Controlling for Spurious Moderation in Marketing: A Review of Statistical Techniques

Ahmad Daryanto*

Bryan Lukas†

Abstract

Studies in marketing rarely control for spurious moderation, which occurs when unmeasured nonlinear terms affect researchers' moderation analyses. Spurious moderation can lead to false-positive research conclusions. With a focus on testing moderation hypotheses with multiple regression analysis and covariance-based structural equation modelling, we discuss techniques and, where relevant, develop programming code for modifying these methods so that spurious moderation can be controlled for. After demonstrating the control techniques' effectiveness with simulated data, we set out guidelines for their use, with particular attention paid to confirmatory research traditions in marketing research.

Keywords: spurious moderation; interaction effects; unmeasured nonlinear effect; multiple regression analysis; covariance-based structural equation modelling

*Lancaster University, a.daryanto@lancaster.ac.uk

†University of Manchester's Alliance Manchester Business School, bryan.lukas@manchester.ac.uk

Introduction

Testing moderation hypotheses underwrites conceptual advancements in the marketing discipline. At the core of such tests is the question of how unique a relationship between an independent and dependent variable is. For a moderation analysis to contribute to theory development and business practice, the test results need to be robust so that they can be relied on with a degree of confidence. An issue that can distort researchers' conclusions about statistically significant interaction effects is the potential for them to be spurious.

Spurious moderation occurs when the variance in a dependent variable is incorrectly attributed to a hypothesized interaction effect although it is actually due to the presence of an undetected nonlinear effect. The result of a spuriously moderated relationship is that the hypothesized relationship will appear to be statistically significant when, in fact, it is not.

The potential for spurious moderation is inherent if the components of the focal interaction term are correlated and the true population model is nonlinear. A discussion of this phenomenon first took place in the psychology literature, led by Busemeyer and Jones (1983) and Lubinski and Humphreys (1990), and extended over the course of nearly a decade with Shepperd (1991), MacCallum and Mar (1995), Ganzach (1997), and Kromrey and Foster-Johnson (1999). Parallel discussions also emerged in the management literature with Cortina (1993) and Ganzach (1998). These discussions recommend controlling for spurious moderation, with Cortina (1993) and Ganzach (1997) emerging as seminal papers on the topic in the business literature.

A citation analysis of the spurious moderation literature reveals that the potential for spurious moderation is rarely addressed in empirical marketing studies. Our methodology was to follow the citation trails of the seminal spurious moderation studies listed in Table 1. As can be seen in Table 1 and Table 2, there are just 21 reported attempts to control for spurious moderation in marketing studies hypothesizing interaction effects published over the past 40 years or so – there is a far higher number of such attempts in the broader business literature and allied disciplines such as psychology.

[INSERT TABLE 1 HERE]

[INSERT TABLE 2 HERE]

The importance of engaging with the topic of spurious moderation is pertinent for the marketing discipline. Relying on multiple traditional academic disciplines, marketing not only makes contributions to marketing theory and practice, but also generates insights for the disciplines relied on. To that end, marketing research has high interdisciplinary responsibility. However, Meyvis and Van Osselaer (2018) report of growing concerns among researchers about the replicability of empirical findings reported in marketing studies.

Why controlling for spurious moderation is uncommon in the marketing literature is not clear; a number of reasons may be at work, including the following four we aim to address in this study. The first reason is quite possibly a limited awareness of the potential for spurious moderation and its statistical implications. The range of available techniques might also have evaded researchers. Another possibility is that researchers are unsure of how to implement available techniques with their preferred software and how the techniques perform, including how the results are interpreted. A further reason is the lack of guidance in the literature on choosing among the suitable techniques.

In this study, we review the options that researchers have to control for spurious moderation. We concentrate on testing hypothesized interaction effects with two methods: multiple regression analysis and covariance-based structural equation modelling (SEM); the former method for testing interactions is the mainstay in marketing, and the latter is gaining currency. We first describe and empirically demonstrate spurious moderation to raise researchers' awareness of the phenomenon and highlight the problems it can pose. Second, we discuss techniques to modify multiple regression analysis and SEM in order to control for spurious moderation and, with that, provide researchers with a catalogue of control options. Third, we provide programming code for the SEM-based control techniques on the basis of recent methodological advances and freely available software to facilitate the techniques' uptake in research practice. Fourth, we demonstrate the use of the techniques with simulated data to showcase their effectiveness. Finally, guidelines with a particular concern for what is a confirmatory research tradition in marketing are provided to support researchers with their technique choices. We make the following recommendations: control for spurious moderation in a post-hoc manner; enter quadratic terms built from the components of the focal interaction term when multiple regression analysis is used (for which we endorse two existing techniques), and enter latent quadratic terms in the case of covariance-based SEM (for which we propose three techniques to modify existing interaction-testing approaches); and disclose the key results of control procedures.

Perhaps beginning with Varadarajan (2003), there is a growing literature stream in marketing focusing on steps that marketing researchers can undertake to improve the methodological rigor of their analyses. This development has accelerated more recently. Examples are discussions of power and effect sizes (Meyvis and Van Osselaer 2018), endogeneity (Rutz and Watson 2019), common method variance (Baumgartner, Weijters, and Pieters 2021), the instrumental variable (IV)-free Gaussian copula approach (Becker, Proksch, and Ringle 2021), and measurement invariance (Steenkamp and Maydeu-Olivares 2021). We place our review of techniques suitable for controlling spurious moderation in marketing research within this literature stream.

Description of spurious moderation

Moderation analysis informs researchers whether the effect of a predictor variable, X, on an outcome variable, Y, depends on the level of another variable, Z, called a moderator variable. A moderation or interaction effect is captured by an interaction term, XZ, created by multiplying X and Z. The model specification to test for an interaction takes the form displayed in Equation 1. If X and Z are manifest variables, the form is known as a moderated regression model. In an SEM framework, the symbols X and Z are replaced by ξ.

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

Cortina (1993) and Ganzach (1997) provide concise descriptions of spurious moderation. An observed moderation is spurious when the coefficient of the interaction term is statistically significant despite there being no true interaction; if unaware of the real reason for the observed moderation, researchers will likely draw false-positive conclusions from their moderation analysis (i.e., commit Type I error).

Spurious moderation arises when the statistical significance of the interaction term is not due to an actual interaction of the term's components but, instead, an overlap between the interaction term and untested nonlinear trends in the analyzed data. An overlap arises if the components of the interaction term are correlated.

When X and Z are correlated, XZ will be correlated with X^2 and Z^2 such that XZ effectively becomes a measure of nonlinear effects (Cortina 1993). For instance, if the correlation between X and Z is as high as 1, then X^2, Z^2, and XZ are identical. While this example could suggest that the potential for spurious moderation is restricted to cases of multicollinearity, we must emphasize that the problem persists as long as a given correlation between X and Z is nonzero.

In sum, the two conditions leading to spurious moderation are (1) a correlation between the variables of the interaction term and (2) nonlinear relationships between these variables and the dependent variable (Ganzach 1998). Since researchers rarely know with certainty what the true population model is, the literature on spurious moderation is unambiguous: if the interaction term components are correlated, recommendations are to rule out that the observed moderation is due to nonlinear trends (especially quadratic) in the population.

Demonstration of spurious moderation

The potential for spurious moderation should not be underestimated in marketing. The following simulation outcomes (described in more detail in Appendix A) showcase this point for multiple regression analysis, the results of which would also apply for latent variables.[1]

Monte Carlo simulations based on two published empirical studies in marketing journals were performed. We randomly selected Slotegraaf and Atuahene-Gima (2011) and Homburg and Kuehnl (2014) from a list of studies that met the following criteria: examination of nonlinear effects without hypothesizing moderation; reported estimated coefficients of the relationships between the predictor and the dependent variables; availability of relevant means and standard deviations; and amenability of the authors' model to a re-specification as a moderation model.[2] We then respecified the nonlinear relationships in the researchers' models to be linear. Next, we generated data using the correlation matrices reported in the studies, and proceeded to estimate an interaction effect as if we were unaware of the true (nonlinear) underlying relationship. The simulations and results are discussed in detail in Appendix A and summarized below.

In both illustrations, the population model is purely quadratic. If one were unaware of this fact and modelled linear relationships involving X, Z, and Y, then the chance of finding a statistically significant interaction between X and Z is larger than 90% in the case of Slotegraaf and Atuahene-Gima (2011), and larger than 27% in the case of Homburg and Kuehnl (2014). Had nonlinearity been controlled for, the percentage would have been reduced to just above 5% in the first illustration case, and to just slightly below 5% in the second case.

The true nature of a given population model is not always known. Our simulations show that it is useful to rule out the possibility that a statistically significant interaction effect is

[1] There are two previous demonstrations by Ganzach (1997) and MacCallum and Mar (1995) but these studies are purely statistical discussions without a context translation.

[2] That both selected studies are based on multiple regression analysis is a reflection of this method's prevalence in the marketing literature.

simply due to unknown trends in the population model.

