Journal of Applied Structural Equation Modeling

eISSN: 2590-4221 Journal of Applied Structural Equation Modeling: 2(2),1-21, June 2018

ESTIMATING MODERATING EFFECTS IN PLS-SEM AND PLSc-SEM: INTERACTION TERM GENERATION*DATA TREATMENT

Jan-Michael Becker[a*], Christian M. Ringle[b] and Marko Sarstedt[c]

[a]*University of Cologne, Albertus-Magnus-Platz 1, 50923 Cologne, Germany*
[b]*Hamburg University of Technology (TUHH), Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany*
[c]*Otto-von-Guericke-University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany*

[a*]*j.becker@wiso.uni-koeln.de*, [b]*c.ringle@tuhh.de*, [c]*marko.sarstedt@ovgu.de*

ABSTRACT

When estimating moderating effects in partial least squares structural equation modeling (PLS-SEM), researchers can choose from a variety of approaches to model the influence of a moderator on a relationship between two constructs by generating different interaction terms. While prior research has evaluated the efficacy of these approaches in the context of PLS-SEM, the impact of different data treatment options on their performance in the context of standard PLS-SEM and consistent PLS-SEM (PLSc-SEM) is as yet unexplored. Our simulation study addresses these limitations and explores if the choice of approach and data treatment option has a pronounced impact on the methods' parameter recovery. An empirical application substantiates these findings. Based on our results, we offer recommendations for researchers wishing to estimate moderating effects by means of PLS-SEM and PLSc-SEM.

Keywords: *Interaction Term, Moderation, Moderator Analysis, Partial Least Squares, Path Modeling, PLS-SEM, PLSc-SEM, Structural Equation Modeling, Unstandardized Data*

INTRODUCTION

Partial least squares structural equation modeling (PLS-SEM; Lohmöller, 1989; Wold, 1982) generally estimates linear relationships between the constructs of interest. However, theory may suggest that a moderator variable influences the strength, or even the direction of the relationship between constructs in the structural model. Figure 1 illustrates a case of a simple path model in which the moderator variable (or construct) M is hypothesized to influence the relationship p_1 between constructs Y_1 and Y_2 (e.g., Baron & Kenny, 1986; Hayes, 2013).

To model the moderating effect p_3, researchers generate an interaction term, which expresses the joint influence of the exogenous construct and moderator variable on the endogenous construct. In this example, the structural model regression equation would have the following form (Jaccard & Turrisi, 2003; Equation 2.3):

$$Y_2 = c + p_1 \cdot Y_1 + p_2 \cdot M + p_3 \cdot (Y_1 \cdot M) + e. \quad (1)$$

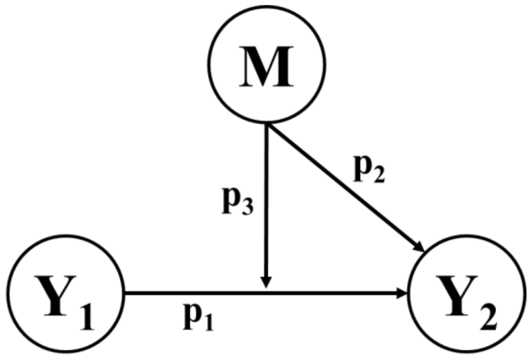

Figure 1: Simple moderator model (Hair, Hult, Ringle, & Sarstedt, 2017, Chapter 7)

Here, Y_2 represents the endogenous construct, Y_1 the exogenous construct, M the moderator variable, ($Y_1 \cdot M$) the interaction term (that represents the moderating effect p_3), and e the error term. In Equation 1, p_1, p_2, and p_3 are the structural model parameters, whereas c represents the constant. It is important to note that the direct or main effect p_1 in the PLS path model becomes the simple effect p_1 in the moderator model which includes the interaction term (Hair, Hult, Ringle, & Sarstedt, 2017, Chapter 7). The estimated equation is still linear in its parameters (i.e., p_1, p_2, and p_3), while accounting for the moderating effect utilizing the interaction term ($Y_1 \cdot M$). Rearranging the equation allows to see the (non-linear) nature of the moderating influence:

$$Y_2 = c + (\, p_1 + p_3 \cdot M)\, Y_1 + p_2 \cdot M + e. \tag{2}$$

The effect of Y_1 on Y_2 is estimated as $(p_1 + p_3 \cdot M)$. Thereby the strength of the effect of Y_1 depends on the level of the moderator M. When the moderator is zero, the effect of Y_1 on Y_2 is p_1. For each unit increase (decrease) in the moderator the simple effect increases (decreases) by p_3. If the latent variable scores are standardized (as it is usually the case in PLS-SEM), the moderator is zero at its mean and p_3 represents an increase (decrease) of the effect of Y_1 on Y_2 of one standard deviation on the moderator.

