimates of the parameters of interest will all be biased whenever $\beta_2$ is not zero. More precisely, the coefficients estimated by the underspecified model in Eq. (4) will be $\gamma_0 = \beta_0 + \beta_2\alpha_0$, $\gamma_1 = \beta_1 + \beta_2\alpha_1$, and $\gamma_3 = \beta_3 + \beta_2\alpha_3$, where the βs are simply the coefficients from the fully specified model in Eq. (1) and the αs are the coefficients from the regression of Z on X and XZ, i.e., $Z = \alpha_0 + \alpha_1 X + \alpha_3 XZ + \varepsilon$. Practically speaking, the extent of the bias will depend on (i) the degree to which $\beta_2$ differs from zero and (ii) the magnitude of $\beta_2$ relative to the magnitudes of $\beta_0$, $\beta_1$, and $\beta_3$. In addition, the extent to which each of the coefficients is individually biased depends on the distribution of the modifying variable.

Are there any circumstances in which omitting a constitutive term would not lead to significant inferential errors? Possibly. However, there are at least two necessary conditions that must be met before an analyst considers omitting constitutive terms. First, the analyst must have a strong theoretical expectation that the omitted variable (Z in this case) has no effect on the dependent variable in the absence of the other modifying variable (X = 0 in this case). Note, though, that the only situation in which this theoretical expectation can be justified a priori is if X is measured with a natural zero. This is because the coefficients on constitutive terms depend on how an analyst scales these variables. For example, using a measure of democracy such as the familiar Polity score that is scaled from −10 to 10 in an interaction model will generate a different coefficient on the variable that it interacts with than the same measure of democracy that is scaled from 0 to 20. This can be shown quite easily (Braumoeller 2004).

Start with our basic fully specified model outlined in Eq. (1). Imagine that some arbitrary constant L such as 10 in our Polity example is now added to X to create $X^*$. The model now becomes

$$Y = \beta_0 + \beta_1(X^* - L) + \beta_2 Z + \beta_3(X^* - L)Z + \epsilon. \tag{5}$$

Rewriting, we get

$$Y = (\beta_0 - \beta_1 L) + \beta_1 X^* + (\beta_2 - \beta_3 L)Z + \beta_3 X^* Z + \epsilon. \tag{6}$$

It should be clear that rescaling X in this arbitrary way changes the coefficient on Z from $\beta_2$ to $\beta_2 - \beta_3 L$. The standard error of the coefficient on Z naturally changes as well (Bernhardt and Jung 1979; Griepentrog et al. 1982).[5] $\beta_2$ may truly be zero, but we have no way of knowing in practice if we are estimating $\beta_2$ or $\beta_2 - \beta_3 L$ if our theory does not tell

[5] Note that Eqs. (5) and (6) clearly show that the arbitrary rescaling of X (which results in a different zero point for this variable) has no effect on the coefficient on the interaction term. This reminds us that the significance of an interaction term does not depend on the zero point of the independent variables in an interaction model. It is also worth remembering at this point that the significance of an interaction term does not depend on the units in which the independent variables are measured either. Although changing the units of the independent variables will change the size of the coefficient on the interaction term, the standard error will also change size proportionately, leaving the statistical significance unchanged.

us which particular scale to use for X (something that is normally the case). In other words, the analyst has no way of predicting a priori what the coefficient on Z will be before actually estimating his or her model. As a result, the analyst *cannot* have a theoretical justification for omitting Z in this case. It is only if X has a natural zero that the analyst can be justified in having a theoretical prediction that the coefficient on Z is zero when X is zero. This is because the zero point of a variable with a natural zero is not subject to rescaling by definition. The widespread use of scales and indices such as the Polity score in the political science literature, though, suggests that analysts will often have variables that do not have a natural zero.

The second condition that must be met before omitting a constitutive term is that the analyst should estimate the fully specified model outlined in Eq. (1) and find that $\beta_2$ is zero. In other words, the analyst should test his or her claim that $\beta_2$ is in fact zero. Note that even if $\beta_2$ is statistically indistinguishable from zero, the other parameters of interest will still be estimated with bias to the extent that $\beta_2$ is not *exactly* zero if the constitutive term is dropped.[6] This may or may not be much of a problem in practice. Much will depend on whether $\beta_2$ is close to zero and whether $\beta_2$ is small relative to $\beta_1$ and $\beta_3$.

Thus, there may be a set of very limited conditions under which it is appropriate to estimate a model that omits constitutive terms. However, given that the second condition requires the analyst to estimate a fully specified model, there is no compelling reason why the analyst should not simply report these results. At the very least, the analyst should report the estimated coefficient and standard error of the omitted constitutive term from the fully specified model.

We should note at this point that all of the interaction models presented in this article have the same basic specification as that shown in Eq. (1). We focus on this particular specification because it easily handles both discrete and continuous modifying variables. In this type of setup, it is necessary to include all of the constitutive terms. However, there is an alternative way of specifying interaction models for those specific cases in which the modifying variable is discrete (Wright 1976):

$$Y = \tau_0 + \tau_1 Z + \tau_2 XZ + \tau_3(X\neg Z) + \epsilon. \tag{7}$$

For example, the modifying variable might be region, where Z = north and $\neg$Z = south (where all observations are coded as either north or south). In this example, $\tau_2$ in Eq. (7) indicates the effect of a unit change in X when Z is north, while $\tau_3$ indicates the effect of X when Z is south. Note in this case that it is only necessary to include Z as a separate variable when the interaction terms included in the model are $XZ$ and $X \neg Y$. Why is this? First, it is not possible to include both X and Z in this setup since this leads to perfect multicollinearity. Second, it is necessary to include Z because omitting it constrains the two regression lines (one for north, one for south) to have the same intercept ($\tau_0$) and potentially biases the estimates in exactly the same way as outlined earlier in Fig. 2. Note that the analyst should not include X instead of Z since this still constrains the two regression lines to have the same intercept. Thus, the analyst must include Z and only Z if

