Journal of Management
1998, Vol. 24, No. 5, 615-622

Nonlinearity, Multicollinearity and the Probability of Type II Error in Detecting Interaction

Yoav Ganzach
Tel Aviv University

The paper analyzes the impact of the inclusion of quadratic terms on the probability of type II error in testing for interaction in the presence of multicollinearity. The analysis focuses on two situations: (a) when the true model includes only linear and interaction terms; and (b) when the true model includes linear, interaction and quadratic terms. The implications of this analysis on the estimation of interaction in multiple regression are discussed.

An interaction between two independent variables is said to occur when the impact of one independent variable on the dependent variable depends on the level of another independent variable. When there are two independent variables, X and Z, and one dependent variable, Y, interaction is usually conceptualized in terms of the effect of the product XZ on Y after the linear effects of X and Z are partialled; it is examined by estimating the model:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ + \varepsilon \quad (1)$$

and by testing whether the value of β_3 is significantly different from zero.

However, examining hypotheses about interaction by estimating model (1) may lead to increased probability of type I error—the error of accepting the hypothesis that an interaction exists (rejecting the hypothesis that an interaction does not exist) when the true model does not include an interaction. Two important conditions leading to this error are the presence of both multicollinearity between the independent variables and curvilinear (and in particular quadratic) relationships between the independent variables and the dependent variable. That is, if the "true" model is:

$$Y = b_0 + b_1 X + b_2 Z + b_3 X^2 + b_4 Z^2 + \varepsilon \quad (2)$$

Direct all correspondence to: Yoav Ganzach, Faculty of Management, Tel Aviv University, Tel Aviv, 69978, Israel; Fax: 972-3-6409560, e-mail: <YOAVGN@POST.TAU.AC.IL>.

and the researcher mistakenly estimates the model of equation 1, the estimated interaction may be significant even though there is no true interaction. The reason for this is that when the correlation between X and Z increases, so does the correlation between XZ and the quadratic terms X^2 and Z^2, which results in an overlap between the variance explained by XZ and the variance explained by the quadratic terms. Busemeyer and Jones (1983) present an extensive discussion of such "spurious" interactions.

Two recent papers have dealt with this issue. Lubinski and Humphreys (1990) analyzed data pertaining to a widely held interactive theory about the antecedents of mathematical achievement, and showed that this theory was erroneous, because researchers failed to include quadratic terms in models that examine for interaction. They suggested that nonlinear monotonic relationships between the predictors and the dependent variable often lead to type I error in testing for interaction effects.

In response to this work, Cortina (1993) argued that many of the significant interactions reported in the literature may be spurious, suggesting that, when the independent variables are correlated, quadratic terms should be examined whenever an interaction hypothesis is being tested; that is, he suggested that the model for testing interactions should be:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 X^2 + \beta_4 Z^2 + \beta_5 XZ \qquad (3)$$

The work of Lubinski and Humphreys (1990) and Cortina (1993) is certainly important in pointing out problems associated with type I error in the estimation of interaction effects. However, not enough attention was paid in these studies to problems associated with type II error, that is, the error of failing to reject the null hypothesis with respect to the interaction, when the true model does include an interaction. The purpose of this paper is to highlight some issues associated with the antecedents of type II error when interaction is estimated in the presence of multicollinearity.

The paper is organized as follows. The first section examines the argument made by Cortina (1993) that entering the quadratic terms into the model prior to testing for interaction does *not* reduce the power of detecting interaction (except for a negligible decrease due to a change in the number of degrees of freedom). The second section presents the results of a simulation that examines how the probability of type II error depends on the inclusion of quadratic terms when the true model does *not* include quadratic terms. The third section demonstrates that, when the true model does include quadratic terms, type II error can also result from *not* including quadratic terms in the regression. Finally, the last section discusses the implications of the results for the examination of interaction in multiple regression.

The Addition of Quadratic Terms Increases the Probability of Type II Error When the True Model Does Not Include Quadratic Terms

Cortina (1993) argued that when the true model does not include quadratic terms, adding these terms to the model when testing interaction hypotheses does not decrease the power of detecting interaction, except for a small decrease associated with a change in the number of degrees of freedom. Since the power of a test equals one minus the probability of type II error, this argument suggests that the probability of type II error is not affected by adding quadratic terms to the model. In this section, it is shown that when the true model does not include quadratic terms, there is loss of power—and, therefore, increase in the probability of type II error—when the curvillinear terms X^2 and Z^2 are added to the model. To see why, note that if model (1) is estimated, the effect size of the interaction is given by:

$$\lambda = \frac{R_{LI}^2 - R_L^2}{1 - R_{LI}^2} \quad (N-1) \tag{4}$$

where R_L is the variance of the dependent variable explained by the linear terms, R_{LI} is the variance explained by the linear *and* interaction terms, and N is the number of observations (Cohen, 1988: 457).