Figure 2 shows this example's structural PLS path model (on how to establish a PLS path model, see Chin, 1998; Hair, Hult, Ringle, & Sarstedt, 2017, Chapter 2)

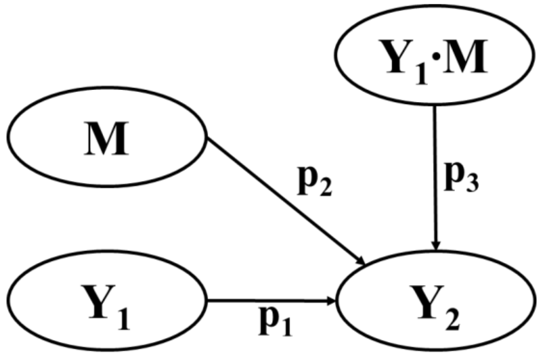

Figure 2: Interaction term in moderation (Hair, Hult, Ringle, & Sarstedt, 2017, Chapter 7)

To conduct moderator analyses in PLS-SEM, methodological research has produced several approaches for generating the interaction term expressed in Equation 1 and shown in Figure 2. Henseler and Chin (2010) give an overview of potential methods, which include the *product-indicator approach*, the *orthogonalizing approach*, the *hybrid approach*, and the *two-stage approach*. As the hybrid approach is not implemented in any PLS-SEM software and since Henseler and Chin's

(2010) simulation study results indicate that it does not perform better than any of the available alternatives in this research, we focus on the available (e.g., in software applications such as SmartPLS) and more frequently used techniques, which include the *product-indicator approach*, the *orthogonalizing approach*, and the *two-stage approach*. Besides Henseler and Chin (2010), Henseler and Fassott (2010), and Hair, Hult, Ringle, and Sarstedt (2017, Chapter 7) explain and discuss in detail these three approaches.

In their seminal paper, Chin, Marcolin, and Newsted (2003) introduced the *product-indicator approach* to PLS-SEM, demonstrating its usefulness in simulations and in an empirical application. This approach involves multiplying each indicator of the exogenous construct with each indicator of the moderator variable to generate the interaction term's product indicators. As such, the product-indicator approach requires a reflective measurement of the exogenous construct and the moderator variable (for a distinction between formative and reflective measurement models in PLS-SEM, for example, see Sarstedt, Hair, Ringle, Thiele, & Gudergan, 2016; further explications, see the Appendix A). Addressing this limitation, Henseler and Fassott (2010) introduced the *two-stage approach* to generate the interaction term when the exogenous construct and/or the moderator variable are measured formatively. Finally, the *orthogonalizing approach* extends the product-indicator approach adapting an idea of Lance's (1988) residual centering approach for moderated multiple regressions. It creates an interaction term that is uncorrelated with the predictor and moderator and thereby avoids collinearity problems from introducing the multiplicative term.

Henseler and Chin (2010) evaluated and compared the performance of these three approaches (and the hybrid approach) in an extensive simulation study with reflectively measured constructs by using common factor model data. Their results suggest that, depending on the research objective (e.g., parameter recovery, prediction, or statistical power), the two-stage approach or the orthogonalizing approach (Little, Bovaird, & Widaman, 2006)—a variant of the product-indicator approach—should be preferred. Finally, Henseler, Fassott, Dijkstra, and Wilson (2012) extended these findings to nonlinear effects between formatively measured constructs in PLS-SEM.