[6]This is particularly problematic for the coefficient on the interaction term. Earlier we showed that this coefficient in the underspecified model is estimated as $\gamma_3 = \beta_3 + \beta_2\alpha_3$, where $\beta_3$ represents the true parameter from the fully specified model and $\beta_2\alpha_3$ represents the bias. Remember that $\alpha_3$ comes from the regression of Z on X and XZ, i.e., $Z = \alpha_0 + \alpha_1 X + \alpha_3 XZ + \varepsilon$. Given the obvious relationship between Z and XZ, it is implausible that $\alpha_3$ would ever be zero. This raises the distinct possibility that the coefficient on the interaction term in the underspecified model ($\gamma_3$) could suffer from considerable bias even if $\beta_2$ is close to zero. Analysts should, therefore, be extremely wary of omitting constitutive terms even if they find that $\beta_2$ is insignificant or close to zero.

the interaction terms are XZ and $X \neg Z$. This line of argument is easily generalized to the case in which X interacts with K logically exhaustive and mutually exclusive dummy variables. In this case, the analyst should include either (i) the K dummy variables and no constant or (ii) K−1 of the dummy variables and a constant as constitutive terms.

### 3.1 *Multicollinearity*

A reader might respond that including all constitutive terms in an interaction model increases multicollinearity, thereby increasing the size of the standard errors and making it less likely that the coefficient on the interaction term will be significant. This may be true but does not justify the omission of constitutive terms. First, as Friedrich (1982) notes, the problem of multicollinearity in multiplicative interaction models has been overstated. Much of the concern about multicollinearity arises when the analyst observes that the coefficients from a linear-additive model change when an interaction term is introduced. In the linear-additive world, the sensitivity of results to the inclusion of an additional variable is often taken as a sign of multicollinearity. However, this need not be the case with interaction models. As we have already indicated, the coefficients in interaction models no longer indicate the average effect of a variable as they do in an additive model. As a result, they are almost certain to change with the inclusion of an interaction term, and this should not be interpreted as a sign of multicollinearity.

Even if there really is high multicollinearity and this leads to large standard errors on the model parameters, it is important to remember that these standard errors are never in any sense "too" large—they are always the "correct" standard errors. High multicollinearity simply means that there is not enough information in the data to estimate the model parameters accurately and the standard errors rightfully reflect this. Even more important to remember is that the analyst is not directly interested in the significance or insignificance of the model parameters per se anyway. Instead, the analyst who employs a multiplicative interaction model is typically interested in the marginal effect of X on Y. In the case of Eq. (1), this is $\frac{\partial Y}{\partial X} = \beta_1 + \beta_3 Z$. As a result, the analyst really wants to know the standard error of this quantity and not the standard error of $\beta_0$, $\beta_1$, $\beta_2$, or $\beta_3$. The standard error of interest is:

$$\hat{\sigma}_{\frac{\partial Y}{\partial X}} = \sqrt{var(\hat{\beta}_1) + Z^2 var(\hat{\beta}_3) + 2Z cov(\hat{\beta}_1 \hat{\beta}_3)}. \tag{8}$$

If the covariance term is negative, as is often the case, then it is entirely possible for $\beta_1 + \beta_3 Z$ to be significant for substantively relevant values of Z even if all of the model parameters are insignificant.[7] The fact that analysts are not directly interested in the significance of model parameters suggests that problems associated with multicollinearity are often exaggerated in the context of multiplicative interaction models.[8] Overall, the danger of inferential errors from omitting constitutive terms outweighs any possible benefits. As a result, analysts should include all constitutive terms in their multiplicative interaction models except in very rare circumstances.

[7] For an example of this, see Brambor et al. (2005).

[8] Of course, if there is perfect multicollinearity because one of the variables in the interaction model is only observed for a subset of the cases being analyzed, then the analyst will simply not be able to include all of the constitutive terms. However, this type of situation should be relatively rare. One example where it was necessary to omit a constitutive term due to perfect multicollinearity is Smith's (2000) analysis of ballot initiatives and voter turnout in U.S. legislative elections.

While the problem of multicollinearity has been overstated in the context of interaction models and should not cause the analyst to omit constitutive terms, some scholars have argued that "centering" the relevant variables can mitigate any multicollinearity issues that exist. This is simply not true. As Kam and Franzese (2003, p. 3) note in their impeccably clear and forthright discussion of this topic, centering "alters nothing important statistically and nothing at all substantively." Again, the basic point is that multicollinearity issues arise because there is too little information in the data. Since centering does not provide us with any new or more accurate data, it cannot help us overcome any problems with multicollinearity. This is relatively easy to show.

Compare our basic "uncentered" interaction model outlined in Eq. (1) with the one below:

$$Y = \delta_0 + \delta_1 X_c + \delta_2 Z_c + \delta_3 X_c Z_c + \epsilon_c, \tag{9}$$

where the variables have been centered by subtracting their means, i.e., $X_c = X - \bar{X}$ and $Z_c = Z - \bar{Z}$. Some substitution and rearranging allows us to rewrite Eq. (9) as

$$Y = \delta_0 - \delta_1\bar{X} - \delta_2\bar{Z} + \delta_3\bar{X}\bar{Z} + (\delta_1 - \delta_3\bar{Z})X + (\delta_2 - \delta_3\bar{X})Z + \delta_3 XZ + \epsilon_c. \tag{10}$$