If model (3) is estimated, the effect size of the interaction is given by:

$$\lambda' = \frac{R_{LQI}^2 - R_{LQ}^2}{1 - R_{LQI}^2} \quad (N-3) \tag{5}$$

where R_{LQ} is the variance of the dependent variable explained by the linear and quadratic terms, R_{LQI} is the variance explained by the linear, quadratic and interaction terms, and N is the number of observations.

From equations 4 and 5, it can be seen that if the true model does not include quadratic terms, the effect size of the interaction is larger when model (1) is estimated than when model (3) is estimated (i.e., $\lambda' < \lambda$). This happens because $R_L < R_{LQ}$ (the quadratic terms "capture" some of the interaction variance), whereas $R_{LI} \approx R_{LQI}$ (once the interaction is included, there is no incremental variance associated with the quadratic terms). Finally, since effect size is directly related to the power of detecting the effect,[1] and since type II error decreases with increase in power, the probability of type II error in detecting interaction is higher when quadratic terms are included in the model.

The Effect of Degree of Multicollinearity, Strength of Interaction, and Sample Size on the Probability of Type II Error When the True Model Does Not Include Quadratic Terms

To assess the effect of the degree of multicollinearity and the strength of the interaction on the probability of type II error, a number of simulations were performed. The data sets for these simulations were generated by a (true) model that included only two linear terms and their interaction.

Table 1. The Percent of Significant Interaction Coefficients When the Estimated Model Included Quadratic Terms (Left Number) and When the Estimated Model Did Not Include Quadratic Terms (Right Number). The True Model Included Only the Linear Terms and Their Interaction

r_{XZ}	*N* = 200				*N* = 500			
	IS ≈ 2%	*IS* ≈ 3%	*IS* ≈ 6%	*IS* ≈ 8%	*IS* ≈ 2%	*IS* ≈ 3%	*IS* ≈ 6%	*IS* ≈ 8%
.1	30% | 31%	71% | 72%	92% | 93%	99% | 100%	76% | 76%	98% | 99%	100% | 100%	100% | 100%
.3	31% | 33%	73% | 77%	91% | 92%	98% | 99%	72% | 75%	97% | 97%	100% | 100%	100% | 100%
.5	29% | 35%	57% | 73%	79% | 90%	89% | 94%	65% | 72%	93% | 98%	100% | 100%	100% | 100%
.7	13% | 34%	32% | 66%	54% | 85%	66% | 95%	33% | 65%	65% | 98%	90% | 100%	96% | 100%

In all the data sets used in the simulations, the variance explained by the true model was kept constant at 25%. Three parameters were varied: (a) N, the number of observations (200 or 500); (b) r_{xz}, the correlation between the predictors (.1, .3, .5 and .7); and (c) IS, the strength of the interaction (about 2%, 3%, 6% and 8% of the variance was explained by the interaction). The values of the parameters were chosen to represent the ranges of values, which are usually found in management research (see Cortina, 1993 and McClelland & Judd, 1993).

Using the following procedure, 1000 data sets were generated for each of the 32 combinations of the parameters. First, values of X and Z were sampled from a bivariate standard normal distribution with the required correlation between X and Z. Second, values for Y' were generated using the equation $Y' = X + Z + bXZ$, where b was .3, .5, .7, and .9, corresponding to the four levels of interaction strength. Finally, Y' was standardized, and a random error, ε, sampled from a standard normal distribution, was added to it, to generate Y, the dependent variable in the simulation. Y was related to Y' and ε by the equation $Y = .5Y' + .866\varepsilon$, resulting in X, Z and XZ explaining 25% of the variance of Y. Note that this procedure resulted in the strength of the interaction being approximately constant within each value of b; that is, it did not depend on the multicollinearity. Therefore, the effect of r_{xz} on type II error in the results cannot be attributed to variations in IS.