While these studies offer valuable insights into the efficacy of different ways to generate the interaction term, prior simulations univocally relied on standardized data. This way of treating the data seems logical as the PLS-SEM algorithm routinely standardizes the input data prior to parameter estimation (for alternative variance-based SEM algorithms such as GSCA, see Hair, Hult, Ringle, Sarstedt, & Thiele, 2017; Hwang & Takane, 2004). However, two of the approaches frequently used to generate the interaction term—the product-indicator approach and the orthogonalizing approach—require a decision on how to calculate the indicator variables' product terms. Possible options involve calculating the product terms based on (1) unstandardized indicator data, (2) mean-centered indicator data, and (3) standardized indicator data. Neither of the prior studies has paid much attention to the implications of using one of the data treatment options compared to the others. Chin et al. (2003) recommend mean-centering to avoid collinearity problems. Henseler and Chin (2010) echo this recommendation without differentiating between the options in their simulation study. Finally, Henseler and Fassott (2010) discuss the scaling of the interaction term and its indicators in terms of collinearity and interpretational confounding, but do not compare these data treatments options empirically.

Addressing this research gap, this study explores the performance of the product-indicator, two-stage, and orthogonalizing approaches on unstandardized, mean-centered, and standardized indicator data. Our simulation study examines both the standard PLS-SEM algorithm, as well as Dijkstra's (2014) and Dijkstra and Henseler's (2015a, 2015b) consistent PLS-SEM (PLSc-SEM) algorithm, when analyzing the different data treatments' efficacy in modeling the interaction term (for PLSc-SEM alternatives and extensions, see Bentler & Huang, 2014; Jung & Park, 2018). As a modified version of Wold's (1982) original PLS-SEM algorithm, the PLSc-SEM method produces model estimates that mimic a common factor model approach to measurement. Our analysis therefore extends Dijkstra and Schermelleh-Engel's (2014) study, which demonstrates PLSc-SEM's efficacy regarding estimating nonlinear effects.

Our results show that the choice of approach to generate the interaction term and data treatment has a pronounced impact on the parameter bias that PLS-SEM and PLSc-SEM produce. Specifically, we find that the two-stage approach clearly outperforms all the other approaches to operationalize the interaction term in terms of parameter recovery, regardless of whether PLS-SEM or PLSc-SEM is used. The two-stage approach performs very much like the product-indicator approach with standardized indicator data in a model that only includes reflective measurement models. However, the two-stage approach is the superior option in PLS path models that include formatively measured constructs. As such, our results suggest the routine application of the two-stage approach in PLS-SEM and PLSc-SEM. Thereby, our findings provide the guidance that researcher need in their frequent application of PLS-SEM-based moderator analyses (Ali, Rasoolimanesh, Sarstedt, Ringle, & Ryu, 2017; Nitzl, 2016; Ringle, Sarstedt, Mitchell, & Gudergan, 2018).

SIMULATION STUDY

Simulation design and preliminary analyses

Our simulation study analyzes the parameter recovery accuracy of both the PLS-SEM and PLSc-SEM algorithms when using the product-indicator, two-stage, and orthogonalizing approaches in combination with three data treatment options (i.e., unstandardized, mean-centered, and standardized data; see Mooi, Sarstedt, & Mooi-Reci, 2018). Since the technical underpinnings of the different approaches for generating the interaction term are well established, we do not elaborate on these details, but refer to the relevant literature (e.g., Hair, Hult, Ringle, & Sarstedt, 2017; Henseler & Chin, 2010; Henseler & Fassott, 2010).

Our simulation draws on a simple path model with one exogenous construct, one moderator construct, and one endogenous construct (Equation 3). In addition, we consider two path models with different measurement model operationalizations. In Model 1, three indicators with unstandardized unit loadings and standardized loadings of [.70; .80; .90] measure all three constructs reflectively (Figure C1 in the Appendix). In Model 2 (Figure C2 in the Appendix), four indictors with weights of w_1 = [.25; .40; .10; .25,] and w_2 = [.20; .35; .20; .25] measure the exogenous construct and the moderator construct formatively, whereas three indicators reflectively measure the endogenous construct as in Model 1. For Model 1 the data generation follows a common factor model approach, while it follows a mixed composite and common factor model approach for Model 2 (Rigdon, Sarstedt, & Ringle, 2017; Sarstedt et al., 2016).