It should now be clear that this centered model is just an algebraic transformation of our uncentered model from earlier where $\beta_0 = \delta_0 - \delta_1\bar{X} - \delta_2\bar{Z} + \delta_3\bar{X}\,\bar{Z}$, $\beta_1 = \delta_1 - \delta_3\bar{Z}$, $\beta_2 = \delta_2 - \delta_3\bar{X}$, and $\beta_3 = \delta_3$. It is true that the algebraic transformation that results from centering the variables will result in different coefficients and standard errors in the centered model compared to those in the uncentered model. However, this is because they measure different substantive quantities in each model and not because one model produces better estimates than the other. For example, $\beta_1$ in the uncentered model captures the marginal effect of a one-unit increase in X *when Z is zero*, while the equivalent coefficient $\delta_1$ in the centered model captures the marginal effect of a one-unit increase in $X_c$ (the same as a unit increase in X) *when Z is at its mean.* If the analyst were to actually calculate the marginal effect of a one-unit increase in X at the same level of Z from the estimates of the centered and uncentered models, he or she would obtain exactly the same marginal effect and measure of uncertainty. Given that the centered and uncentered models are algebraically equivalent, we can unequivocally state that centering does not change the statistical certainty of the estimated effects and, therefore, cannot really mitigate any multicollinearity issues that exist. Scholars should stop justifying the use of centered variables or the omission of constitutive terms in interaction models by claiming that this reduces multicollinearity. In the case of centered variables, this claim is simply unfounded. While it is true that the omission of constitutive terms may well reduce multicollinearity, it is probably unwise, almost always unnecessary, and can be justified only in extremely rare circumstances.

## 4 Do Not Interpret Constitutive Terms as Unconditional Marginal Effects

Scholars should refrain from interpreting the constitutive elements of interaction terms as unconditional or average effects—they are not. Notice that the reason why multiplicative interaction models capture the intuition behind conditional hypotheses so effectively is because they make the effect of the independent variable X on the dependent variable Y depend on some third variable Z. As a consequence, the coefficient on the constitutive term X *must not* be interpreted as the average effect of a change in X on Y as it can in a linear-additive regression model. As the above discussion should have made clear, the coefficient on X only captures the effect of X on Y when Z is zero. Similarly, it should be obvious that the coefficient on Z only captures the effect of Z on Y when X is zero.[9] It is, therefore, incorrect to say that a positive and significant coefficient on X (or Z) indicates that an increase in X (or Z) is expected to lead to an increase in Y. Many papers in the political science literature make this mistake and slip into interpreting the coefficients on constitutive terms as if they capture the unconditional effect of these variables.[10] Because the terminology used to discuss constitutive terms varies in the literature, we would like to reiterate that claiming that the coefficients on constitutive terms represent the "unconditional," "main," "independent," or "average" effect is simply wrong since it implies that these coefficients can be interpreted in the same way as they would in a linear-additive model. Just as we have come to recognize that coefficients in logit and probit models cannot be interpreted as unconditional marginal effects, we should recognize that the coefficients on constitutive terms in interaction models cannot be interpreted in this way either.

A brief example will illustrate some of the pitfalls that occur if this point is ignored. Imagine the following hypothesis:

> $H_2$: An increase in X is associated with an increase in Y when condition Z is present, but a decrease in Y when condition Z is absent.

This hypothesis can easily be tested with the same model outlined in Eq. (1). However, now our expectations are that $\beta_1 < 0$ and $\beta_1 + \beta_3 > 0$. This necessitates that $|\beta_3| > |\beta_1|$. Imagine the case in which our estimates were $\beta_1 = -0.5$ and $\beta_3 = 1.0$, such that $\beta_1 + \beta_3 = 0.5$. Although these estimates support our hypothesis very nicely, they say nothing about the average or unconditional effect of X on Y. The estimates only tell us that X is associated with a decrease in Y when condition Z is absent and an increase in Y when condition Z is present. To know the average effect of X on Y we would have to know how frequently condition Z is present in our sample. For example, if we knew that condition Z was present half of the time and absent the other half, then a one-unit increase in X would have no affect on Y on average, i.e., $0.5(-0.5) + 0.5(0.5) = 0$. Absent any knowledge about the distribution of condition Z, the only clear way to gauge the average effect of X on Y is to run an unconditional model in which X is not included in a multiplicative interaction term.

---

[9]It is a feature of multiplicative interaction models that they are symmetric. As a result, these models (and the data) cannot distinguish between the causal story where Z modifies the effect of X on Y and the causal story where X modifies the effect of Z on Y. It is up to the analyst to determine which of these causal stories is theoretically more accurate. This choice has consequences because it is likely to determine whether the analyst calculates the marginal effect of X on Y for different values of Z or the marginal effect of Z on Y for different values of X. While both of these quantities of interest are meaningful, one may be more theoretically or substantively useful. In our examples throughout this article, we have assumed that Z modifies the effect of X on Y and, therefore, that the quantity of interest is the marginal effect of X on Y.

[10]In fact, it is even possible to find this mistake in some econometric textbooks. Consider the following model:

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_2 X_3 + u. \tag{17}$$

Gujarati (2003, p. 284) mistakenly refers to $\beta_2$ as the marginal effect of $X_2$ on Y. He overlooks the fact that $\beta_2$ only indicates the marginal effect of $X_2$ on Y when $X_3$ is zero.

However, if we do expect the effect of X on Y to depend on Z, then the standard interpretation of this unconditional model has to be modified somewhat. This is because the coefficient on X in the unconditional linear-additive model is the weighted average of the conditional marginal effects in the interaction model. Consequently, the marginal effect of X in the unconditional model is sensitive to the distribution of the conditioning variable in the sample. For instance, if in the hypothetical example there were twice as many cases in which condition Z was absent as where Z was present, then the unconditional test would indicate that a one-unit increase in X would be associated with a $-0.17$ unit decrease in Y, i.e., $0.66(-0.5) + 0.33(0.5) = -0.17$. Note that the unconditional or average effect of X on Y is now different; it had been zero. This is the case even though the hypothesized conditional relationship has not altered. The point here is that one should use conditional tests to examine conditional hypotheses since the failure to do this produces an estimate of the unconditional relationship that reflects not only the underlying causal relationship between X and Y but also the distribution of the conditioning variable Z. In other words, it makes little sense to talk about the unconditional or average effect of X on Y when you have a conditional hypothesis.