Remedy for spurious moderation

The literature is clear on how to control for spurious moderation when testing moderation hypotheses: nonlinear covariates made from the components of the interaction term should be entered into the hypothesized model. In marketing research practice, this means testing XZ together with X^2 and Z^2. This principle applies to controlling for spurious moderation in multiple regression analysis and SEM.

Using only squared terms (and not exponents greater than two) as covariates when controlling for spurious moderation is a pragmatic convention advocated in the research methods literature. The main reason is that, while nonlinear terms can take a variable to the nth degree polynomial, imperfectly measured variables taken to powers greater than two may then be mostly comprised of measurement error (Busemeyer and Jones 1983). We should also point out that a quadratic function is usually a good approximation of nonlinear trends displayed by organizational and market phenomena—just as it is for most trends displayed by psychological phenomena for which the convention of using quadratic terms as covariates was first proposed (Ganzach 1997).

In the next sections, we review specific spurious moderation control techniques involving squared terms as covariates. We then provide illustrations using simulated data.

Testing for spurious moderation in multiple regression analysis

Below, we discuss techniques proposed in the literature to control for spurious moderation if multiple regression analysis is used to test a conceptual model. Our review of the literature reveals four control techniques: stepwise regression, simultaneous entry, hierarchical, and information-theoretic. The stepwise technique is exploratory as it involves interventions for avoiding spurious moderation prior to, or during, the assessment of the significance of a hypothesized interaction effect. Therefore, researchers do not know whether or not their specified interaction is statistically significant prior to testing for the potential influence of nonlinear effects. The hierarchical, simultaneous entry, and information-theoretic techniques, on the other hand, can be employed in an exploratory or confirmatory manner because they are conducted either before (i.e., exploratory) of after (i.e., confirmatory) the predicted interaction effect is known to be statistically significant. Table 2 lists marketing studies utilizing these control techniques.

Stepwise regression technique

A stepwise regression technique to control for spurious moderation in multiple regression analysis was suggested by Lubinski and Humphreys (1990) and is conducted in two steps. In Step 1, a linear, additive regression model is constructed with X and Z as predictors:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \varepsilon \tag{2}$$

In Step 2, XZ and two quadratic terms, X^2 and Z^2, are added to the Step 1 equation using a stepwise regression algorithm. This algorithm determines the sequences by which the main and quadratic terms are added to the model. After these steps have been taken, the fit of the competing models obtained from both steps is evaluated using R^2. The best model would be selected according to the one that produces the largest increment in R^2 over that provided by the linear, additive regression model obtained in Step 1. An F test is used to evaluate whether the selected model provides a statistically significant better fit than the model in Step 1.

Kromrey and Foster-Johnson (1999) conducted Monte Carlo simulation studies to examine the effectiveness of the stepwise regression technique, thereby varying the correlation between X and Z, the population model R^2, sample sizes, the reliabilities of X and Z, as well as the effect size of X^2, Z^2, and XZ. The linear, additive model built in Step 1 was treated as the true model. Competing models were generated in Step 2 according to the stepwise procedure. To measure the effectiveness of this approach, Type 1 error was calculated as the proportion of the times that the model in Step 1, as the true model, was not selected. These Monte Carlo simulations showed that the stepwise technique is problematic because of the difficulty of maintaining adequate statistical power to detect interactions and quadratic effects, with power reduction especially pronounced when sample sizes are less than 175, even when X and Z have high reliabilities. Also, the obtained results were not generalizable beyond the sample in which the approach was applied. Kromrey and Foster-Johnson (1999) concluded that these problems are statistical problems characteristic of stepwise regression analyses.

Hierarchical technique

The use of a hierarchical technique to detect spurious interaction effects was suggested by Cortina (1993) and involves three steps. In Step 1, a model with linear, additive terms is tested according to Equation (2). In Step 2, two quadratic terms, X^2 and Z^2, are added to the model, followed by XZ in Step 3. The essence of this technique is that the quadratic terms are treated as covariates before adding the interaction term to the model; entering XZ in the last step enables an examination of the variance in Y explained by XZ over and above the variance explained by X, Z, X^2, and Z^2. An F test or t-test is used to evaluate whether XZ is statistically significant or not. If the technique is undertaken in a confirmatory manner and XZ is not statistically significant after implementing the technique, the originally observed interaction is highly likely to be spurious.

A drawback of the hierarchical technique is that the inclusion of X^2 and Z^2 in one model together with X, Z, and XZ reduces the statistical power of detecting a statistically significant interaction effect. Aware of this issue, Cortina (1993) argued that the decrease in power is not substantial as it is limited to that associated with the loss of degrees of freedom due to the inclusion of X^2 and Z^2 and, hence, should normally not be a concern.

Simultaneous entry technique

The use of a simultaneous entry technique to control for spurious moderation was suggested by Ganzach (1997). The inspection of quadratic terms with this technique is undertaken concurrently with an assessment of the interaction effects. As with Cortina (1993), an F test or t-test is used to evaluate whether XZ is statistically significant or not. If the technique is undertaken in a confirmatory manner and XZ is not statistically significant after entering all predictors at once, the originally observed interaction is likely to be spurious. The model specification is:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ + \beta_4 X^2 + \beta_5 Z^2 + \varepsilon \tag{3}$$

The simultaneous entry technique suffers from the same drawback as the hierarchical technique. The inclusion of X^2 and Z^2 together with X, Z, and XZ in one model reduces the statistical power of detecting an interaction effect. Unlike Cortina (1993), Ganzach (1998) was more concerned with this power reduction, which appeared to be particularly notable in his analyses when the correlation between X and Z was greater than 0.7. Nevertheless, Ganzach (1998) concluded that a loss of power was an acceptable trade-off for minimizing the occurrence of spurious moderation effects.

Information-theoretic technique

This control technique suggested by Daryanto (2019) is based on a multimodel inference approach (Burnham and Anderson 2002). The technique compares a given moderated regression model with competing, plausible models, formulated by varying the slope coefficient of the following linear additive model: $Y = \beta_0 + \beta_1 X + \beta_2 Z + \varepsilon$.

As a result of the derivations, seven model specifications are generated. One model with two quadratic and one interaction term (see model LM6 below) that cannot be derived from the linear additive model is then added to the the seven models; the reason for which is that LM6 cannot be derived from the linear additive model. The model specifications are listed in Equation 4:

$$
\begin{aligned}
&LM1: Y = \beta_0 + \beta_1 X + \beta_2 Z + \varepsilon \\
&LM2: Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon \\
&LM3: Y = \beta_0 + \beta_1 Z + \beta_2 Z^2 + \varepsilon \\
&LM4: Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ + \varepsilon \\
&LM5: Y = \beta_0 + \beta_1 X^2 + \beta_2 Z^2 + \beta_3 X + \beta_4 Z + \varepsilon \\
&LM6: Y = \beta_0 + \beta_1 X^2 + \beta_2 Z^2 + \beta_3 X + \beta_4 Z + \beta_5 XZ + \varepsilon \\
&LM7: Y = \beta_0 + \beta_1 X^2 + \beta_2 Z^2 + \beta_3 X + \beta_4 Z + \beta_5 XZ + \beta_6 ZX^2 + \varepsilon \\
&LM8: Y = \beta_0 + \beta_1 X^2 + \beta_2 Z^2 + \beta_3 X + \beta_4 Z + \beta_5 XZ + \beta_6 XZ^2 + \varepsilon
\end{aligned} \tag{4}
$$

To assess whether the moderated regression model (i.e., lm4) is the best among the plausible models, a distance coefficient (D) is used—also known as Kullback-Leibler distance coefficient. The D coefficient is the rescaled value of the Akaike Information Criterion (AIC). The D coefficient of a model can be interpreted as the likelihood of the model given the data. A model is considered to be highly plausible, given the data, if D $\leq$ 2; if D $>$ 10, the model should not be considered (Burnham and Anderson 2002).

.

Testing for spurious moderation in covariance-based structural equation models

If SEM is used to test moderation hypotheses, specific techniques to control for spurious moderation—such as those for multiple regression analysis approaches reviewed above—have not been discussed in the literature. One reason might be that testing for spurious moderation is more complicated in SEM, as latent variables are measured with multiple indicators. Nevertheless, the principle of controlling for possible nonlinear effects by treating them as covariates is applicable to SEM just as it is to multiple regression analysis. We should emphasize here that preventing spurious moderation is particularly important when using SEM because of the violations of multivariate normality that can introduce bias in the interaction effects when nonlinear terms are not taken into account (Marsh, Wen, and Hau 2004).

In an exploratory SEM context then, latent quadratic and latent interaction terms should be tested simultaneously. In a confirmatory SEM context, these terms should be tested using a model-comparative approach, whereby a model with only a latent interaction term is estimated separately from a model with both latent quadratic and latent interaction terms—in other words, the former is nested within the latter and researchers can use model selection criteria (e.g., AIC) to select the best model (Lin, Huang, and Weng 2017).