In Model 1, we generated two random normal variables with means of 3 and 4 and unit standard deviations for the exogenous and moderator constructs. The endogenous construct's scores were then calculated as follows:

$$Y_2 = 5 + 6 \cdot Y_1 + 3 \cdot M + 3 \cdot (Y_1 \cdot M) + \varepsilon \quad (3)$$

When mean-centering the independent (exogenous) variables, the population values of this structural model regression equation become:[1] c=71; p_1=18; p_2=12; p_3=3. Setting the variance of the error term ε to 99 (the variance of the non-error part is 477)[2] yields the following

[1] p_1=18 (=6*3); p_2=12 (=4*3); c = 71 (=18+12+3*3*4+5);

[2] Non-error part: 477 = 18^2 (=324) + 12^2 (=144) + 3^2 (=9); Total variance (Var(Y_2)) = 477+99=576; Std(Y_2)=sqrt(576)=24

standardized population values:[3] c=0; p_1=0.75; p_2=0.50; p_3=0.125. We added indicator-specific random error terms with a mean of zero and variance specified according to the standardized indicator loading and construct scores' variance to generate the observed variable scores.

In Model 2, we generated data of eight random normal variables (i.e., X_1 to X_8) with pre-specified mean values [4; 2; 4.5; 3; 3; 4; 5; 4] and variances of $1/w_i$ (i.e., the outer model weight). Thus, the resulting exogenous constructs have a variance of $\sigma(Y_1) = \sigma(M) = 1$, and mean values of $\mu(Y_1) = 3$ and $\mu(M) = 4$. Hence,

$$Y_1 = 0.25 \cdot X_1 + 0.40 \cdot X_2 + 0.10 \cdot X_3 + 0.25 \cdot X_4 \quad (4)$$

$$M = 0.20 \cdot X_5 + 0.35 \cdot X_6 + 0.20 \cdot X_7 + 0.25 \cdot X_8 \quad (5)$$

The endogenous construct is generated in the same way as in Model 1 according to Equation 1.

We generated 1,000 datasets from this population model for both Model 1 and Model 2 with sample sizes of 100 and 500. For the data generation we use the R framework and the MASS library (R Core Team, 2014; Venables & Ripley, 2002). Appendix B documents the data generation R code for both conditions while Appendix C illustrates the path models used.

Tables 1 and 2 show the mean regression estimates across all the samples, using the simulated factor scores as input for both models. The regression estimates are very close to the population values, supporting the adequacy of our data generation procedure.

Table 1. Mean regression estimates for Model 1 (sample size: 500)

	c **Mean** (SD)	p_1 **Mean** (SD)	p_2 **Mean** (SD)	p_3 **Mean** (SD)
Unstandardized	**4.96** (5.790)	**6.03** (1.850)	**3.00** (1.394)	**3.00** (.446)
Mean-Centered	**71.00** (1.055)	**18.01** (.481)	**11.99** (.472)	**3.00** (.446)
Standardized		**.751** (.021)	**.499** (.024)	**.124** (.019)

SD = standard deviation

Table 2. Mean regression estimates for Model 2 (sample size: 500)

	c Mean (SD)	p_1 Mean (SD)	p_2 Mean (SD)	p_3 Mean (SD)
Unstandardized	**4.86** (5.834)	**6.09** (1.845)	**3.05** (1.408)	**2.97** (.446)
Mean-Centered	**71.00** (1.090)	**17.99** (.471)	**11.97** (.469)	**2.97** (.446)
Standardized		**.750** (.021)	**.499** (.024)	**.124** (.019)

SD = standard deviation

As the sample size does not influence the implications of our results, we focus our results discussion on the sample size of 500. Appendix D shows the additional results of simulations with a sample size of 100.

Next, we applied the PLS-SEM and PLSc-SEM algorithms, using the three data treatment options and the product-indicator, two-stage, and orthogonalizing approaches on the artificially generated data. In the two-stage approach the product terms are always standardized, because

[3] p_1=0.75 (=18/24); p_2=0.50 (=12/24); p_3=0.125 (=3/24)

the standardized latent variable scores from the first stage are used to generate the interaction term for the second stage. Hence, in this approach, varying the interaction term generation is not meaningful. For all approaches and in keeping with Henseler & Chin (2010), we applied the correction of the interaction term's variance, thus ensuring that the resulting interaction term was unstandardized when computing the final structural model results. We used the software SmartPLS 3, which supports the required calculations for the simulated data (Ringle, Wende, & Becker, 2015).[4]

SIMULATION RESULTS

Table 3 shows the structural model estimates for Model 1 in terms of different combinations of approaches and data treatment options regarding both PLS-SEM and PLSc-SEM. The results reveal pronounced differences, depending on the choice of approach to generate the interaction term and the data treatment.