For each of the 32,000 data sets, interaction was examined by estimating two models. One model included only the linear and interaction terms (the LI model) and the other included the linear, interaction, and quadratic terms (the LQI model). For each combination of parameters, the number of significant ($p < .05$) interaction coefficients in the LI model and the number of significant interaction coefficients in the LQI model were counted. The difference between these two counts indicates the degree of type II error which is attributable to the inclusion of quadratic terms.

The results of the simulations are presented in Table 1. It is clear from these results that, when the true model does not include quadratic terms, adding these terms to the model increases the probability of type II error. For example, when IS $\approx$ 3% and $n = 200$, the probability of type II error increases from 23% to 27% when $r_{xz} = .3$ and from 27% to 43% when $r_{xz} = .5$. This increase is even higher when the $r_{xz} = .7$. However, for the parameters typically encountered in management research, the increase in type II error associated with adding the quadratic terms is not large, and as the number of observations rises, it becomes quite minimal. Only when multicollinearity is very high (i.e., above .7), does the addition of quadratic terms have a substantial impact on the probability of type II error. Thus, whereas Cortina's (1993) suggestion that the addition of quadratic terms has little impact on the power of detecting interaction is appropriate for the conditions most frequently discussed in management research (the first two rows of Table 1), when multicollinearity is high, substantial loss of power may occur.

The Addition of Quadratic Terms May Decrease the Probability of Type II Error When the True Model Does Include Quadratic Terms

The previous sections showed that including quadratic terms in the regression equation increases the probability of type II error in detecting interaction.

However, *not* including these terms may also increase the probability of type II error with respect to the significance test for the interaction. This happens when the true model includes quadratic terms, and the signs of the coefficients of the quadratic terms are opposite to the sign of the product term. That is, if the true model is given by

$$Y = b_0 + b_1X + b_2Z + b_3X^2 + b_4Z^2 + b_5XZ \quad (6)$$

and the signs of b_3 and b_4 are different from the sign of b_5 (for simplicity we assume that b_1 and b_2 have the same sign; otherwise either X or Z can usually be transformed by multiplying by -1 so that the signs will be the same). In this case, not including the quadratic terms increases the probability of type II error. The reason for this is that when the multicollinearity between X and Z increases, so does the correlation between XZ and the quadratic terms. As a result, the observed interaction "captures" some of the variance associated with the quadratic terms, and the true interactive relationship is canceled out by the curvilinear relationships.

As an example, consider the following simulation in which $N = 500$; X, Z and ε (an error term) are normally distributed with the same mean and standard deviation; the correlation between X and Z is .701; and the true value of Y is given by $Y = X + Z + X^2 + Z^2 - 2XZ + 8\varepsilon$. In only 4% of 1000 such simulations did the regression $Y = \beta_0 + \beta_1X + \beta_2Z + \beta_3XZ$ yield a significant coefficient for XZ. Thus, despite the existence of a true negative interaction, the omission of the quadratic terms resulted in a non-significant interaction coefficient in the estimated model.

Lubinski and Humphreys (1990) and Cortina (1993) suggest that some of the interactions reported in the literature may be spurious, since researchers failed to include quadratic terms in their models when testing hypotheses about interactions in the presence of multicollinearity. However, as the current discussion shows, the conclusion that the results of tests for interaction in the literature are biased towards type I error is premature; they may often be biased towards type II error. Note also that if b_3 and b_4 in equation 6 are opposite in sign to b_5 and their absolute values are larger than the absolute value of b_5, estimating an interaction model without including the quadratic terms may lead to significant interaction coefficients whose values are opposite to the value of the true coefficient. Ganzach (1997) supplies real world examples for such a situation.

Finally, a more informative cost/benefit analysis of the effect of adding quadratic terms on the estimation of interaction hypotheses when the true model includes quadratic terms should involve the study of the probability of type I error, as well as the probability of type II error. However, to evaluate the probability of this type I error, information about typical quadratic effects in management research is required. Since so far researchers have seldom included quadratic terms in their models, such information is not yet available.

Discussion

The occurrence of type II error in the detection of interaction effects has received much attention in the literature (Bobko & Russell, 1994; Sackett, Harris,

& Orr, 1987), and researchers agree that the power of the methods commonly used to detect these effects is very low (e.g., Cronbach, 1987). A number of reasons for this have been offered, the most important being the low reliability of the product term (Dunlap & Kemery, 1988; MacCallum & Marr, 1995) and the small residual variance of this term when the main effects are controlled (Bobko, 1986; McClelland & Judd, 1993). Note, however, that these issues are relevant to the detection of interaction in general, and not necessarily to the detection of interaction in the presence of multicollinearity, which is the subject of the current paper. Thus, the problem of type II error, which usually occurs in testing interaction hypotheses, may be compounded in the presence of multicollinearity.