## 5 Do Not Forget to Calculate Substantively Meaningful Marginal Effects and Standard Errors

Analysts should provide a substantively meaningful description of the marginal effects of the independent variables and the uncertainty with which they are estimated. As we just illustrated, the coefficient on the constitutive term X cannot be interpreted as an unconditional marginal effect since it indicates only the effect of a one-unit change in X on Y when the conditioning variable is zero. As a result, the interpretation of multiplicative interaction models differs in an important way from linear-additive regression models. To see this more clearly, compare the marginal effect of X in the following two models:

$$Y = \omega_0 + \omega_1 X + \omega_2 Z + \mu \tag{11}$$

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_2 XZ + \epsilon. \tag{12}$$

The marginal effect of X in Eq. (11) is $\frac{\partial Y}{\partial X} = \omega_1$. This is true of any linear-additive model such as this. In contrast, the marginal effect of X in a multiplicative interaction model such as Eq. (12) is calculated as

$$\frac{\partial Y}{\partial X} = \beta_1 + \beta_3 Z. \tag{13}$$

In other words, the additive model asserts that X has a constant effect on Y, while the interaction model asserts that the effect of a change in X on Y depends on the value of the conditioning variable Z.

Typically, traditional results tables report only model parameters (and perhaps some goodness of fit measures). In a linear-additive model, these are the quantities of interest since the coefficients and standard errors describe what we know about the marginal effect of each independent variable. However, this is not the case with multiplicative interaction models. Although we still want to know about the marginal effect of some independent variable X on Y in an interaction model ($\beta_1 + \beta_3 Z$), typical results tables will report only the marginal effect of X when the conditioning variable is zero, i.e., $\beta_1$. Similarly, these tables report only the standard error for this particular effect. As a result, the only inference that can be drawn in this situation is whether X has a significant effect on Y for the unique case in which $Z = 0$.[11] This can often be substantively uninformative, particularly if we never observe real-world situations in which the conditioning variable is actually zero. For example, if Z were government size, then we would know the effect of X only when government size was zero. Since there are no observations for which government size is zero, it is easy to see that $\beta_1$ would tell us nothing meaningful about the effect of X on Y. There is simply no way of knowing from the typical results table if X has a significant effect on Y when the conditioning variable is not zero.

The analyst cannot even infer whether X has a meaningful conditional effect on Y from the magnitude and significance of the coefficient on the interaction term either.[12] As we showed earlier, it is perfectly possible for the marginal effect of X on Y to be significant for substantively relevant values of the modifying variable Z even if the coefficient on the interaction term is insignificant. Note what this means. It means that one cannot determine whether a model should include an interaction term simply by looking at the significance of the coefficient on the interaction term. Numerous articles ignore this point and drop interaction terms if this coefficient is insignificant. In doing so, they potentially miss important conditional relationships between their variables.

The point here is that the typical results table often conveys very little information of interest because the analyst is not concerned with model parameters per se; he or she is primarily interested in the marginal effect of X on Y ($\beta_1 + \beta_3 Z$) for substantively meaningful values of the conditioning variable Z. While it is often possible to calculate the marginal effect of X for any value of Z from the typical results table using a little algebra, the problem is that it is not possible to do the same for the standard errors. This is because the relevant elements of the variance-covariance matrix necessary to calculate the standard error in Eq. (8) are rarely reported, i.e., $cov(\beta_1 \beta_3)$.

If a multiplicative interaction model is employed, it is nearly always necessary for the analyst to go beyond the traditional results table in order to convey quantities of interest such as the marginal effect of X on Y. If the modifying variable is dichotomous, this simply requires the analyst to present four numbers—the marginal effect of X when Z is 0 and when Z is 1, along with the two corresponding standard errors. If the conditioning variable is continuous, the analyst must work a little harder. A simple figure can be used to succinctly illustrate the marginal effect of X and the corresponding standard errors across a substantively meaningful range of the modifying variable(s). Below is an example taken from Golder (forthcoming) addressing the effect of presidential elections on legislative fragmentation.[13] The central hypothesis is the following:

[11]Though accurate, even this inference may not be sensible. As we showed earlier, this is because it is possible to manipulate both the size and significance of the coefficients on constitutive terms in multiplicative interaction models that employ variables without a natural zero by arbitrarily rescaling the variables. In such models, we might not want to read too much into these coefficients.

[12]Ai and Norton (2003) illustrate that the inferences that can be drawn from the coefficient on the interaction term are even more limited when analysts employ nonlinear models such as logit and probit. Somewhat disturbingly, they found in a review of the 13 economics journals listed on JSTOR that none of the 72 articles published between 1980 and 1999 that included nonlinear models with explicit interaction terms interpreted this coefficient correctly.

[13]This example employs a single modifying variable and a continuous dependent variable. However, it is just as easy to use a similar figure for models with multiple modifying variables (Brambor et al. 2004; Clark 2003) or for models with limited dependent variables (Golder 2005).

**Table 1** The impact of presidential elections on the effective number of electoral parties. Dependent variable: Effective number of electoral parties

| *Regressor* | *Model* |
|---|---|
| Proximity | −3.44** (0.49) |
| PresidentialCandidates | 0.29* (0.07) |
| Proximity*PresidentialCandidates | 0.82** (0.22) |
| Controls | —— |
| Constant | 3.01** (0.33) |
| $R^2$ | 0.34 |
| N | 522 |

*$p < 0.05$; **$p < 0.01$ (two-tailed). Control variables not shown here.
Robust standard errors clustered by country in parentheses.

$H_3$: Temporally-proximate presidential elections will reduce the effective number of legislative parties if and only if the number of presidential candidates is sufficiently low.

Using data from 522 legislative elections from 1946 to 2000, the author tests his hypothesis with the following model:

$$\begin{aligned} ElectoralParties = {} & \beta_0 + \beta_1 Proximity + \beta_2 PresidentialCandidates \\ & + \beta_3 Proximity{*}PresidentialCandidates \\ & + \beta_4 Controls + \epsilon, \end{aligned} \tag{14}$$

where *ElectoralParties* and *PresidentialCandidates* measure the effective number of electoral parties and presidential candidates, *Proximity* is a continuous measure of the temporal proximity of presidential and legislative elections, and *Controls* refers to a series of control variables that include social heterogeneity and electoral system characteristics. Results from this model are shown in Table 1.