Table 3. Simulation results of Model 1 (sample size: 500)

Population parameters: $p_1 = .750$; $p_2 = .500$; $p_3 = .125$									
	PLS-SEM		**PLSc-SEM**			**PLS-SEM**		**PLSc-SEM**	
	Mean	**SD**	**Mean**	**SD**		**Mean**	**SD**	**Mean**	**SD**
Product-indicator (Mean-Centered)					**Orthogonalizing (Mean-Centered)**				
p_1	.641	.024	.749	.030	p_1	.643	.024	.751	.031
p_2	.427	.028	.498	.035	p_2	.428	.028	.499	.035
p_3	.064	.021	.065	.023	p_3	.067	.017	.070	.017
Product-indicator (Unstandardized)					**Orthogonalizing (Unstandardized)**				
p_1	.418	.105	-.128	.067	p_1	.643	.024	.751	.031
p_2	.259	.079	-.158	.055	p_2	.428	.028	.499	.035
p_3	.045	.020	.138	.012	p_3	.067	.017	.070	.017
Product-indicator (Standardized)					**Orthogonalizing (Standardized)**				
p_1	.641	.024	.748	.031	p_1	.643	.024	.751	.031
p_2	.427	.028	.498	.035	p_2	.428	.028	.499	.035
p_3	.100	.033	.124	.042	p_3	.105	.026	.133	.033
Two-Stage (*)					**Main Effects Only**				
p_1	.643	.024	.751	.031	p_1	.643	.024	.751	.031
p_2	.428	.028	.500	.035	p_2	.428	.028	.499	.035
p_3	.098	.028	.124	.038					

SD = standard deviation

Using unstandardized data as input for the product-indicator approach produces considerable biases in both PLS-SEM and PLSc-SEM. PLSc-SEM clearly underestimates the effects of the exogenous and moderator constructs, yielding model estimates of -.128 and -.158, which are far from the pre-specified values of .75 and .50. Conversely, with an estimate of .138, the approach

[4] Technically, we used a single dataset in SmartPLS, which included an indicator for the different factor level combinations. We used this indicator as a grouping variable and, then, ran the PLS-SEM analyses for each group representing a specific factor level combination. The SmartPLS results report included the outcomes for each group-specific model estimation.

slightly overestimates the interaction term's pre-specified effect (.125). Similarly, PLS-SEM underestimates the structural model parameters, particularly the interaction effect.

On the contrary, estimating the model with PLSc-SEM by means of the product-indicator approach and standardized data, produces estimates with practically no bias and low standard deviations. These results support PLSc-SEM's adequacy to estimate models using common factor model data (Dijkstra & Henseler, 2015a, 2015b; Sarstedt et al., 2016) and is consistent with the results by Dijkstra and Schermelleh-Engel (2014). The standard PLS-SEM algorithm, however, underestimates all structural model parameters, while producing lower standard deviations, a behavior well documented in prior literature (e.g., Goodhue, Lewis, & Thompson, 2012; Henseler et al., 2014; Reinartz, Haenlein, & Henseler, 2009).

Applying PLS-SEM and PLSc-SEM, using the product-indicator approach and mean-centered data yields results like those of the standardized data option for the simple effects p_1 and p_2. However, both approaches show a pronounced tendency to underestimate the interaction effect p_3, rendering this data treatment option inadequate in terms of parameter recovery.

The results of the orthogonalizing approach correspond largely to those of the product-indicator approach across all data treatment options. The exception is, however, that the results of the unstandardized data option parallel those of the mean-centered option for both PLS-SEM and PLSc-SEM. In addition, the interaction effect p_3 is slightly overestimated for the PLSc-SEM method on standardized data. This overestimation is even more pronounced for a sample size of 100 (Appendix D). This overestimation, combined with PLS-SEM's underestimation of the relationships between common factors in the structural model, yields a parameter estimate closer to the expected population value when using PLS-SEM and the orthogonalizing approach on standardized indicator data. This could be a potential reason why Henseler and Chin (2010) conclude that the orthogonalizing approach is particularly advantageous for precise parameter recovery. In their simulation study, they only investigate the different approaches for PLS-SEM on standardized data. However, the analysis of the PLSc-SEM results reveals that the orthogonalizing approach tends to overestimate the interaction term's coefficient.