Techniques to include latent quadratic terms will be shaped by the approach taken to include latent interactions. A number of approaches have been developed to include latent interactions, and these fall into three procedural types: constrained, unconstrained, and distribution-analytic; respectively, the most common approaches in marketing are the single indicator approach by Ping (1995), the unconstrained product-indicator approach by Marsh, Wen, and Hau (2004), and the latent moderated structural equation approach by Klein and Moosbrugger (2000). We outline these approaches below and show how to include latent quadratic terms to control for spurious moderation by way of newly developed code. Only four studies known to us have attempted to control for spurious moderation when using covariance-based SEM: Falk, Hammerschmidt, and Schepers (2010), Tudoran, Olsen, and Dopico (2012), Wang, Dou, and Zhou (2012), and Gabler, Hill, and Landers (2021) (see Table 1 and Table 2). How latent quadratic terms were included is not discussed in these studies.

Single indicator technique

Ping (1995) uses a single indicator constrained SEM approach for estimating a model with latent interactions, whereby a latent construct is measured with only one indicator assumed to be normally distributed.[3] In the first step, the loading and error variance of the single indicator are specified and estimated using the additive part of the model. This step is equivalent to running a confirmatory factor analysis (CFA). In the second step, parameters of the model with the latent interaction construct are estimated while the loading and error variance are fixed using estimates found in Step 1. Fixing the loadings and error variances is justified when the latent variables are unidimensional (Anderson and Gerbing 1988). Ping (1995) uses a maximum likelihood method as the method of estimation. The single indicator for the latent interaction variable is built by multiplying the sum of the X and Z indicators.

In general, if a latent variable X has J number of indicators, and a latent variable Z has K number of indicators, Ping (1995) shows that loadings for the single indicator can be calculated from the estimates of loadings of the latent variables X and Z (see Equation 5) as well as the estimates of variance of the error term of the single indicators (see Equation 6). These estimates are obtained from the additive part of the model (i.e., an SEM model without the nonlinear terms) provided that X and Z are unidimensional. If X and Z are not unidimensional, the estimates obtained from the additive model will be significantly different from those of the SEM model with nonlinear terms (Anderson and Gerbing 1988).

[3] The same author also proposed a two-step procedure (Ping 1996). In the following, we only discuss a modification of Ping (1995), since Ping (1996) requires the specification of non-linear constraints when both latent interaction and latent quadratic terms exist in an SEM model, which is complex and prone to human error; we are unaware of marketing papers that implemented Ping (1996).

We present the graphical representation of Ping (1995)'s model in Figure 1 where latent variables X and Z are measured with two indicators.

$$\begin{aligned}\lambda_{XZ} &= \sum_{i=1}^{J}\lambda_{x_i}\sum_{i=1}^{K}\lambda_{z_i} \\ \lambda_{XX} &= (\sum_{i=1}^{J}\lambda_{x_i})^2\end{aligned} \tag{5}$$

$$\begin{aligned}\sigma^2_{\mu_{XZ}} &= (\sum_{i=1}^{J}\lambda_{x_i})^2\sigma^2_X\sum_{i=1}^{K}\sigma^2_{u_{z_i}} + \\ &\quad (\sum_{i=1}^{K}\lambda_{z_i})^2\sigma^2_Z + \sum_{i=1}^{J}\sigma^2_{u_{x_i}} + \\ &\quad \sum_{i=1}^{K}\sigma^2_{u_{z_i}}\sum_{i=1}^{J}\sigma^2_{u_{x_i}} \\ \sigma^2_{\mu_{XX}} &= 4(\sum_{i=1}^{J}\lambda_{x_i})^2\sigma^2_X\sum_{i=1}^{J}\sigma^2_{u_{x_i}} + 2(\sum_{i=1}^{J}\sigma^2_{u_{x_i}})^2\end{aligned} \tag{6}$$

Where σ^2_X and σ^2_Z are the variance of X and Z; and $\sigma^2_{u_{x_i}}$ and $\sigma^2_{u_{z_i}}$ are the variance of the error terms that belong to the latent variables X and Z.

[INSERT FIGURE 1 HERE]

To safeguard against spurious moderation when using Ping (1995) for testing interactions, we suggest the following single indicator technique, code for which is found in Appendix B. The code contains the estimation of the error variance and loadings associated with three single indicators, each serving as a measurement item for two latent quadratic variables (XX and ZZ) and a latent interaction variable (XZ). The technique is run in two steps. In the first step, a main effect-only model is tested to get the estimates of the error variance and loadings for the single indicators. In the second step, the estimates obtained in the first step are used to fix the error variance and the loading of the single indicator into fixed values. In addition, the covariance between the latent interaction and its components as well as the covariance between the latent quadratic terms and their components are set to zero. Subsequently, the coefficient of (XZ) is inspected to assess the significance of the observed interaction, as potential spuriousness has now been controlled for.

Unconstrained product-indicator technique

To test hypothesized interaction effects, Marsh, Wen, and Hau (2004) allow a latent interaction to covary with its constituents; thus, the interaction is estimated freely. This unconstrained product-indicator approach assumes that the latent variables X and Z are bivariate normal, unlike the nonlinear terms XZ, X^2, and Z^2, which are not normally distributed. To create product indicators as measurement items for a latent interaction, the procedure requires that the measurement indicators (say, items for X) are first mean-centered and then multiplied with the mean-centered indicators of another latent construct (say, Z). Multiplication implies pairing the indicators of X with those of Z. The indicators for the latent interaction can be built using a matched-paired cross-product indicator strategy, in which the indicators do not appear twice in the constructions of the pair. As there are multiple ways of creating unique pairs, Marsh, Wen, and Hau (2004) recommend matching the indicators based on their reliabilities (e.g., the item with the highest reliability in X is matched with that with the highest reliability in Z). If X and Z do not have the same number of indicators, the items for latent interaction can be created using parcels of indicators (e.g., using the average scores). We present the path diagram of the unconstrained model in Figure 2.

[INSERT FIGURE 2 HERE]

To control for spurious moderation, we propose the following unconstrained product-indicator technique to be used with Marsh, Wen, and Hau (2004)'s interaction-testing approach, the code for which is in Appendix C. In the code, two latent quadratic terms (XX and ZZ) and one latent interaction (XZ) are included in the SEM model as covariates. The code utilizes the *indProd* function as a part of the *semTools* package as proposed by Jorgensen et al. (2016) that can be used to mean center or double-mean center product indicators, or create residual centering. We suggest a matched-paired cross-product indicator strategy, in which the indicators do not appear twice in the construction of indicators for the latent quadratic and latent interaction variables. A matched-pair strategy is practical if the latent variables X and Z have the same numbers of indicators. Before creating the cross-product indicators, the indicators for each latent variable are mean centered by setting meanC = True using the *indProd* function. In the estimation results, the coefficient of XZ informs whether the hypothesized interaction effect is statistically significant and not subject to potential spuriousness.

Latent moderated structural equation (LMS) technique

Klein and Moosbrugger (2000) proposed a LMS approach for including interaction and latent variable models. This distribution-analytic procedure does not require cross-product terms and is based on the analysis of the distribution of indicators represented as a finite mixture of normal distributions. To estimate the parameters in an SEM model with interactions, iterative procedures with the expectation-maximization (EM) algorithm are used. The approach can be implemented in Mplus.

Code (for MPlus) provided in Appendix D illustrates our proposed LMS technique for controlling spurious moderation if Klein and Moosbrugger (2000) is used to test interaction hypotheses. Once again, two latent quadratic terms (XX and ZZ) as covariates and one latent interaction are included in the SEM model. The latent variable interaction XZ and latent quadratic terms (e.g., XX) are specified by using the | symbol of the $MODEL$ command used jointly with the $XWITH$ option of the MODEL command (e.g., XZ | X $XWITH$ Z). In the code, for the estimation procedure, the maximum likelihood is used with robust standard errors (MLR) by specifying $ALGORITHM = INTEGRATION$ in the $ANALYSIS$ command (Muthén and Muthén 2017). For clarity from the user's point of view, we also declare $ESTIMATOR = MLR$. In the $OUTPUT$ command, users can request the standardized estimates of the model parameters by using the *standardized* option. To inspect the parameter specifications and starting values, users can use the $TECH1$ option, whilst to inspect the optimization history, users can use the $TECH8$ option (Muthén and Muthén 2017). In the estimation results, the coefficient of XZ is inspected to assess its significance; and if it is significant, it would be unlikely to be spurious.

Illustration

In this section, we illustrate the capabilities of the control techniques discussed in this paper, except the stepwise regression technique.[4] For our illustration, we first generated data according to a specified formula that contained two main effect variables (X and Z) and one quadratic term (X^2). Next, we used this data to estimate a multiple regression model (MR_1) that only contained main effects and one interaction term, and a model with two quadratic terms (MR_2). We also used this data to illustrate the three covariance-based SEM techniques (single indicator, unconstrained product-indicator, and LMS techniques).

