4

THE MEANING, USE, AND ABUSE OF SOME COMMON GENERAL-PRACTICE RULES

Having discussed formulation of interactive hypotheses, and interpretation and presentation of effects, we turn now to clarify some general-practice rules often applied in the social sciences.

Colinearity and Mean-Centering the Components of Interaction Terms

One common concern regarding the estimation of interactive models is the (*multi*)*colinearity*, or high correlation, among independent variables induced by multiplying regressors together to create additional regressors. Colinearity, as social scientists well know, induces large standard errors, reflecting our low confidence in the individual coefficients estimated on these highly correlated factors. What is sometimes forgotten is that these large standard errors are *correctly* large; the effect of x controlling for other terms (i.e., holding them constant) *is* hard to determine with much certainty if x and other terms correlate highly. These large standard errors accurately reflect our high degree of uncertainty in these conditions. These perhaps unfortunate, but very real, facts regarding colinearity led Althauser (1971), for example, to argue against the use of interactive terms at all. However, to omit interactions simply because including them invites a greater degree of uncertainty in parameter estimates is to misspecify intentionally our theoretical propositions. This assures at least inefficiency but most likely induces bias due to standard omitted-variable-bias considerations: namely, if the omitted factor, xz, (partially) correlates with the included factor, x, and (partially) correlates with the dependent variable, y, then bias results. The sign and magnitude of the bias are given by the product of these two partial coefficients.

Scholars therefore struggled valiantly for some technical artifice to reduce interaction-induced colinearity. However, the problem of colinearity is "too little information." As such, the only routes around the problem available to researchers are to ask the data questions that require less information (e.g., only first-order questions, like those in table 10 or table 12) or to obtain more information by drawing new data (preferably less correlated, but more data will help regardless) or by relying more heavily upon the theoretical arguments/assumptions to specify models that ask more precise questions of the data than do generic linear-interactive models (e.g., Franzese 1999, 2002, 2003a).

Scholars have instead devoted inordinate attention to illusory colinearity "cures." The most commonly prescribed "cure" is to "center" the variables (i.e., subtract their sample means or "mean-deviate" them) that comprise the interactive terms. Smith and Sasaki (1979) offered centering as a technique that would improve substantive interpretation of the individual coefficients, and we agree that it might facilitate interpretation in some substantive contexts. Tate (1984) argued that, although centering should not change the substantive effects (actually, it *will* not: see the discussion that follows), it "may improve conditioning through reduction of colinearity" (253). Others, including Morris, Sherman, and Mansfield (1986) and Dunlap and Kemery (1987), recommend centering less circumspectly. The centering technique of Cronbach (1987) has attained considerable acceptance in social science, perhaps due to the promotion of it by Jaccard, Turrisi, and Wan (1990). Unfortunately, Cronbach's clarification on the extremely limited value of centering seems less widely known.

To be sure, the centering procedure of Cronbach (1987) is harmless; however, it also offers no help against the "too little information" problem of colinearity, if understood correctly. Our concern is that centering seems widely misunderstood and misinterpreted. Some existing scholarly research claims, wrongly, that centering helps evade colinearity in some manner that actually produces more certain effect estimates. Centering adds no new information of any sort to the empirical estimation, and so it cannot possibly produce more precise estimates. Centering merely changes the substantive question to which the coefficients and t-tests of those coefficients refer.

Consider this standard linear-interactive model:

$$y = \beta_0 + \beta_x x + \beta_z z + \beta_{xz} xz + \varepsilon \tag{31}$$

Cronbach (1987) suggested subtracting the sample means from each of the independent variables involved in the interaction and multiplying the resulting demeaned variables together for the interaction term. The mean-centered model, then (using γ to represent coefficient values resulting from use of the centered data), is as follows:

$$y = \gamma_{0^*} + \gamma_{x^*} x^* + \gamma_{z^*} z^* + \gamma_{x^*z^*} x^* z^* + \varepsilon^* \tag{32}$$

where $x^* = x - \bar{x}$ and $z^* = z - \bar{z}$.