Finally, the two-stage approach's performance parallels that of the product-indicator approach with standardized data. PLSc-SEM produces almost no bias, while PLS-SEM tends to slightly underestimate the structural model parameters. The underestimation of PLS-SEM is expected as we are using purely reflective (common factor model) data. As PLS-SEM is a method of composites, it shows its known structural model parameter attenuation for common factor model data.

For Model 2, comprising reflectively and formatively measured constructs, we expect the standard PLS-SEM results to be closer to the population parameters, as attenuation should occur to a smaller degree in this model (Becker, Rai, & Rigdon, 2013). However, since part of our model's data generation follows a common factor model rather than a composite model (Hair et al., 2017; Sarstedt et al., 2016), structural model estimates between composites and common factors (i.e., the reflectively measured endogenous construct) should still be downward biased, but to a lesser degree than in a purely reflective model (i.e., Model 1).

The results in Table 4 confirm these expectations. The PLS-SEM results mirror those of Model 1 across all the data treatment and approach combinations, albeit with smaller biases in most cases. An exception is the use of the product-indicator approach with unstandardized data or mean-centered data, for which PLS-SEM produces considerably higher (downward) biases for the interaction effect p_3 than in Model 1. The same holds for PLSc-SEM, which generally performs weaker regarding recovering the interaction term's parameter across practically all combinations of approaches for generating the interaction term and data treatment options. Again, the two-stage approach excels in both PLS-SEM and PLSc-SEM. Interestingly, in this model with formative and reflective measures, the orthogonalizing approach overestimates the interaction term p_3 for both PLS-SEM and PLSc-SEM on standardized indicator data. This reconfirms the problematic overestimation tendency of the orthogonalizing approach observed for PLSc-SEM in Model 1.

Table 4. Simulation results of Model 2 (sample size: 500)

Population parameters: p_1 = .750; p_2 = .500; p_3 = .125									
	PLS-SEM		PLSc-SEM			PLS-SEM		PLSc-SEM	
	Mean	SD	Mean	SD		Mean	SD	Mean	SD
Product-indicator (Mean-Centered)					**Orthogonalizing (Mean-Centered)**				
p_1	.680	.022	.735	.024	p_1	.690	.022	.746	.024
p_2	.449	.027	.486	.030	p_2	.456	.028	.493	.030
p_3	.020	.009	.022	.010	p_3	.033	.006	.036	.007
Product-indicator (Unstandardized)					**Orthogonalizing (Unstandardized)**				
p_1	.529	.055	.571	.059	p_1	.690	.022	.746	.024
p_2	.328	.050	.354	.054	p_2	.456	.028	.493	.030
p_3	.019	.006	.021	.007	p_3	.033	.006	.036	.007
Product-indicator (Standardized)					**Orthogonalizing (Standardized)**				
p_1	.679	.022	.734	.024	p_1	.690	.022	.746	.024
p_2	.449	.027	.485	.030	p_2	.456	.028	.493	.030
p_3	.096	.040	.103	.043	p_3	.146	.022	.158	.024
Two-Stage (*)					**Main Effects Only**				
p_1	.690	.022	.746	.024	p_1	.690	.022	.746	.024
p_2	.456	.027	.493	.030	p_2	.456	.028	.493	.030
p_3	.112	.025	.121	.027					

SD = standard deviation

EMPIRICAL EXAMPLE

To illustrate the different approaches for generating the interaction term and data treatment, we draw on the simple corporate reputation model used in Hair et al. (2017; 2018). The goal of this model is to explain the effects of competence (*COMP*) and likeability (*LIKE*), representing the two dimensions of corporate reputation (Schwaiger, 2004), on customer satisfaction (*CUSA*) and ultimately customer loyalty (*CUSL*). Drawing on Hair et al. (2017), we consider customers' perceived switching costs (*SWITCH*) as a moderator variable that can be assumed to negatively influence the relationship between satisfaction and loyalty. The interaction term *CUSA* * *SWITCH* establishes the negative effect of the *SWITCH* moderator variable on the path from *CUSA* to *CUSL*. Figure 1 shows the path model, which, in a similar form, has frequently been used to illustrate the PLS-SEM methods and its extensions (Henseler, Ringle, & Sarstedt, 2016; Matthews, Sarstedt, Hair, & Ringle, 2016; Sarstedt, Ringle, & Hair, 2017).