For our illustration, we generated data according to $Y = .3X + Z + .3X^2 + \varepsilon$, $\varepsilon \sim N(0, 1)$. Although this data-generating equation does not include an intercept term, it is known to be non-zero when the model is estimated (Jöreskog and Yang 1996). We generated X and Z according to a bivariate normal distribution with both means equal to zero and standard deviations equal to one. We set the correlation between X and Z to .2. After generating X, Z, and Y, we treated them as latent variables. In doing so, we generated three indicators for each latent variable. We set loadings for all indicators to .7 and the reliability (ρ) of each latent variable to .7.

[4]We opt out of illustrating the stepwise technique, as this is an exploratory technique, which as we note later, we do not advocate.

Next, we calculated the values for each indicator as a linear combination of the values of the latent variables and standard normal distribution scores. For example, the values of the *ith* indicator of latent variable X are calculated as: $x_i = \rho X + \sqrt{1-\rho^2}\varepsilon_i$. The additive model with the coefficient estimates of the simulated data is displayed in Figure 3.

[INSERT FIGURE 3 HERE]

For the single indicator, the unconstrained product-indicator, and the LMS techniques, we estimated an SEM model with a latent interaction and two latent quadratic terms simultaneously: $Y = \beta_0 + \beta_1 X^2 + \beta_2 Z^2 + \beta_3 X + \beta_4 Z + \beta_5 XZ + \varepsilon$.

As can be seen in Table 3, the multiple regression model without controlling for the quadratic terms X^2 and Z^2 (see $MR1$) produced a significant interaction effect ($b_{xz} = .057, p < .001$). When the quadratic terms are controlled for (see $MR2$), the interaction term is no longer significant ($b_{xz} = .032, p > .05$).

[INSERT TABLE 3 HERE]

The results for the information-theoretic technique are presented in Table 4. They show that lm4, which is a multiple regression model that contains two linear terms (X, Z) and one interaction term (XZ), has a large D value ($D > 46$). Given the large D value and, thus, a small Akaike weight (W), lm4 should not be used for a moderation effect test. The most plausible models are lm5, lm6, and lm7 because they have a very small D value (each less than 2). These models contain quadratic terms (see Equation 4). These results indicate that the interaction effect captured by lm4 is highly likely to be spurious.

[INSERT TABLE 4 HERE]

Table 3 also shows results for the SEM-based control techniques. As can be seen, results from the single indicator technique ($b_{xz} = .420, p > .05$) show that the interaction term is not statistically significant. The interaction term is also not significant for the unconstrained product-indicator ($b_{xz} = .037, p > .05$) and LMS techniques ($b_{xz} = .040, p > .05$). Together, the results in Table 3 and Table 4 demonstrate the usefulness of controlling for possible nonlinear effects by treating them as covariates.

Guidelines for using a control technique

In this section, we turn to our last objective and offer four suggestions for using a technique to safeguard against spurious moderation. These suggestions accommodate both the regression-based and latent variable approaches discussed in this paper.

Suggestion 1: Control for quadratic effects in a post-hoc manner

Our preference would be to test a theoretically justified interaction effect in a confirmatory manner. This suggestion is consistent with Aiken and West (1991)'s recommendation to not include quadratic terms without sufficient theoretical justification and a theory-driven testing approach often favoured in the marketing literature. An added benefit of a confirmatory approach is that the chance of committing Type I error is less likely because the presence of moderation is already verified. Apart from stepwise, all techniques discussed in this paper can be conducted in a confirmatory manner.

Suggestion 2: When multiple regression analysis is used, enter quadratic terms

While our second suggestion is a straightforward one, namely to enter quadratic terms when multiple regression analysis is used, a critical question is at which stage of the modelling procedure should this be done. Recall, there are three available techniques to enter quadratic terms: stepwise, hierarchical, and simultaneous entry. We do not advocate the stepwise technique because of, primarily, the inherent difficulty to maintain adequate statistical power, as noted earlier. We also do not advocate the hierarchical technique because it is inconsistent with a theory-driven testing approach. Rather, we advocate the simultaneous entry technique, but we must highlight two important variations. One is that Ganzach (1998)'s approach is to enter quadratic terms a priori to knowing whether or not an interaction effect is significant, which is effectively an exploratory-simultaneous entry technique. The other is to enter quadratic terms post-hoc to finding a statistically significant interaction effect, which is a confirmatory-simultaneous entry technique.

Our preference is the latter variation, which is to enter quadratic terms into a hypothesized model *after* that model has been tested, resulting in a full model (i.e., a model with linear, quadratic, and interaction terms). This suggestion is once again in observation of a theory-driven testing approach. A variation of this confirmatory-simultaneous technique is the information-theoretic technique as it always incorporates the functional form of the full model (see lm6 in Equation 4).

Suggestion 3: When covariance-based SEM is used, enter latent quadratic terms

Entering quadratic terms can also be used to assess nonlinearity in covariance-based SEM. Echoing our first suggestion, we advise to first assess the statistical significance of the latent interaction. After the latent interaction is found to be significant, researchers can proceed with controlling for spurious moderation by including latent quadratic terms in their SEM model. This model-comparison technique is akin to a hierarchical technique in multiple regression analysis.

In comparison to the single indicator technique, our unconstrained product-indicator technique is easier to implement, avoiding the need for specifying nonlinear constraints as required in the former. This technique can be easily implemented in lavaan, which is an R open-source package.

Alternatively, researchers can use our LMS technique without a need to create products of indicators to represent interaction terms and the quadratic terms of latent factors. The execution of LMS in Mplus is straightforward for researchers who are familiar with Mplus syntax. One limitation, however, is that Mplus does not report model fit indices commonly reported in SEM, such as RMSEA, CLI, and TLI.

Suggestion 4: Report key results of the chosen control procedure

Communicating about whether and how spurious moderation was controlled for is, of course, a necessity for establishing confidence in moderation analysis results. But, controlling for spurious moderation also provides further, non-hypothesized insights into the true nature of population data because it can offer new perspectives and theoretical insights into relationships thought to be only linear. Indeed, it is an informative procedure that reaches beyond minimizing Type I error in relation to hypothesized interaction effects. In particular, by controlling for spuriousness, researchers receive not only insights into the possibility of nonlinearity, but also information on whether nonlinearity and interactions co-occur. The latter possibility of joint effects adds the exploration dimension to the investigation. It is therefore good practice to report key results stemming from a chosen control procedure. We suggest reporting in table format rather than making reassuring references in the text of a study.

Conclusion

The marketing literature – especially more recently – has decidedly advanced the rigour of scholarly activities in marketing, effectively developing its own body of work on the topic of how to enhance research quality. These developments are having a positive impact on the marketing discipline's level of maturity. Improving methods to test how unique a relationship of research interest is would seem to be an important aspect of these developments. Minimizing the occurrence of spurious moderation is such a test-improvement.

Spurious moderation occurs when variance in a dependent variable is incorrectly attributed to an observed interaction effect although it is actually due to an undetected nonlinear effect. The likely outcome of undetected spurious moderation is Type I error and, with that, the erroneous conclusion that, for example, a particular marketing relationship is unique when it is not. To undertake checks for how robust statistically significant interactions are by ruling-out spurious moderation is a good course of action in marketing research, one even the authors of this study have fallen short of in the past.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Aiken, Leona S, and Stephen G West. 1991. *Multiple Regression: Testing and Interpreting Interactions*. sage.

Alavi, Sascha, Johannes Habel, Paolo Guenzi, and Jan Wieseke. 2018. "The Role of Leadership in Salespeople's Price Negotiation Behavior." *Journal of the Academy of Marketing Science* 46 (4): 703–24.

Anderson, James, and David Gerbing. 1988. "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach." *Psychological Bulletin* 103 (3): 411.

Baumgartner, Hans, Bert Weijters, and Rik Pieters. 2021. "The Biasing Effect of Common Method Variance: Some Clarifications." *Journal of the Academy of Marketing Science* 49 (2): 221–35.

Becker, Jan-Michael, Dorian Proksch, and Christian M Ringle. 2021. "Revisiting Gaussian Copulas to Handle Endogenous Regressors." *Journal of the Academy of Marketing Science*, 1–21.

Burnham, KP, and DR Anderson. 2002. "Model Selection and Multimodel Inference. Springer Verlag." *New York*.

Busemeyer, Jerome R, and Lawrence E Jones. 1983. "Analysis of Multiplicative Combination Rules When the Causal Variables Are Measured with Error." *Psychological Bulletin* 93 (3): 549.

Cortina, Jose M. 1993. "Interaction, Nonlinearity, and Multicollinearity: Implications for Multiple Regression." *Journal of Management* 19 (4): 915–22.

Daryanto, Ahmad. 2019. "Avoiding Spurious Moderation Effects: An Information-Theoretic Approach to Moderation Analysis." *Journal of Business Research* 103: 110–18.

DeCarlo, Thomas E, and Son K Lam. 2016. "Identifying Effective Hunters and Farmers in the Salesforce: A Dispositional–Situational Framework." *Journal of the Academy of Marketing Science* 44 (4): 415–39.