The results in Table 1 indicate that temporally proximate presidential elections have a significant reductive effect on the effective number of electoral parties when there are no presidential candidates ($\beta_1$ is negative). Note that this is substantively meaningless since there are no cases in which there are presidential elections and no presidential candidates. We know from the fact that the coefficient on *Proximity*PresidentialCandidates* is positive that this reductive effect declines as the number of presidential candidates increases. However, there is no way of knowing from the information in Table 1 what the impact of temporally proximate presidential elections is when the number of candidates is greater than zero. Recognizing that this traditional table of results can only throw limited light on his hypothesis, the author presents a simple figure (Fig. 3) that graphically illustrates how the marginal effect of temporally proximate presidential elections changes across the observed range of presidential candidates.[14]

The solid sloping line in Fig. 3 indicates how the marginal effect of temporally-proximate presidential elections changes with the number of presidential candidates. Any particular point on this line is $\frac{\partial ElectoralParties}{\partial Proximity} = \beta_1 + \beta_3 PresidentialCandidates$. 95%

[14]It is not always the case that the observed range of the modifying variable is the most substantively meaningful or informative range to employ in these figures. The specific question being addressed by the analyst should determine the relevant range of the modifying variable.

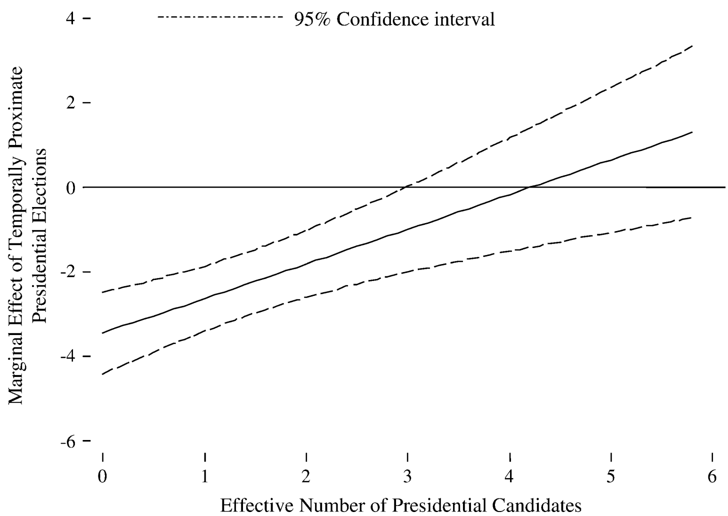


**Fig. 3** The marginal effect of temporally proximate presidential elections on the effective number of electoral parties.

confidence intervals around the line allow us to determine the conditions under which presidential elections have a statistically significant effect on the number of electoral parties—they have a statistically significant effect whenever the upper and lower bounds of the confidence interval are both above (or below) the zero line. It is easy to see that temporally proximate presidential elections have a strong reductive effect on the number of electoral parties when there are few presidential candidates. As predicted, this reductive effect declines as the number of presidential candidates increases. Once there are more than 2.9 effective presidential candidates, presidential elections no longer have a significant reductive impact on legislative fragmentation. If one examines only those countries in which presidential elections take place, roughly 60% of legislative elections between 1946 and 2000 have occurred when there are fewer presidential candidates than this. A conscientious analyst will report the percentage of the sample that falls within the region of significance as the author does here so that the reader can better judge the substantive implications of the results. The point is that simply having a significant marginal effect across some values of the modifying variable is not particularly interesting if real-world observations rarely fall within this range.

This example illustrates that it is extremely difficult and often impossible to evaluate conditional hypotheses using only the information provided in traditional results tables. Moreover, it shows that it is relatively simple to evaluate the marginal effect of some variable X across a substantively meaningful range of its modifying variable(s) with a simple figure. Our hope is that analysts who use interaction models will move beyond traditional results tables when evaluating their conditional hypotheses. Here, we have presented just one way in which analysts might want to go about doing this.

The examples that we have used in this article have all involved a continuous dependent variable. However, it is important to note that all of the points that we have made regarding the specification and interpretation of interaction models are directly applicable to situations in which the analyst has a limited dependent variable. Nothing of importance changes. The analyst who wishes to test a conditional hypothesis should still use an interaction model. This requires explicitly modeling the hypothesized conditionality by

**Table 2** Survey of articles with interaction models (APSR, AJPS, JOP, 1998–2002)

| *Recommendation* | *Yes* | *No* | *Total* |
|---|---|---|---|
| Include all constitutive terms | 107 (69%) | 49 (31%) | 156 |
| Interpret constitutive terms correctly* | 38 (38%) | 63 (62%) | 101 |
| Provide range for marginal effect | 86 (55%) | 70 (45%) | 156 |
| Provide measure of uncertainty | 34 (22%) | 122 (78%) | 156 |

*Only 101 articles interpreted constitutive terms.

including an interaction term just as one would if one had a continuous dependent variable. Some scholars have argued that it is not necessary to include an interaction term because nonlinear models such as logit and probit implicitly force the effect of all the independent variables to depend on each other anyway (Berry and Berry 1991). However, this line of reasoning is wrong (Frant 1991; Nagler 1991; Gill 2001). Consider an additive probit model, i.e., no explicit interaction term:

$$E[Y] = \Phi(\gamma_0 + \gamma_1 X + \gamma_2 Z) = \Phi(\cdot). \tag{15}$$

The marginal effect of X is $\partial\Phi(\cdot)/\partial X = \gamma_1\Phi'(\cdot)$. It is easy to see that the marginal effect of X depends on the value of the other independent variables even in this additive model. However, note that this dependence occurs whether the analyst's hypothesis is conditional or not—it is just part and parcel of deciding to use a nonlinear model such as probit; it is always there. If one wants to test a conditional hypothesis in a meaningful way, then the analyst has to include an explicit interaction term as in the model below.