The current paper demonstrates that when multicollinearity exists and the true model does not include curvilinear trends, adding quadratic terms does increase the probability of type II error in detecting interaction. On the other hand, *not* including quadratic terms may also increase the probability of type II error (as demonstrated in the previous section), as well as the probability of type I error (Lubinski & Humphreys, 1990). It is, thus, natural to ask whether quadratic terms should be added to the regression equation when examining hypotheses about interaction.

The answer to this question depends on factors such as the cost of type I error versus the cost of type II error, or prior beliefs about the probability that the true model includes interaction and/or quadratic terms. Nevertheless, the current simulations indicate that in most of the situations which are encountered by researchers in management, adding quadratic terms does not result in a considerable increase in the probability of type II error in detecting interaction if the true model does not include quadratic terms. Thus, including quadratic terms affords protection against type I and type II error associated with the estimation of interaction when the true model includes quadratic terms—for a relatively small increase in the probability of type II error associated with the case of the true model not including quadratic terms.

On the premise that theory should serve as the prime guideline in deciding which terms are to be included in a regression model, it could be argued that when the theory predicts interaction, but not quadratic, relationships—as is often the case in applied psychology—quadratic terms should not be included in the equation (Shepperd, 1991). However, in my view, two basic considerations, which are unrelated to any specific interaction theory being tested, suggest that quadratic terms should be introduced into the model, even if the theory being tested is about interaction. First, psychological theories are usually associated with a conditionally *monotone*, rather than a conditionally linear, relationship between independent and dependent variables. Thus, a model that includes quadratic terms is usually a better representation of underlying theories than one that includes only linear terms. Second, psychological measurements are usually associated with a monotone, rather than linear, relationship between the true score of the variable and the measure of this variable (e.g., Krantz & Tversky, 1971), which may result in significant quadratic terms in a regression even if the true relationships among underlying constructs are linear (Busemeyer & Jones, 1983).

Acknowledgment: Financial support for this research was provided by the Israel Institute of Business Research.

Notes

1. The relationship between effect size and power depends only on the sample size and the significance criterion, which are equal in the two models.

References

Bobko, P. 1986. A solution to some dilemmas when testing hypotheses about ordinal interaction. *Journal of Applied Psychology*, 71: 323–326.

Bobko, P., & Russell, C. J. 1994. On theory, statistics and the search for interactions in the organizational sciences. *Journal of Management*, 20: 193–200.

Busemeyer, J. R., & Jones, L. E. 1983. Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin*, 93: 549–562.

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cortina, J. M. 1993. Interaction, nonlinearity, and multicollinearity: Implications for multiple regression. *Journal of Management*, 19: 915–922.

Cronbach, L. J. 1987. Statistical tests for moderator variables: Flaws in analyses recently proposed. *Psychological Bulletin*, 102: 414–417.

Dunlap, W. P., & Kemery, E. R. 1988. Effects of predictor intercorrelations and reliabilities on moderated multiple regression. *Organizational Behavior and Human Decision Processes*, 41: 248–258.

Ganzach, Y. 1997. Misleading interaction and curvilinear terms. *Psychological Methods*, 3: 235–247.

Krantz, D. H., & Tversky, A. 1971. Conjoint measurement analysis of composition rules in psychology. *Psychological Review*, 78: 151–169.

Lubinski, D., & Humphreys, L. G. 1990. Assessing spurious "moderator effects": Illustrated substantively with the hypothesized ("synergistic") relation between spatial and mathematical ability. *Psychological Bulletin*, 107: 385–393.

MacCallum, R. C., & Marr, C. M. 1995). Distinguishing between moderator and quadratic effects in multiple regression. *Psychological Bulletin*, 118: 405–421.

McClelland, G. H., & Judd, C. M. 1993. Statistical difficulties in detecting interactions and moderator effects. *Psychological Bulletin*, 114: 376–390.

Sackett, P. R., Harris, M. M., & Orr, J. M. 1987. On seeking moderator variables in meta-analysis of correlational data: A Monte Carlo investigation of statistical power. *Journal of Applied Psychology*, 71: 302–310.

Shepperd, J. A. 1991. Cautions in assessing spurious "moderator effects." *Psychological Bulletin*, 110: 315–317.