Eisend, Martin. 2010. "A Meta-Analysis of Gender Roles in Advertising." *Journal of the*

Academy of Marketing Science 38 (4): 418–40.

Falk, Tomas, Maik Hammerschmidt, and Jeroen JL Schepers. 2010. "The Service Quality-Satisfaction Link Revisited: Exploring Asymmetries and Dynamics." *Journal of the Academy of Marketing Science* 38 (3): 288–302.

Gabler, Colin B, Ronald Paul Hill, and V Myles Landers. 2021. "Saving Behavior Within and Across Developing Nations: Implications for Public Policy Makers." *Journal of Public Policy & Marketing* 40 (3): 354–71.

Ganzach, Yoav. 1997. "Misleading Interaction and Curvilinear Terms." *Psychological Methods* 2 (3): 235.

———. 1998. "Nonlinearity, Multicollinearity and the Probability of Type II Error in Detecting Interaction." *Journal of Management* 24 (5): 615–22.

Godfrey, Andrea, Kathleen Seiders, and Glenn B Voss. 2011. "Enough Is Enough! The Fine Line in Executing Multichannel Relational Communication." *Journal of Marketing* 75 (4): 94–109.

Habel, Johannes, Sascha Alavi, and Doreén Pick. 2017. "When Serving Customers Includes Correcting Them: Understanding the Ambivalent Effects of Enforcing Service Rules." *International Journal of Research in Marketing* 34 (4): 919–41.

Homburg, Christian, Martin Klarmann, and Sabine Staritz. 2012. "Customer Uncertainty Following Downsizing: The Effects of Extent of Downsizing and Open Communication." *Journal of Marketing* 76 (3): 112–29.

Homburg, Christian, and Christina Kuehnl. 2014. "Is the More Always Better? A Comparative Study of Internal and External Integration Practices in New Product and New Service Development." *Journal of Business Research* 67 (7): 1360–67.

Jayanti, Rama K. 2021. "An Institutional Theory Perspective on Retailer Loyalty." *Journal of Relationship Marketing* 20 (2): 156–76.

Jöreskog, Karl G, and Fan Yang. 1996. "Nonlinear Structural Equation Models: The Kenny-Judd Model with Interaction Effects." *Advanced Structural Equation Modeling: Issues* *and Techniques* 3: 57–88.

Jorgensen, Terrence D, Sunthud Pornprasertmanit, Alexander M Schoemann, Yves Rosseel, Patrick Miller, Corbin Quick, Mauricio Garnier-Villarreal, et al. 2016. "Package 'semTools'."

Klein, Andreas, and Helfried Moosbrugger. 2000. "Maximum Likelihood Estimation of Latent Interaction Effects with the LMS Method." *Psychometrika* 65 (4): 457–74.

Kromrey, Jeffrey D, and Lynn Foster-Johnson. 1999. "Statistically Differentiating Between Interaction and Nonlinearity in Multiple Regression Analysis: A Monte Carlo Investigation of a Recommended Strategy." *Educational and Psychological Measurement* 59 (3): 392–413.

Lin, Li-Chung, Po-Hsien Huang, and Li-Jen Weng. 2017. "Selecting Path Models in SEM: A Comparison of Model Selection Criteria." *Structural Equation Modeling: A Multidisciplinary Journal* 24 (6): 855–69.

Lubinski, David, and Lloyd G Humphreys. 1990. "Assessing Spurious" Moderator Effects": Illustrated Substantively with the Hypothesized (" Synergistic") Relation Between Spatial and Mathematical Ability." *Psychological Bulletin* 107 (3): 385.

MacCallum, Robert C, and Corinne M Mar. 1995. "Distinguishing Between Moderator and Quadratic Effects in Multiple Regression." *Psychological Bulletin* 118 (3): 405.

Madrigal, Robert. 2000. "The Influence of Social Alliances with Sports Teams on Intentions to Purchase Corporate Sponsors' Products." *Journal of Advertising* 29 (4): 13–24.

Marsh, Herbert W, Zhonglin Wen, and Kit-Tai Hau. 2004. "Structural Equation Models of Latent Interactions: Evaluation of Alternative Estimation Strategies and Indicator Construction." *Psychological Methods* 9 (3): 275.

Meyvis, Tom, and Stijn MJ Van Osselaer. 2018. "Increasing the Power of Your Study by Increasing the Effect Size." *Journal of Consumer Research* 44 (5): 1157–73.

Muthén, LK, and BO Muthén. 2017. "Mplus." *Statistical Analysis with Latent Variables. User's Guide* 8.

Nagengast, Liane, Heiner Evanschitzky, Markus Blut, and Thomas Rudolph. 2014. "New Insights in the Moderating Effect of Switching Costs on the Satisfaction–Repurchase Behavior Link." *Journal of Retailing* 90 (3): 408–27.

Ogilvie, Jessica, Adam Rapp, Raj Agnihotri, and Daniel G Bachrach. 2017. "Translating Sales Effort into Service Performance: It's an Emotional Ride." *Journal of Personal Selling & Sales Management* 37 (2): 100–112.

Ordanini, Andrea, and Ananthanarayanan Parasuraman. 2011. "Service Innovation Viewed Through a Service-Dominant Logic Lens: A Conceptual Framework and Empirical Analysis." *Journal of Service Research* 14 (1): 3–23.

Ping, Robert A. 1995. "A Parsimonious Estimating Technique for Interaction and Quadratic Latent Variables." *Journal of Marketing Research* 32 (3): 336–47.

———. 1996. "Latent Variable Interaction and Quadratic Effect Estimation: A Two-Step Technique Using Structural Equation Analysis." *Psychological Bulletin* 119 (1): 166.

Qian, Cheng, Murali Chandrashekaran, and Kangkang Yu. 2015. "Understanding the Role of Consumer Heterogeneity in the Formation of Satisfaction Uncertainty." *Psychology & Marketing* 32 (1): 78–93.

Rutz, Oliver J, and George F Watson. 2019. "Endogeneity and Marketing Strategy Research: An Overview." *Journal of the Academy of Marketing Science* 47 (3): 479–98.

Shepperd, James A. 1991. "Cautions in Assessing Spurious" Moderator Effects."." *Psychological Bulletin* 110 (2): 315.

Slotegraaf, Rebecca J, and Kwaku Atuahene-Gima. 2011. "Product Development Team Stability and New Product Advantage: The Role of Decision-Making Processes." *Journal of Marketing* 75 (1): 96–108.

Steenkamp, Jan-Benedict EM, and Alberto Maydeu-Olivares. 2021. "An Updated Paradigm for Evaluating Measurement Invariance Incorporating Common Method Variance and Its Assessment." *Journal of the Academy of Marketing Science* 49 (1): 5–29.

Taylor, Steven A. 1997. "Assessing Regression-Based Importance Weights for Quality Perceptions and Satisfaction Judgements in the Presence of Higher Order and/Orinteraction Effects." *Journal of Retailing* 73 (1): 135–59.

Ting, Ding Hooi. 2004. "Service Quality and Satisfaction Perceptions: Curvilinear and Interaction Effect." *International Journal of Bank Marketing*.

Tudoran, Ana Alina, Svein Ottar Olsen, and Domingo C Dopico. 2012. "Satisfaction Strength and Intention to Purchase a New Product." *Journal of Consumer Behaviour* 11 (5): 391–405.

Varadarajan, P Rajan. 2003. "Musings on Relevance and Rigor of Scholarly Research in Marketing." SAGE Publications.

Vlachos, Pavlos A, and Adam P Vrechopoulos. 2012. "Consumer–Retailer Love and Attachment: Antecedents and Personality Moderators." *Journal of Retailing and Consumer Services* 19 (2): 218–28.

Wang, Guangping, Wenyu Dou, and Nan Zhou. 2012. "The Interactive Effects of Sales Force Controls on Salespeople Behaviors and Customer Outcomes." *Journal of Personal Selling & Sales Management* 32 (2): 225–43.

Zablah, Alex R, Brian P Brown, and Naveen Donthu. 2010. "The Relative Importance of Brands in Modified Rebuy Purchase Situations." *International Journal of Research in Marketing* 27 (3): 248–60.

Appendix A: Monte carlo simulation

To demonstrate the likelihood of spurious moderation effect and, with that, to highlight the usefulness of checking whether a detected significant interaction effect is robust, we performed Monte Carlo simulations based on two randomly selected empirical marketing studies that contain quadratic effects. Our approach was to take the quadratic relationship from the researchers' models, generate data, and then hypothesize an interaction effect as if we were unaware of the true underlying (quadratic) relationship. This approach allows us to demonstrate the danger of estimating interaction effects without controlling for nonlinear relationships in the data.