Cronbach (1987) argued that rescaling the variables thusly could insure against computational errors—that is, errors that are literally computational: deriving from inescapable rounding errors in translating from computer binary to human base-ten—that severe colinearity might induce.[1] Cronbach (1987) also noted that centered and noncentered models "are logically interchangeable, and under most circumstances it makes no difference which is used" (415). Given the many thousands of times computing precision has increased since Cronbach's writing, the computational concern has no current practical relevance in social science, and so it now makes no difference under any circumstances.

Because centering does not affect the substance of any empirical estimation in any way, because it will not affect the computational algorithms of any modern statistical software, and because it is so widely misunderstood in the field, we join Friedrich (1982), Southwood (1978),

1. The computational issue here involves matrix inversion, namely, the $(\mathbf{X}'\mathbf{X})^{-1}$ in OLS formulas for coefficient and standard-error estimates, some of whose columns (i.e., independent variables) may correlate nearly perfectly. If columns of $\mathbf{X}$ correlate perfectly, the determinant of $(\mathbf{X}'\mathbf{X})$, which appears in the denominator of the formula for $(\mathbf{X}'\mathbf{X})^{-1}$, is zero. Division by zero is, of course, impossible; therefore, obtaining distinct coefficient estimates (and thus standard errors) when (some) columns of $\mathbf{X}$ correlate perfectly is impossible. All modern regression software warns of perfect colinearity when it obtains a zero determinant before allowing the computer to crash trying to divide by zero. Most warn of near-perfect colinearity well short of obtaining identically zero for that critical determinant, that is, well short of perfect colinearity, because the translation from the base-ten data matrix to the binary of computers involves rounding error. When something near zero appears in a denominator and contains slight rounding error, the final answer could exhibit massive error. This is the concern that Cronbach raised. The multiplicative terms in interactive regressions, he feared, could be near enough to perfect colinearity to cause severe binary-to-base-ten rounding-error problems. However, since his writing, computers have become many thousands of times more exact in their binary calculations' approximation to base ten, meaning that even this computational concern is no longer present in any practical social-science context.

and others in strongly advising the abandonment of the practice or, at least, far greater care in interpreting and presenting the results following its implementation. To clarify what centering does to the numeric and substantive estimates of an interactive analysis, which is something and nothing, respectively, consider again our basic linear-interaction model and its centered version, which appear in equations (31) and (32), respectively. Starting from equation (32), and substituting terms, we see that

$$y = \gamma_{0^*} + \gamma_{x^*}(x - \bar{x}) + \gamma_{z^*}(z - \bar{z}) + \gamma_{x^*z^*}(x - \bar{x})(z - \bar{z}) + \varepsilon^* \tag{33}$$

$$y = \gamma_{0^*} + \gamma_{x^*}x - \gamma_{x^*}\bar{x} + \gamma_{z^*}z - \gamma_{z^*}\bar{z} + \gamma_{x^*z^*}xz - \gamma_{x^*z^*}\bar{x}z - \gamma_{x^*z^*}x\bar{z} + \gamma_{x^*z^*}\overline{xz} + \varepsilon^*$$

$$y = (\gamma_{0^*} - \gamma_{x^*}\bar{x} - \gamma_{z^*}\bar{z} + \gamma_{x^*z^*}\overline{xz}) + (\gamma_{x^*} - \gamma_{x^*z^*}\bar{z})x + (\gamma_{z^*} - \gamma_{x^*z^*}\bar{x})z + \gamma_{x^*z^*}xz + \varepsilon^* \tag{34}$$

Comparing the centered equation in (34) with the original model in (31) highlights the exact correspondence of results between the centered and uncentered regression models:

$$\begin{aligned}
\beta_0 &= \gamma_{0^*} - \gamma_{x^*}\bar{x} - \gamma_{z^*}\bar{z} + \gamma_{x^*z^*}\overline{xz} \\
\beta_x &= \gamma_{x^*} - \gamma_{x^*z^*}\bar{z} \\
\beta_z &= \gamma_{z^*} - \gamma_{x^*z^*}\bar{x} \\
\beta_{xz} &= \gamma_{x^*z^*}
\end{aligned}$$