$$E[Y] = \Phi(\beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ) = \Phi(\cdot) \tag{16}$$

The marginal effect of X is now $\frac{\partial\Phi(\cdot)}{\partial X} = (\beta_1 + \beta_3 Z)\Phi'(\cdot)$. Just as with interaction models with continuous dependent variables, analysts employing interaction models with limited dependent variables should also include all constitutive terms. They should also remember not to interpret the coefficients on the constitutive terms as if they are unconditional or average marginal effects. Finally, they should calculate substantively meaningful marginal effects and standard errors. In other words, nothing changes if the dependent variable is limited.[15]

## 6 A Survey of the Literature

To many, the points that we have made will seem like common sense. We agree. Some will say that they are so obvious that surely everyone already knows how to deal with interaction models. Here, however, we beg to differ. While it is true that much of what we have presented is well established in the political methodology literature, a systematic examination of three leading nonspecialized political science journals (*American Political Science Review, Journal of Politics, American Journal of Political Science*) from 1998 to 2002 indicates that the points made in this article are ignored by political science practitioners using interaction models. Table 2 provides summary results from our survey of the literature.

[15] All of our examples in this article have also included a single modifying variable. However, we should reiterate that just as our points concerning the specification and interpretation of interaction models apply to models with limited or continuous dependent variables, they also apply equally well whether the analyst has one modifying variable (Z modifies the effect of X on Y) or multiple modifying variables (Z and J both modify the effect of X on Y).

During the five-year period from 1998 to 2002, 156 articles employed interaction models of one variety or another.[16] While a large proportion of these articles are in American politics, all of the subfields in political science (comparative politics, international politics, comparative and international political economy, etc.) are represented, with the exception of political theory. We found that 49 of the 156 articles presented results from models in which at least one constitutive term was omitted. In several cases, multiple terms were omitted. It is not the case that these analyses are necessarily wrong or that incorrect inferences have been drawn; after all, we noted earlier that there is a specific (if limited) set of cases in which the omission of constitutive terms may be appropriate. All that we wish to suggest is that there is a potential for bias in these articles and that this is significant given that several of them have already generated substantial research agendas.

The inclusion of all constitutive terms is primarily a specification issue. What if we turn to the interpretation of results from interaction models? Of the 101 articles that actually interpreted one or more of the constitutive terms, 63 interpreted them as unconditional or average effects. While we cannot say with any certainty that any inferential errors were made by omitting constitutive terms without re-running the analyses in the 49 articles just mentioned, we can say with absolute certainty that interpreting constitutive terms as unconditional or average effects is wrong. It is simply not appropriate to interpret the results of interaction models as if they came from a linear-additive model.

Our final point was that analysts should calculate marginal effects and standard errors across a substantively meaningful range of the modifying variable. Here, we employed a very liberal approach to coding the articles. In order to comply with this point, all the authors had to do was calculate the marginal effect of an independent variable for at least one value of the modifying variable other than zero and provide a measure of the uncertainty with which it was estimated. In fact, we coded articles that reported predicted probabilities or outcomes under two or more different situations as having complied with this point, even though this is not strictly speaking an illustration of a marginal effect or arguably even the quantity of interest. Despite our liberal approach to coding, we found that only 86 articles reported marginal effects at multiple values of the modifying variable. Of these, most reported marginal effects at only two of the possible values of the modifying variable. More significantly, only 34 articles in our survey actually provided any measure of the uncertainty with which these marginal effects were estimated (other than when the modifying variable was zero). In other words, only 22% of the articles in our survey are in a position to plausibly say whether some variable X had a statistically significant effect on some Y when the modifying variable was not 0. Of those articles that actually did provide suitable measures of uncertainty, almost none went on to discuss the substantive significance of their results by indicating how many real-world observations fell in the range of statistical significance. The situation would not be so troubling if it were a relatively small subset of the 156 articles that accounted for the majority of cases in which the points that we have made were ignored. However, this is not the case. As we mentioned in the introduction, only 16 (10%) articles actually included all constitutive

[16]Some scholars test conditional hypotheses by splitting their data into categories (such as male and female or north and south) across which the effect of some independent variable X is supposed to differ. Instead of using explicit interaction terms, they simply run separate regressions on each of these categories. While this is a perfectly reasonable way to test conditional hypotheses, there is no real gain in terms of interpretability and there will always be a loss of efficiency due to the smaller sample sizes. We omit analyses of this type from our literature survey because we wish to focus purely on those models with an explicit interaction term. A complete list of the articles in our survey can be found on the Web site associated with this article.

terms, did not make mistakes interpreting these terms, and calculated substantively meaningful marginal effects and standard errors.[17]

The overwhelming message from Table 2 is that although we may well have thought we knew how to use interaction models, we obviously did not. In fact, we believe that our survey suggests that our well-established "common-sense" points are so rarely reflected in common practice that analysts should critically re-evaluate, and where necessary re-specify, models employing interaction terms before using their results as the basis for future research. Substantively different conclusions from those in the original analyses often arise when this is done. As an illustration of this we briefly summarize the results from three replications that we have recently conducted during our research on party systems and electoral institutions.[18]

In an award-winning article in the *American Political Science Review*, Boix (1999) examines the factors that determine electoral system choice in advanced democracies. He makes two main conclusions. First, ethnic or religious fragmentation encourages the adoption of proportional representation in small and medium-sized countries (621). He draws this conclusion based on a model that includes an interaction term between ethno-religious fragmentation and country size. However, he does not include either of the constitutive terms. When these terms are included, there is no longer any evidence that ethno-religious fragmentation ever affects the adoption of proportional representation. The second conclusion is that countries are more likely to shift to proportional representation when the proportion of socialist votes and the effective number of non-socialist parties are both large. This conclusion comes from a model in which there is an interaction term between the strength of socialist parties and the number of non-socialist parties but no constitutive terms. In this case, the coefficient on the interaction term from the primary model remains significant (albeit only at the 90% level now) once the constitutive terms are included, but its magnitude increases by 340%. Thus, the original analysis considerably underestimates the interactive effect of these two variables. Moreover, the failure to include constitutive terms means that the predicted electoral thresholds reported by Boix are off by up to 80%.