First, we considered a study by Slotegraaf and Atuahene-Gima (2011), which examines the role of decision-making in new product development teams. We selected a subset of these authors' conceptual model, focusing only on the effect of Project Stability and Team Size on Team-level Debate.[5] Slotegraaf and Atuahene-Gima (2011) hypothesized Project Stability to have an inverted U-shaped relationship with Team-level Debate. One of the control variables used in the empirical model was Team Size. For our illustration, we re-specified the curvilinear relationship into one that is a linear, moderated relationship, with Team Size as the moderator. We reasoned that the effect of Project Stability on Team-level Debate will be enhanced if Team Size is smaller rather than larger, because in smaller teams the potential for disagreement is plausibly lower than in larger teams.

Given the correlation between Project Stability and Team Size, the mean and the standard deviation of the three variables, and the coefficient estimates of Team Stability and Team Size on Team-level Debate, we generated data with 10,000 replications. The data was generated according to the curvilinear relationship between Project Stability and Team-level Debate. We assumed the regression error to be normally distributed, with the mean

[5]Project Stability refers to the likelihood of a team member to remain in a team for the duration of the team project; Team Size is the number of members in a team; and Team-level Debate refers to the extent to which team members express disagreement, opinions, and dissent regarding a team project's goals and priorities.

equalling zero and the standard deviation equalling the reported standard deviation of Team-level Debate. For each replication, we estimated a moderated regression model (referred to as the "moderated model", see Equation 7) and a linear, quadratic moderated regression model (referred to as the "full model", see Equation 8). Both models are represented by the following equations, respectively:

$$DEBATE = \beta_0 + \beta_1 STAB + \beta_2 SIZE + \beta_3 STAB \times SIZE + \mu \tag{7}$$

$$\begin{aligned} DEBATE = \pi_0 + \pi_1 STAB + \pi_2 SIZE + \pi_3 STAB^2 + \\ \pi_4 SIZE^2 + \pi_5 STAB \times SIZE + \nu \end{aligned} \tag{8}$$

Where variables DEBATE, STAB, and SIZE refers to Team-level debate, Project Stability, and Team size, respectively. The coefficients β_0, β_1, β_2 and β_3 are the regression estimates of the moderated regression model with μ as the error term. The coefficients π_0, π_1, π_2, π_3, π_4 and π_5 are the regression estimates of the full model with ν as the error term. We calculated the proportion of p-values associated with the interaction term smaller than .05 and the means of the coefficient estimate associated with XZ.

Our results are reported in Table A1 and Figure A1. For the moderated model, the mean for the coefficient estimate associated with XZ is .11, and the percentage of the p-values smaller than .05 in the moderated regression model is 91.5%. For the full model, the mean for the coefficient estimate associated with XZ is .00, and the percentage of the p-values smaller than .05 is 5.17%. We present the distribution of the p-values associated with XZ in case of Slotegraaf and Atuahene-Gima (2011) in Figure A1. The distribution shows that when the moderated regression model is used, the distribution is skewed to the right, confirming that there is a higher chance of finding the interaction term to be significant. However, when the full regression model is used, the distribution of the p-value resembles a uniform distribution with a small peak around zero, confirming that the chance of finding the interaction to be statistically significant is low.

In Homburg and Kuehnl (2014) — our second randomly selected demonstration study – Customer Integration was hypothesized to have an inverted U-shaped relationship with Innovation Success. Interfirm Collaboration was also hypothesized to have a nonlinear effect on Innovation Success.[6] For our illustration, we re-specified the curvilinear relationship between Customer Integration and Innovation Success into one that is a linear, moderated relationship, with Interfirm Collaboration as the moderator. Our reasoning in this case was that the effect of Customer Integration on Innovation Success will be enhanced if the level of Interfirm Collaboration is high rather than low, because external collaborating parties can contribute additional market input. The moderated regression model and the full model that we considered are described in Equation 9 and Equation 10 as follows:

$$SUCCESS = \beta_0 + \beta_1 INT + \beta_2 COLLAB + \beta_3 INT \times COLLAB + \mu \tag{9}$$

$$\begin{aligned} SUCCESS = \pi_0 + \pi_1 INT + \pi_2 COLLAB + \pi_3 INT^2 + \\ \pi_4 COLLAB^2 + \pi_5 INT \times COLLAB + \nu \end{aligned} \tag{10}$$

Where variables SUCCESS, INT, and COLLAB refers to Innovation Success, Customer Integration, and Interfirm Collaboration, respectively. The coefficients β_0, β_1, β_2 and β_3 are the regression estimates of the moderated model with μ as the error term. The coefficients π_0, π_1, π_2, π_3, π_4 and π_5 are the regression estimates of the full model with ν as the error term.

Homburg and Kuehnl (2014) differs from Slotegraaf and Atuahene-Gima (2011) in that the correlation between X and Z is smaller ($r(X, Z) = .12$ vs. .39), as are the regression coefficients associated with X (b = - .03 vs. b = - .25) and X^2 (b = .14 vs. b = .59) in the population model. Despite these differences, our results resemble those of Slotegraaf and Atuahene-Gima (2011). Specifically, for the moderated model, the mean for the coefficient

[6] Customer integration refers to the extent to which a company includes customers in the product development process. Innovation success is the extent to which a firm has achieved its product

estimate associated with XZ is .03, and the percentage of the p-values smaller than .05 in the moderated model is 27.9%. For the full model, the mean for the coefficient estimate associated with XZ is .00, and the percentage of the p-values smaller than .05 is 4.8% (see Table A1). Similar graphs to those of the Slotegraaf and Atuahene-Gima (2011) data were obtained for the distributions of the p-values (see Figure A2).

As we discuss in the main text, if the quadratic terms associated with X and Z had been controlled for, there would have been a lower chance than otherwise of finding the interaction term of interest to be significant. In both illustrations, the chance of detection would have been reduced to approximately 5%.

[INSERT TABLE A1 HERE]

[INSERT FIGURE A1 HERE]

[INSERT FIGURE A2 HERE]

Appendix B: Single indicator technique

R code for the implementation of the single indicator technique to control for spurious moderation when testing an interaction effect.

```
# Ping (1995)'s two-step procedure.
# Step 1. estimate error variance and loading of single indicator
# Step 2, the estimates are used to fix the error variance and loading of  # the single indicator.

rm(list = ls())
set.seed(12346)
library(haven)
library(tidyverse)
library(lavaan)

#dat<-read_sav("/Users/ahmaddaryanto/Documents/Data/dat95.sav")
dat<-read_sav("/Users/ahmaddaryanto/Documents/Data/quadratic.sav")

#dat <- data.frame(scale(dat, scale = FALSE)) # mean-centered.

centering <- function(x) {
    xcenter = colMeans(x)
    x - rep(xcenter, rep.int(nrow(x), ncol(x)))
}

dat <- centering(dat)
# calculate single indicators

xs <- dat %>%
  select(x1, x2, x3) %>%
  rowSums()

zs <- dat %>%
  select(z1, z2, z3) %>%
  rowSums()

xx <- xs * xs
zz <- zs * zs
xz <- xs * zs

dat <- cbind(dat, xx, zz, xz)
# Step 1. Run additive model to get estimates of the error variances
# and loadings of single indicators.

ping.add.model <-'

#latent variables

X =~ x1 + x2 + x3
Z =~ z1 + z2 + z3
Y =~ y1 + y2 + y3

# regression
Y ~ X + Z
```

```
'
fit <- sem(ping.add.model,
           data = dat,
           estimator = "ML")
summary(fit, standardized = TRUE, fit.measures = TRUE)

collect <- lavInspect(fit, what = "est")
loadings <- as.data.frame(collect$lambda)
error.variances <- as.data.frame(collect$theta)
lv.variances <- as.data.frame(collect$psi) # latent variable variances

############### Step 2 #################

# Extract loadings
a <- loadings["x1","X"]
b <- loadings["x2","X"]
c <- loadings["x3","X"]

e <- loadings["z1","Z"]
f <- loadings["z2","Z"]
g <- loadings["z3","Z"]

# Estimate loadings for single indicators
lambda.xx = (a + b + c) ** 2
lambda.zz = (e + f + g) ** 2
lambda.xz = (a + b + c) * (e + f + g)

print(lambda.xx)
print(lambda.zz)
print(lambda.xz)

# Estimate variances of latent constructs
V_X <- lv.variances["X","X"]
V_Z <- lv.variances["Z","Z"]

# Estimate variances of error terms
v_ux1 <- error.variances["x1","x1"]
v_ux2 <- error.variances["x2","x2"]
v_ux3 <- error.variances["x3","x3"]
v_uz1 <- error.variances["z1","z1"]
v_uz2 <- error.variances["z2","z2"]
v_uz3 <- error.variances["z3","z3"]

# Estimate of the variances of the single indicators

V_uxx <- 4*((a+b+c)**2)*V_X*(v_ux1+v_ux2+v_ux3)+2*((v_ux1+v_ux2+v_ux3)**2)
V_uzz <- 4*((e+f+g)**2)*V_Z*(v_uz1+v_uz2+v_uz3)+2*((v_uz1+v_uz2+v_uz3)**2)
V_uxz <- ((a+b+c)**2)*V_X*(v_ux1+v_ux2+v_ux3) +
         ((e+f+g)**2)*V_Z*(v_ux1+v_ux2+v_ux3) +
         (v_ux1+v_ux2+v_ux3)*(v_uz1+v_uz2+v_uz3)

print(V_uxx)
print(V_uzz)
print(V_uxz)

##################################################
```

```
# Estimate parameters of SEM model

ping.sem.model <-'

#latent variables

X =~ x1 + x2 + x3
Z =~ z1 + z2 + z3
Y =~ y1 + y2 + y3

# fix loadings
XX =~ 8.107 * xx
ZZ =~ 8.271 * zz
XZ =~ 8.189 * xz

# regression
Y ~ X + Z + XX + ZZ + XZ

# label the error variances

xx ~~ V_uxx * xx
zz ~~ V_uzz * zz
xz ~~ V_uxz * xz

# fix parameters obtained from Step 1

V_uxx == 31.76554
V_uzz == 32.50927
V_uxz == 16.23857

# fix covariances
XX~~0*X
ZZ~~0*Z
XZ~~0*X
XZ~~0*Z
XZ~~0*XX
'
fit2 <- sem(ping.sem.model,
            data = dat,
            estimator = "ML")
summary(fit2, standardized = TRUE, fit.measures = TRUE)
fitMeasures(fit,c("chisq","df","cfi","tli","rmsea","srmr"))
```