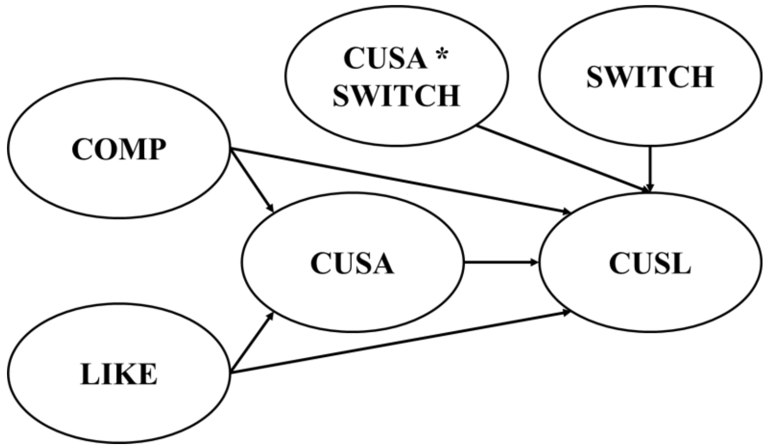

Figure 3. Corporate reputation path model

The measurement models of *COMP*, *LIKE*, and *CUSL* draw on three reflective items each, whereas *SWITCH* is measured with four reflective items. Finally, a single item (i.e., overall satisfaction) represents *CUSA*. The model estimation draws on data from two major German mobile communications network providers and two smaller competitors. A total of 344 customers rated each item on a seven-point Likert scale. We used the software SmartPLS 3 (Ringle et al., 2015) to create and estimate the model. We use bootstrapping with 5,000 samples and the no sign changes option to test for the coefficients' significance. We find that the measurement models meet all the relevant evaluation criteria.

Comparing the results of the empirical example in Table 5 with those of the simulation study (Table 3) reveals a consistent pattern. The product-indicator approach with standardized data and the two-stage approach produce highly similar results. We find that both methods have a significant ($p < .05$) negative moderating effect, which is slightly more pronounced in PLSc-SEM. Conversely, using the orthogonalizing approach with standardized data suggests a considerably stronger moderating effect, which, however, is not significant ($p > .10$). As in the simulation study, using the product-indicator approach with unstandardized data produces highly divergent results. Finally, mean-centering seems to deflate the parameter estimates, rendering the interaction effect nonsignificant ($p > .10$) when using the orthogonalizing approach.

Jointly, these results illustrate that the choice of approach to generate the interaction term and the data treatment option matter. Depending on the modus operandi, researchers should expect different results and even changes in significance. The latter is not surprising, given that moderating effects are usually small. For example, Aguinis et al.'s (2005) review of all studies published from 1969 to 1998 in the *Journal of Applied Psychology*, *Personnel Psychology*, and *Academy of Management Journal* reports a mean moderating effect size of .017 in latent variable models. In light of such a marginal mean effect size, which does not even correspond to a small effect (Cohen, 1988), the choice of approach to generate the interaction term and the data treatment option can produce divergent findings—as in our empirical example.

DISCUSSION AND CONCLUSION

When estimating moderating effects using PLS-SEM and PLSc-SEM, researchers can choose from a variety of approaches to model the moderator's influence on the relationship between two constructs. While Henseler and Chin (2010) evaluated the efficacy of these approaches in the context of PLS-SEM, prior research did not explore the impact of different data treatment options on their performance. Furthermore, prior methodological research on moderating effects focused on PLS-SEM, neglecting the PLSc-SEM algorithm proposed by Dijkstra (2014) and Dijkstra and Henseler (2015a, 2015b). Addressing these limitations, we report the results of a simulation study that not only allows us to draw conclusions on each approach's suitability regarding generating the interaction term in PLS-SEM and PLSc-SEM (i.e., product-indicator, two-stage, orthogonalizing), but also assesses how the data treatment influences these approaches' performance.