Collecting terms, we see that the first parenthetical expression in equation (34) contains its set of constant terms and thus equals the intercept, β_0, from (31). The second parenthetical expression in (34) is its ultimate coefficient on x, which is equal to β_x from (31), and the third parenthetical expression is the ultimate coefficient on z in (34), which equals β_z in (31). The fourth term is the coefficient on xz in each model. Trivially, since the right-hand-side models are mathematically interchangeable, the estimated residuals and therefore the estimated residual variance from the centered and uncentered models are also identical.

As we explained previously, researchers' common troubles arise when they confuse coefficients with effects. We know, for example, that the marginal effect of x on y in equation (31) would be $\partial y/\partial x = \beta_x + \beta_{xz}z$. The marginal effect of x^* on y given equation (32) would be $\partial y/\partial x^* = \gamma_{x^*} + \gamma_{x^*z^*}z^*$. Since $\beta_x = \gamma_{x^*} - \gamma_{x^*z^*}\bar{z}$, we can express $\gamma_{x^*} = \beta_x + \gamma_{x^*z^*}\bar{z}$. Therefore

$$\partial y/\partial x^* = \beta_x + \gamma_{x^*z^*}\bar{z} + \gamma_{x^*z^*}z^*$$

Then, given that $z^* = z - \bar{z}$, we have

$$\partial y/\partial x^* = \beta_x + \gamma_{x^*z^*}\bar{z} + \gamma_{x^*z^*}z - \gamma_{x^*z^*}\bar{z} = \beta_x + \gamma_{x^*z^*}z$$

Finally, since $\beta_{xz} = \gamma_{x^*z^*}$, we conclude

$$\partial y/\partial x^* = \beta_x + \beta_{xz}z = \partial y/\partial x$$

Stated directly, the point is obvious: the effect of a marginal increase in the centered version of x is identical to the effect of a marginal increase in uncentered x. The same identity applies to the effects of z, of course. We reiterate: centering does not change the estimated *effects* of any variables.

Further, the estimated variance-covariances (i.e., standard errors, etc.) of those effects are also identical. Thus, the estimated statistical certainty of the estimated effects is also unchanged by centering. For the uncentered data, $V(\partial\hat{y}/\partial x) = V(\hat{\beta}_x) + z^2V(\hat{\beta}_{xz}) + 2zC(\hat{\beta}_x,\hat{\beta}_{xz})$. Using the mean-centered model:

$$V(\partial\hat{y}/\partial x^*) = V(\hat{\gamma}_{x^*}) + (z^*)^2V(\hat{\gamma}_{x^*z^*}) + 2z^*C(\hat{\gamma}_{x^*},\hat{\gamma}_{x^*z^*})$$

Substituting $\hat{\gamma}_{x^*} = \hat{\beta}_x + \hat{\gamma}_{x^*z^*}\bar{z}$ and $\hat{\gamma}_{x^*z^*} = \hat{\beta}_{xz}$

$$V(\partial\hat{y}/\partial x^*) = V(\hat{\beta}_x + \hat{\beta}_{xz}\bar{z}) + (z^*)^2\, V(\hat{\beta}_{xz}) + 2z^*C(\hat{\beta}_x + \hat{\beta}_{xz}\bar{z},\, \hat{\beta}_{xz})$$

$$\begin{aligned} V(\partial\hat{y}/\partial x^*) = {} & V(\hat{\beta}_x) + \bar{z}^2\, V(\hat{\beta}_{xz}) + 2\bar{z}C(\hat{\beta}_x,\hat{\beta}_{xz}) + (z^*)^2\, V(\hat{\beta}_{xz}) \\ & + 2z^*C(\hat{\beta}_x + \hat{\beta}_{xz}\bar{z},\hat{\beta}_{xz}) \end{aligned}$$