Appendix C: Unconstrained product-indicator technique

R code for the implementation of the unconstrained product-indicator technique to control for spurious moderation when testing an interaction effect.

```
#######################################
# Unconstrained product-indicator using matched-pair strategy.
# Indicators are mean-centered before product indicators are created.
# The resulting product-indicators are double-mean centered: in the indprod function, set meanC=True, and doubleMC = TRUE.
# With the matched-pair strategy, the error of product indicators need not be correlated because no components are shared.
#######################################

set.seed(1234)
library(haven)
library(tidyverse)
library(lavaan)
library(semTools)
###########

dat<-read_sav("/Users/ahmaddaryanto/Documents/Data/quadratic.sav")

xl <- dat %>%
  select(x1,x2,x3)

zl <- dat %>%
  select(z1, z2, z3)

prod<-cbind(xl, zl)

xz.ind <- indProd(prod, var1 = 1:3, var2 = 4:6, match = TRUE,
                    meanC = TRUE, residualC = FALSE, doubleMC = FALSE) %>%
                    select(x1.z1,x2.z2,x3.z3)

xx.ind<- indProd(prod, var1 = 1:3, var2 = 1:3, match = TRUE,
                    meanC = TRUE, residualC = FALSE, doubleMC = FALSE) %>%
                    select(x1.x1,x2.x2,x3.x3)

zz.ind<- indProd(prod, var1 = 4:6, var2 = 4:6, match = TRUE,
                    meanC = TRUE, residualC = FALSE, doubleMC = FALSE) %>%
                    select(z1.z1,z2.z2,z3.z3)

dat<-cbind(dat, xx.ind, zz.ind, xz.ind)

lv.model <-'

#latent variables
```

```
X =~ x1 + x2 + x3
Z =~ z1 + z2 + z3
Y =~ y1 + y2 + y3
XX =~ x1.x1 + x2.x2 + x3.x3
ZZ =~ z1.z1 + z2.z2 + z3.z3
XZ =~ x1.z1 + x2.z2 + x3.z3

# regression
Y ~ X + Z + XX + ZZ + XZ

 '
fit <- sem(lv.model,
           data = dat,
           estimator = "ML")
summary(fit, standardized = TRUE)
fitMeasures(fit,c("chisq","df","cfi","tli","rmsea","srmr"))
```

Appendix D: LMS technique

R code for the implementation of the LMS technique to control for spurious moderation when testing an interaction effect.

```
TITLE: LATENT INTERACTION AND QUADRATIC

DATA:

    FILE = quadratic.dat;

VARIABLE:
    NAMES =   x1-x3 z1-z3 y1-y3;
    USEVARIABLES = x1-x3 z1-z3 y1-y3;

ANALYSIS: ESTIMATOR = MLR;
        TYPE = RANDOM;
        ALGORITHM = INTEGRATION;

MODEL: X BY x1-x3;
       Z BY z1-z3;
       Y BY y1-y3;
       XZ | X XWITH Z;
          XSQ | X XWITH X;
           ZSQ | Z XWITH Z;

       Y ON X Z XZ XSQ ZSQ;

OUTPUT: STANDARDIZED TECH1 TECH8;
```

Table 1. Citation analysis

Key articles	N	Citation in marketing
Busemeyer and Jones (1983)	602	None
Lubinski and Humpreys (1990)	442	Taylor (1997); Madrigal (2000); Ting (2004); Vlachos, Pavlos A and Vrechopoulos (2012); Tudoran, Olsen, and Dopico (2012)
Cortina (1993)	561	Madrigal (2000); Eisend (2010); Zablah (2010); Godfrey(2011); Eisend (2015); Ogilvie, Rapp, Agnihotri, and Bachrach (2017)
McCallum and Mar (1995)	182	Taylor (1997); Ogilvie, Rapp, Agnihotri, and Bachrach (2017)
Ganzach (1997)	316	Eisend (2010); Ordanini and Parasuraman (2011); Homburg, Klarmann, and Staritz (2012); Nagengast, Evanschitzky, Blut, and Rudolph (2014); DeCarlo and Lam (2016); Ogilvie, Rapp, Agnihotri, and Bachrach (2017); Falk, Hammerschmidt, and Schepers (2010); Tudoran, Olsen, and Dopico (2012); Wang, Dou, and Zhou (2012); Qian, Chandrashekaran, and Yu (2015); Ogilvie, Rapp, Agnihotri, and Bachrach (2017); Gabler, Hill, and Landers (2021); Jayanti (2021)
Total	1787	25

Note:
Key articles = seminal spurious moderation studies, N = Number of articles citing seminal studies; Citation in marketing = Marketing studies that control for spurious moderation (25 studies without double counting; Eisend (2010), Madrigal (2000), Ogilvie et al. (2017), and Tudoran et al. (2012) cite more than one key seminal articles). Literature search conducted on October 26, 2021.

Table 2. Marketing studies that test for the presence of quadratic effects with the aim of avoiding spurious moderation using techniques discussed in this paper.

Author	Journal	Technique
Taylor (1997)	JR	Hierarchical
Madrigal (2000)	JA	Hierarchical
Ting (2004)	IJBM	Stepwise
Eisend (2010)	JAMS	Hierarchical
Falk, Hammerschmidt, and Schepers (2010)	JAMS	Product indicator (exploratory)
Zablah (2010)	IJRM	Simultaneous (confirmatory)
Godfrey(2011)	JM	Simultaneous (exploratory)
Ordanini and Parasuraman (2011)	JSR	Simultaneous (exploratory)
Homburg, Klarmann, and Staritz (2012)	JR	Simultaneous (confirmatory)
Tudoran, Olsen, and Dopico (2012)	JCB	Single indicator (exploratory)
Vlachos, Pavlos A and Vrechopoulos (2012)	JRCS	Simultaneous (exploratory)
Wang, Dou, and Zhou (2012)	JPSSM	Single indicator (exploratory)
Nagengast, Evanschitzky, Blut, and Rudolph (2014)	JR	Simultaneous (confirmatory)
Eisend (2015)	IJA	Hierarchical
Qian, Chandrashekaran, and Yu (2015)	PM	Hierarchical
DeCarlo and Lam (2016)	JAMS	Simultaneous (confirmatory)
Habel, Alavi, and Pick (2017)	IJRM	Simultaneous (exploratory)
Ogilvie, Rapp, Agnihotri, and Bachrach (2017)	JPSSM	Hierarchical
Alavi, Habel, Guenzi, and Wieseke (2018)	JAMS	Simultaneous (confirmatory)
Gabler, Hill, and Landers (2021)	JPM	LMS (exploratory)
Jayanti (2021)	JRM	Hierarchical

Note:
Google scholars as well as EBSCO and ABI/INFORM databases were used in the literature search; JAMS = Journal of the Academy of Marketing Science; JM = Journal of Marketing; JR = Journal of Retailing; JSR = Journal of Service Research; JPSSM = Journal of Personal Selling and Sales Management; JCB = Journal of Consumer Behaviour; PM = Psychology and Marketing; JPM = Journal of Psychology and Marketing; JRM = Journal of Relationship Marketing; JA = Journal of Advertising; IJRM = International Journal of Research in Marketing; IJA = International Journal of Advertising; IJBM = International Journal of Bank Marketing; JRCS = Journal of Retailing and Consumer Services; product indicator = Unconstrained product indicator; LMS = Latent moderated structural equations.