In an article in the *Journal of Politics*, Samuels (2000) examines the relative impact of presidential and gubernatorial coattails on the composition of the Brazilian party system. Theory would suggest that temporally proximate presidential and gubernatorial elections should exert a reductive effect on the number of electoral lists in legislative elections. However, this reductive effect should decline (and may become positive) as the number of presidential and gubernatorial candidates increases. Samuels argues that the unusual importance of the governor for office-seeking candidates in Brazilian legislative elections means that we should observe a gubernatorial coattails effect but not a presidential coattails effect in Brazil. This would help to explain why the party system at the national level is highly fragmented (6.3 effective parties), while the party system at the state level is

[17]We remind the reader that our first point was that analysts should include interaction terms whenever they have a conditional hypothesis. In our survey of the literature, we did not take account of articles that should have employed an interaction model but did not. Kam and Franzese (2003, p. 9) estimate that there are at least as many articles in the political science literature making conditional hypotheses that do not include interaction terms as there are those that do include interaction terms. If this is true (and we think it probably is), then the results in Table 2 seriously underestimate the potential errors in the existing literature regarding conditional hypotheses.

[18]The results from the replications will be made publicly available at http://homepages.nyu.edu/~mrg217 on publication. We encourage analysts to conduct replications of the other studies in our survey of the literature. This may be a good assignment for a first or second class in quantitative methods. We would be happy to post these replications on the Web site associated with this article.

more concentrated (only 3.3 effective parties). The results from three models seem to support his conjecture. However, Samuels draws conclusions from an interaction model that omits constitutive terms. Once these omitted terms are included, none of the coefficients on the variables of interest are significant at the 90% level.[19] Plots of the marginal effect of gubernatorial elections on the number of electoral lists across the observed range of the modifying variable from all three models indicate that gubernatorial elections never exert a coattails effect. While two of the three models indicate that there is no presidential coattails effect either, one suggests that temporally proximate presidential elections will increase the number of electoral lists if the number of presidential candidates is sufficiently high. Thus, contrary to the conclusions reached by Samuels, the evidence from a fully specified model indicates that if there is a coattails effect in Brazilian elections then it is a presidential one and not a gubernatorial one. This indicates that gubernatorial coattails cannot explain why the state party system in Brazil is so much less fragmented than the national party system.

Finally, in a recent article in the *American Political Science Review*, Mozaffar et al. (2003) examine how ethnopolitical cleavages and electoral institutions interact to determine the number of parties in African democracies. Using results from an interaction model, they conclude that "high ethnopolitical fragmentation is likely to reduce the number of parties" (p. 381), that "district magnitude substantially reduces the number of electoral and legislative parties" (p. 387), that the effect of presidential elections on the number of parties "depends on the degree of fractionalization of presidential elections" (p. 384), and that there is an "almost universal tendency of presidential regimes to constrict the structure of party systems" (p. 387). At least two of these conclusions are quite shocking given that they run directly counter to almost 50 years of research on party systems suggesting that ethnic fragmentation and district magnitude actually increase the number of parties (Duverger 1954). The conclusions from this particular article are open to question, though, since the authors draw inferences from a number of multiplicative interaction models in which they omit multiple constitutive terms, interpret constitutive terms as unconditional marginal effects, and fail to calculate marginal effects and standard errors over a sufficiently large range of their modifying variables. We find that by correcting these practices the only conclusion made by the authors that can be sustained is that presidential elections tend to reduce the number of parties. Far from being exceptional as the authors' results suggest, African party systems seem to respond to specific institutional and sociological factors such as district magnitude, ethnic fragmentation, and presidentialism in the same ways as party systems in more established democracies elsewhere in the world.

Our goal here is not to point fingers at anyone in particular.[20] As our survey of the literature suggests, the mistakes noted in these replications are common among practitioners of multiplicative interaction models across the discipline. As a result, we should all hold our hands up and take some responsibility. What these replications show, though, is that ignoring the points that we have made in this article can have serious consequences for drawing inferences—the substantive conclusions supported by the data can be the exact opposite of those reported by the analyst. While these particular

[19]Samuels includes the log of average district magnitude as a control variable. This has the predicted sign and is significant at the 90% level.

[20]If we were to do so, at least one of us would have to point fingers at himself as well. For example, Clark (1998) fails to include all of the constitutive terms in his interaction model of political business cycles in OECD economies. In subsequent work, Clark (2003) finds that his substantive conclusions change when he estimates a fully specified model.

replications reflect our current interest in party systems and electoral institutions, it should be noted that similar replications in the area of comparative/international political economy, international relations, and comparative politics have also produced results that run directly counter to the conclusions presented in the original analyses (Clark 2003; Golder 2003; Braumoeller 2004).

## 7 Conclusion

In this article we have shown several ways in which analysts can improve empirical analyses that employ multiplicative interaction models. Because arguments about strategic behavior, institutional analysis, and cross-national comparisons are typically conditional in nature, empirical tests of such arguments should include some form of interaction model. Many authors, nevertheless, test such conditional arguments in linear-additive models. The good news is that scholars are becoming increasingly aware of this mistake and are now frequently including interactions in their analyses (Franzese 2003b). The bad news is that many of these scholars ignore well-established methodological points concerning the specification and interpretation of interaction models. For example, they omit constitutive terms, make invalid inferences about marginal effects, and fail to provide substantively meaningful estimates of marginal effects and their standard errors. Consequently, large portions of our empirical knowledge about politics are based on models that are misspecified or misinterpreted in at least one of the four ways we identify in this article. For this reason, we believe considerable progress in our understanding of the political world can occur if scholars recognize the points made in this article. While more computationally complex and impressive methodological advances are to be welcomed, their contributions will be diminished to the extent that we, as a discipline, ignore the basics.