Rearranging terms and substituting $z^* = z - \bar{z}$:

$$\begin{aligned} V(\partial\hat{y}/\partial x^*) = {} & V(\hat{\beta}_x) + \bar{z}^2V(\hat{\beta}_{xz}) + 2\bar{z}C(\hat{\beta}_x,\hat{\beta}_{xz}) \\ & + (z - \bar{z})^2V(\hat{\beta}_{xz}) + 2(z - \bar{z})C(\hat{\beta}_x,\hat{\beta}_{xz}) + 2(z - \bar{z})\bar{z}V(\hat{\beta}_{xz}) \end{aligned}$$

$$V(\partial\hat{y}/\partial x^*) = V(\hat{\beta}_x) + z^2V(\hat{\beta}_{xz}) + 2zC(\hat{\beta}_x,\hat{\beta}_{xz}) = V(\partial\hat{y}/\partial x)$$

The variances of the estimated marginal effects of the centered x and of the uncentered x are identical. The same holds for the variances of the estimated marginal effects of z and mean-centered z, of course. As with the coefficients, the numeric values of the elements in the variance-covariance matrices for the coefficients using uncentered and centered data will naturally differ from each other, but exact correspondence in the estimated effects and the estimated variances of effects can be derived through algebraic manipulation of these values. As an example, recall that $\beta_x = \gamma_{x^*} - \gamma_{x^*z^*}\bar{z}$. This implies that $V(\hat{\beta}_x) = V(\hat{\gamma}_{x^*} - \hat{\gamma}_{x^*z^*}\bar{z}) = V(\hat{\gamma}_{x^*})$

$+\ \bar{z}^2 V(\hat{\gamma}_{x^*z^*}) - 2\bar{z}C(\hat{\gamma}_{x^*}, \hat{\gamma}_{x^*z^*})$. Hence, while the estimated *coefficients* and variance-covariance matrices of *coefficients* will differ numerically $\left(\text{i.e., } \hat{\beta}_x \neq \hat{\gamma}_{x^*} \text{ and } \widehat{V(\hat{\beta}_x)} \neq \widehat{V(\hat{\gamma}_{x^*})}\right)$, the estimated *effects* and the precision of the estimated *effects* of the variables will be identical, regardless of whether the data are centered or uncentered. Again, we warn the reader against confusing *coefficients* with *effects*.

If all estimates of the substantive effects and all estimates of the certainty of those substantive effects are identical whether the data are mean-deviated or left uncentered, how, one might wonder, can some key coefficient estimates, their standard errors, and the corresponding t-statistics differ? The answer is simply that the coefficients and associated standard errors and t-statistics do not refer to the effects at the same substantive values of the regressors across centered and uncentered models. For example, in our standard model, $y = \beta_0 + \beta_x x + \beta_z z + \beta_{xz} xz + \varepsilon$, the coefficient β_x gives the effect of a unit increase in x when z equals zero; its standard error and the resulting t-ratio refer to the certainty of that x effect at that particular z value. In $y = \gamma_{0^*} + \gamma_{x^*} x^* + \gamma_{z^*} z^* + \gamma_{x^*z^*} x^* z^* + \varepsilon^*$, the coefficient γ_{x^*} gives the effect of a unit increase in x^* (or x, since a unit increase in x or x^* is the same thing) when z^* equals zero, which is not at all the same value as when $z = 0$ (assuming, of course, that $\bar{z} \neq 0$). Since $z^* = z - \bar{z}$, the centered z^* equals zero when the uncentered z equals its mean, not when $z = 0$ (except in the specific case where $\bar{z} = 0$). The standard error of this coefficient estimate, $\hat{\gamma}_{x^*}$, and the resulting t-ratio also refer to the certainty of the effect of a one-unit change in x at this *different* $z = \bar{z}$ value. Coefficients, standard errors, and t-statistics differ in the centered and the noncentered models because they refer to different substantive quantities, not because either model produces different, much less any better, estimates of effects than does the other.