Table 3. Comparison of results under various estimation methods: population model is quadratic

Model fit	MR1	MR2	Single indicator	Unconstrained	LMS
Chisq	NA	NA	22.279	206.182	NA
df	NA	NA	24	120	NA
CFI	NA	NA	1	0.974	NA
TLI	NA	NA	1	0.967	NA
RMSEA	NA	NA	0.004	0.030	NA
SRMR	NA	NA	0.019	0.032	NA
bx	0.185***	0.184***	0.195***	0.179***	0.204***
bz	0.521***	0.521***	0.627***	0.623***	0.612***
bx^2	NA	0.202***	0.267***	0.242***	0.167***
bz^2	NA	0.008 ns	0.010 ns	0.002 ns	-0.004 ns
bxz	0.057***	0.032 ns	0.0421 ns	0.033 ns	0.040 ns

Note:
Estimates are in the standardized forms. Chisq = Chi-square, df = degree of freedom, *** p < 0.001, n.s = not significant. MR1 is a moderated regression model, MR2 is a full model. Except for LMS, predictors in other methods are mean-centered. Except for LMS, all techniques are implemented using the lavaan package in R. LMS is conducted using Mplus.

Table 4. Information-theoretic approach

Model	df	RSS	AIC	D	W
LM1	4	740.416	2216.382	40.608	< 0.000
LM2	4	986.031	2445.562	277.788	< 0.000
LM3	4	777.911	2255.902	88.128	< 0.000
LM4	5	736.779	2214.442	46.668	< 0.000
LM5	6	693.293	2167.774	0.000	0.403
LM6	7	692.207	2168.520	0.746	0.277
LM7	7	691.962	2168.237	0.463	0.400
LM8	7	736.581	2218.228	50.454	< 0.000

Note:
df = degree of freedom, RSS = residual sum of squares, AIC = Akaike Information Criterion value, D = distance, W = weight. LM1-8 = linear models whose specification refers to each equation included in Equation 4.

Table A1. Results of simulation from published studies (N=10,000)

Study	Estimated Model	b	Percent
First demonstration	Moderated	0.11	91.50
	Ful	0.00	5.17
Second demonstration	Moderated	0.03	27.90
	Ful	0.00	4.80

Note:
b = mean of the coefficient estimate associated with an interaction term; Percent = percentage of p-values associated with an interaction term smaller than .05. Full model is a model with linear, quadratic and interaction terms. Moderated is a model with linear and interaction terms.The first demonstration study generated data using correlation matrix reported in Slotegraaf and Atuahene-Gima (2011). The second demonstration study generated data using correlation matrix reported in Homburg and Kuehnl (2014).

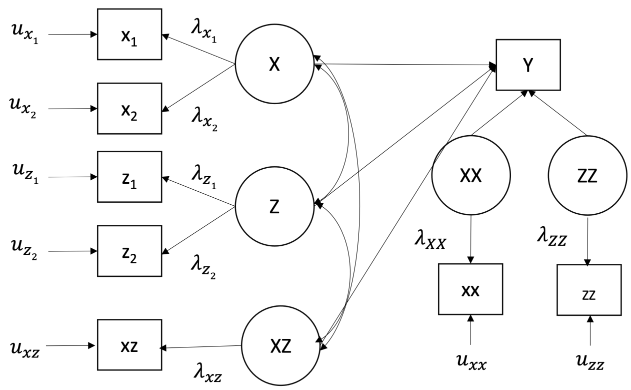

Figure 1. single product indicator

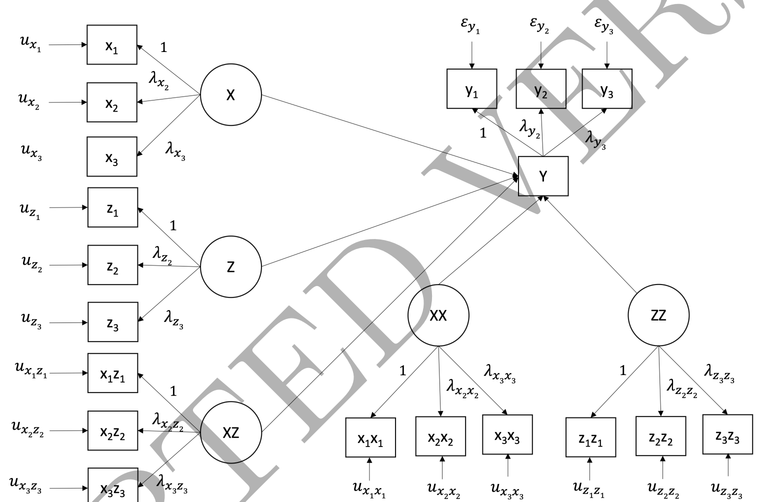

Figure 2. Unconstrained product-indicator model

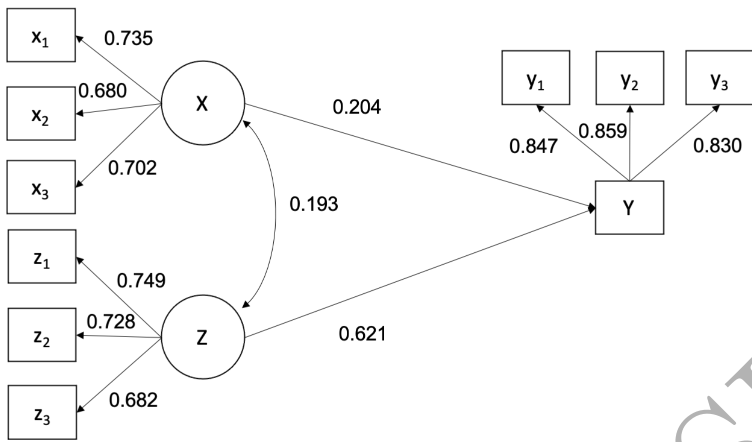

Figure 3. Unconstrained product-indicator model

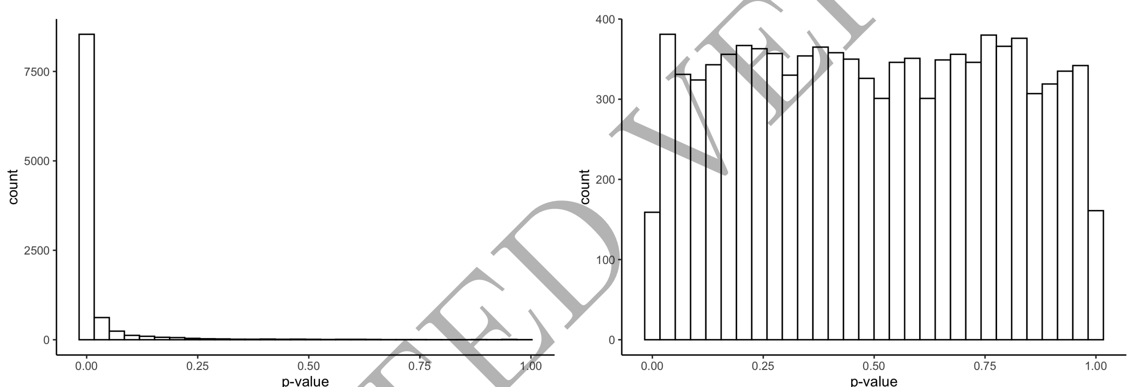

Figure A1. Distribution of p-values associated with XZ in the first simulation study, $r(X, Z)$ = 0.39, n=10000, left = moderated regression model, right = full model.

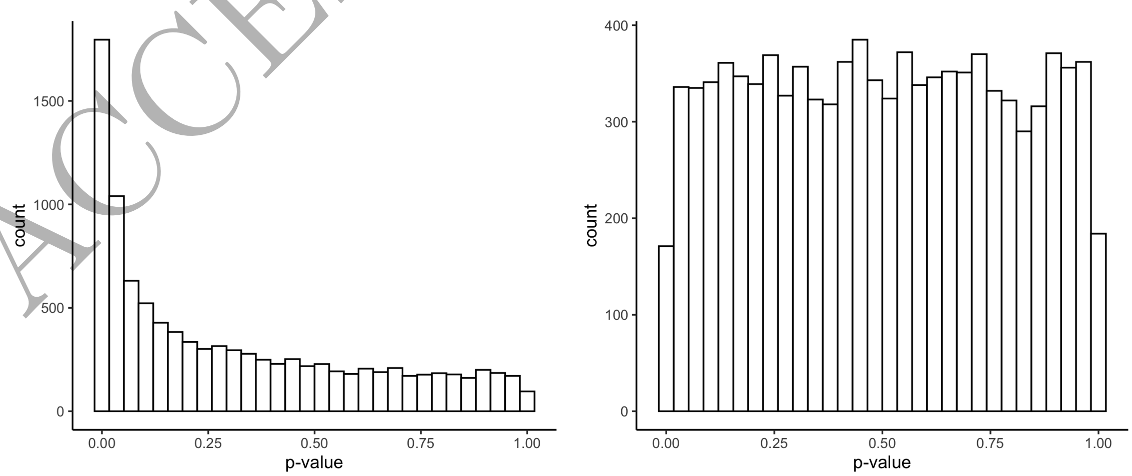

Figure A2. Distribution of p-values associated with XZ in the second simulation study, $r(X, Z) = 0.12$, n = 10000, left = moderated regression model, right = full model.