## References

Ai, Chunrong, and Edward Norton. 2003. “Interaction Terms in Logit and Probit Models.” *Economics Letters* 80:123–129.

Aiken, Leona, and Stephen West. 1991. *Multiple Regression: Testing and Interpreting Interactions.* London: Sage.

Allison, Paul D. 1977. “Testing for Interaction in Multiple Regression.” *American Journal of Sociology* 83:144–153.

Bernhardt, Irwin, and Bong S. Jung. 1979. “The Interpretation of Least Squares Regression with Interaction or Polynomial Terms.” *Review of Economics and Statistics* 61:481–483.

Berry, Frances Stokes, and William Berry. 1991. “Specifying a Model of State Policy Innovation.” *American Political Science Review* 85:573–579.

Boix, Carles. 1999. “Setting the Rules of the Game: The Choice of Electoral Systems in Advanced Democracies.” *American Political Science Review* 93:609–624.

Brambor, Thomas, William Clark, and Matt Golder. 2005. “Are African Party Systems Different?” Unpublished manuscript, New York University.

Braumoeller, Bear. 2004. “Hypothesis Testing and Multiplicative Interaction Terms.” *International Organization* 58:807–820.

Busemeyer, Jerome, and Lawrence Jones. 1983. “Analysis of Multiplicative Combination Rules When the Causal Variables Are Measured with Error.” *Psychological Bulletin* 93:549–562.

Clark, William Roberts. 2003. *Capitalism, Not Globalism: Capital Mobility, Central Bank Independence and the Political Control of the Economy.* Ann Arbor: University of Michigan Press.

Clark, William Roberts, and Usha Nair Reichert. 1998. “International and Domestic Constraints on Political Business Cycles in OECD Economies.” *International Organization* 52:87–120.

Cleary, Paul, and Ronald Kessler. 1982. “The Estimation and Interpretation of Modifier Effects.” *Journal of Health and Social Behavior* 23:159–169.

Cox, D. R. 1984. "Interaction." *International Statistical Review* 52:1–31.

Duverger, Maurice. 1954. *Political Parties.* New York: Wiley.

Frant, Howard. 1991. "Specifying a Model of State Policy Innovation." *American Political Science Review* 85:571–573.

Franzese, Robert. 2003a. "Multiple Hands on the Wheel: Empirically Modeling Partial Delegation and Shared Policy Control in the Open and Institutionalized Economy." *Political Analysis* 11:445–474.

Franzese, Robert. 2003b. "Quantitative Empirical Methods and Context Conditionality." *APSA-CP: Newsletter of the Organized Section on Comparative Politics of the American Political Science Association* 14:20–24.

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

Gill, Jeff. 2001. "Interpreting Interactions and Interaction Hierarchies in Generalized Linear Models: Issues and Applications." Presented at the Annual Meeting of the American Political Science Association, San Francisco.

Golder, Matt. 2003. "Electoral Institutions, Unemployment, and Extreme Right Parties: A Correction." *British Journal of Political Science* 33:525–5346.

Golder, Matt. Forthcoming. "Presidential Candidates and Legislative Fragmentation." Presented at the Annual Meeting of the American Political Science Association, Philadelphia.

Golder, Sona Nadenichek. 2005. "Pre-electoral Coalitions in Comparative Perspective: A Test of Existing Hypotheses." *Electoral Studies.*

Greene, William. 2003. *Econometric Analysis.* New Jersey: Prentice Hall.

Griepentrog, Gray, J. Michael Ryan, and L. Douglas Smith. 1982. "Linear Transformations of Polynomial Regression Models." *American Statistician* 36:171–174.

Gujarati, Damodar. 2003. Basic Econometrics. New York: McGraw Hill.

Jaccard, James, and Choi Wan. 1995. "Measurement Error in the Analysis of Interaction Effects between Continuous Predictors Using Multiple Regression: Multiple Indicator and Structural Equation Approaches." *Pyschological Bulletin* 117:348–357.

Jaccard, James, Choi Wan, and Robert Turrisi. 1990. "The Detection and Interpretation of Interaction Effects between Continuous Variables in Multiple Regression." *Multivariate Behavioral Research* 25:467–478.

Kam, Cindy, and Robert Franzese. 2003. "Modeling and Interpreting Interactive Hypotheses in Regression Analysis: A Brief Refresher and Some Practical Advice." Unpublished manuscript, University of Michigan.

Mozaffar, Shaheen, James R. Scarritt, and Glen Galaich. 2003. "Electoral Institutions, Ethnopolitical Cleavages and Party Systems in Africa's Emerging Democracies." *American Political Science Review* 97:379–390.

Nagler, Jonathan. 1991. "The Effect of Registration Laws and Education on U.S. Voter Turnout." *American Political Science Review* 85:1393–1405.

Nagler, Jonathan. 1994. "Scobit: An Alternative to Logit and Probit." *American Journal of Political Science* 38:230–255.

Norton, Edward, Hua Wang, and Chunrong Ai. 2004. "Computing Interaction Effects and Standard Errors in Logit and Probit Models." *STATA Journal* 4:103–116.

Rosnow, Ralph, and Robert Rosenthal. 1989. "Definition and Interpretation of Interaction Effects." *Psychological Bulletin* 105:143–146.

Rosnow, Ralph, and Robert Rosenthal. 1991. "If You're Looking at the Cell Means, You're Not Looking at *Only* the Interaction (Unless All Main Effects Are Zero)." *Psychological Bulletin* 110:574–576.

Samuels, David. 2000. "The Gubernatorial Coattails Effect: Federalism and Congressional Elections in Brazil." *Journal of Politics* 62:240–253.

Smith, Mark. 2000. "The Contingent Effects of Ballot Initiatives and Candidate Races on Turnout." *American Journal of Political Science* 45:700–706.

Wright, Gerald. 1976. "Linear Models for Evaluating Conditional Relationships." *American Journal of Political Science* 2:349–373.