Centering can, in this manner, actually be useful for substantive interpretation in some contexts. If interpreted carefully and understood fully, centering sometimes can facilitate a more substantively grounded discussion of the empirical analysis. If z cannot logically equal zero, then substantive interpretation of β_x is vacuous, but examining the effect of x when z is equal to its sample mean might be substantively revealing. If so, researchers might advantageously center z around its mean to aid substantive interpretation and discussion of β_x. That is, centering z around its mean allows one to interpret the coefficient on x as the effect of x when z equals its mean rather than when z equals zero. Further, it allows the researcher to interpret the t-statistic on $\hat{\gamma}_x$ as the statistical significance of x when z happens to equal its mean, which may likewise simplify discussion in some contexts.

Accordingly, our concern is that researchers too often misinterpret the results of centering—and have come to the mistaken conclusion that centering alters the estimates of effects or the estimated significance of effects. We recommend that centering transformations, if applied at all, be applied only with the aim to improve substantive presentation, not, mistakenly, to improve (apparent) statistical precision and certainly not, reprehensibly, to move the value of z to which the standard t-ratio refers so as to maximize the number of asterisks of statistical significance on reported t-tests. The substantive interpretation of the effects and the certainty of those effects are completely unaffected by this algebraic sleight-of-hand.

Including x and z when xz Appears

To estimate models containing multiplicative interaction terms, most texts advise a hierarchical testing procedure: that is, if xz enters the model, then x and z must also. If wxz appears, then all (six) of the lower order combinations (x, w, z, xw, xz, wz) must appear also, and so on analogously for higher order interactions. Allison (1979), for example, writes, "[The] common rule . . . is that testing for interaction in multiple regression should only be done hierarchically . . . If a rationale for this rule is given at all, it is usually that additive relationships somehow have priority over multiplicative relationships" (149–50). This rule is probably an advisable one, if researchers must have a rule. Certainly it is a much safer rule than an alternative proviso that one can include or not include components to interactions with little concern or consideration. However, we believe that researchers must understand the logical foundations of the models they estimate and the meaning and purpose of any proffered rule, instead of merely following such rules by rote. We argue instead for theoretically driven empirical specifications with better appreciation of the assumptions underlying alternative models. While the rule of including x and z if including xz may be a quite reasonable application of Occam's razor and is often practically advisable, it is neither logically nor statistically strictly necessary.

As proof that the rule is not logically necessary, notice that one can decompose any variable into the product of two or more others; therefore, strict adherence to this rule would actually entail infinite regress. As a substantive example, note that real GDP (per capita) equals nominal GDP times a price-index deflator (times the population inverse); conversely, nominal GDP (per capita) is real GDP times a price index (times the population inverse). Nothing statistically or logically requires researchers to include all of these components in every model containing some subset of them. Researchers should, instead, estimate the models their theories suggest.

That said, several good reasons to follow the rule exist. First, given the state of social-science theory, the models implied by theory will often be insufficiently specified as to whether to include x and/or z in an interactive model. Due scientific caution would then suggest including x and z to allow the simpler linear-additive theory a chance. (This is Occam's razor.) Failing to do so would tend to yield falsely significant estimates of coefficients on xz if, in fact, x or z or both had just linear-additive effect on y. Second, inclusion of the x and z terms in models involving xz allows a nonzero intercept to the conditional effect lines, such as those plotted in chapter 3. This is important because, even if the effect of x on y is truly zero when z is zero, if this conditional relationship is nonlinear, allowing a nonzero intercept to the linear-interactive estimates of the truly nonlinear interaction (by including x and z) will enhance the accuracy of the linear approximation. Third, and perhaps most important, even when the theory clearly excludes x and/or z from the model, that is, when it unequivocally establishes the effect of one (or both) variable(s) to be zero when the other is zero, the researcher can and should test that prediction and report the certainty with which the data support the exclusion. If that test supports exclusion, then both theory and evidence recommend exclusion of the components, and continued inclusion would be the misspecification of the model. For this sort of empirical exploration, only finding a coefficient expected to be zero in fact to be estimated as (very close to) zero and, highly preferably, with small standard error is clear evidence from the data that the assumption holds. That is, clearest support for the assumption comes from failure to reject *because the estimate is with considerable certainty near zero* rather than because the estimate has very large standard error. In sum, then, this rule, as an application of Occam's razor, is a safer adage than its opposite, but researchers should still, first, understand the basis for the rule and, second, should not shy from breaking it if their theory *and the data* strongly suggest doing so.

We now elaborate these points more fully and formally. If the theory expressly excludes z from having any effect on y when x is zero—that is, nonzero presence of x is a necessary condition for z to affect y, the correct model is

$$y = \beta_0 + \beta_x x + \beta_{xz} xz + \varepsilon \tag{35}$$

By this model, as theory demands, the effect of z on y, $\partial y/\partial z$, equals $\beta_{xz}x$, which is zero when $x = 0$. Estimating this model assumes that x

must be present for z to affect y but does not allow the data to adjudicate the question. If z does affect y even when x is zero, equation (35) would suffer omitted-variable bias, with coefficient estimates wrongly attributing the omitted variable's effects to the variable(s) that do enter the model and that correlate with the omissions. In this case, the omission will most likely imply a biased β_{xz} estimate (primarily). Regression estimation will attribute some of the true-but-omitted effect of z when $x = 0$ to z's interaction with x, and so the estimate of β_{xz} will be too large (small) when this true-but-omitted effect is positive (negative). Thus, if the omitted effect is positive, the estimated effect of z on y ($\partial\hat{y}/\partial z = \hat{\beta}_{xz}x$) will reflect a greater conditional effect than truly exists (i.e., greater slope to this effect line), with underestimation of the effect of z on y at low values of x and overestimation at high values of x. Conversely, the effect of x on y ($\partial\hat{y}/\partial x = \hat{\beta}_x + \hat{\beta}_{xz}z$) will be estimated as more conditional upon z than it truly is, implying too great a slope to this effect line and, likely, also too low an intercept ($\hat{\beta}_x$) to that effect line.

Rather than assume such necessity clauses by omitting key interaction components, we suggest that researchers test them by first estimating the model including all lower order components:

$$y = \beta_0 + \beta_x x + \beta_z z + \beta_{xz} xz + \varepsilon \qquad (36)$$

An insignificant coefficient of β_z here might then support the exclusion theory and provide some justification for proceeding with the necessity clause in place. But recall that a t-test on $\hat{\beta}_z$ only refers to the effect of z when x equals zero. The theory concludes that β_z should equal zero, and so we would hardly want to accept that hypothesis merely because we fail to reject it at some generous significance level like $p < 0.10$. Recall that failure to reject can occur with small coefficient estimates and small standard errors, small coefficient estimates and large standard errors, or large coefficient estimates and larger standard errors. Only the first of these should give the researcher great comfort that he or she may estimate the model that assumes the necessity clause by omitting (an) interaction component(s); the second gives less support for such a restriction; and the last gives very little or none at all.

In summary, estimating models like (36) that include all interaction components when true models, for instance, (35), actually exclude them will cost researchers some inefficiency if not bias. Estimating (36) when the true model is (35) involves trying to estimate more coefficients than necessary, which implies inflated standard errors. Moreover, these included-but-unnecessary coefficients, β_x or β_z, are on variables, x or z, that are likely highly correlated with the necessary ones, xz, which implies greatly inflated standard errors. Thus, the inefficiency of overcautious interaction-component inclusion could easily and often be severe enough to lead researchers to miss many interactions actually present in their subject. Especially as theory advances to grapple with the complex conditionality of the subjects that social scientists study, and as empirical models attempt to follow even though data remain stubbornly scarce, such inefficiency can very easily become unaffordable. Thus, we recommend that researchers (a) acknowledge and discuss the assumptions/arguments underlying the decision to omit *or to include* components of their interaction terms, (b) gauge statistically the certainty with which the data support those assumptions, and then (c) apply Occam's razor by following hierarchical procedures unless theory and data clearly indicate that doing so is unnecessary and overly cautious.