Table 5. Results of the empirical example

	PLS-SEM	PLSc-SEM		PLS-SEM	PLSc-SEM
Product-indicator (Mean-Centered)			**Orthogonalizing (Mean-Centered)**		
COMP→CUSA	.162***	.076	COMP→CUSA	.162**	.077
LIKE→CUSA	.424***	.519***	LIKE→CUSA	.424***	.519***
LIKE→CUSL	.318***	.522***	LIKE→CUSL	.305***	.495***
COMP→CUSL	-.016	-.114	COMP→CUSL	0	-.090
CUSA→CUSL	.465***	.457***	CUSA→CUSL	.497***	.496***
SWITCH→CUSL	.068	.022	SWITCH→CUSL	.070	.026
CUSA * SWITCH →CUSL (Interaction Term)	-.053**	-.058**	CUSA * SWITCH →CUSL (Interaction Term)	-.071	-.068
Product-indicator (Unstandardized)			**Orthogonalizing (Unstandardized)**		
COMP→CUSA	.162***	.076	COMP→CUSA	.162**	.076
LIKE→CUSA	.424***	.519***	LIKE→CUSA	.424***	.519***
LIKE→CUSL	.318***	.550***	LIKE→CUSL	.305***	.495***
COMP→CUSL	-.017	-.091	COMP→CUSL	0	-.090
CUSA→CUSL	.700***	.389***	CUSA→CUSL	.497***	.496***
SWITCH→CUSL	.394***	-.205*	SWITCH→CUSL	.070	.026
CUSA * SWITCH →CUSL (Interaction Term)	-.052**	.026	CUSA * SWITCH →CUSL (Interaction Term)	-.071*	-.068
Product-indicator (Standardized)			**Orthogonalizing (Standardized)**		
COMP→CUSA	.162**	.076	COMP→CUSA	.162**	.076
LIKE→CUSA	.424***	.519***	LIKE→CUSA	.424***	.519***
LIKE→CUSL	.319***	.523***	LIKE→CUSL	.305***	.494***
COMP→CUSL	-.016	-.112	COMP→CUSL	0	-.085
CUSA→CUSL	.465***	.454***	CUSA→CUSL	.497***	.496***
SWITCH→CUSL	.068	.020	SWITCH→CUSL	.069	.023
CUSA * SWITCH →CUSL (Interaction Term)	-.076**	-.090**	CUSA * SWITCH →CUSL (Interaction Term)	-.106	-.112
Two-Stage			**Main Effects Only**		
COMP→CUSA	.162**	.076	COMP→CUSA	.162**	.076
LIKE→CUSA	.424***	.519***	LIKE→CUSA	.424***	.519***
LIKE→CUSL	.319***	.523***	LIKE→CUSL	.314***	.511***
COMP→CUSL	-.017	-.115	COMP→CUSL	-.028	-.142
CUSA→CUSL	.467***	.456***	CUSA→CUSL	.498***	.497***
SWITCH→CUSL	.069		SWITCH→CUSL	.082	.057
CUSA * SWITCH →CUSL (Interaction Term)	-.071**	-.084**			

Notes: * $p<.10$, ** $p < .05$, *** $p < .01$.

The results of our simulation study show that the combination of approaches for generating the interaction term and data treatment options has a pronounced effect on the parameter estimates. Specifically, our simulation results show that the two-stage approach, which, by design, draws on standardized product terms, outperforms the other approaches in terms of parameter recovery. This performance generalizes to both PLS-SEM and PLSc-SEM, as well as models comprising reflectively and formatively measured constructs.

We also find that using unstandardized data should be avoided when generating the interaction term, as both PLS-SEM and PLSc-SEM's performance is confusing, particularly when using the product-indicator approach. Our results also advise against the use of mean-centering before creating the interaction term. This finding is striking, given that several researchers, such as Henseler and Chin (2010, p. 729), call this option to be used, concluding that "centering is advantageous for metric independent and moderator variables." While prior research has questioned mean-centering's efficacy regarding alleviating multicollinearity problems (Echambadi & Hess, 2007), our results suggest that this approach can trigger considerable biases in the estimation of the interaction effect when using PLS-SEM and PLSc-SEM.

Our findings are partly contrary to those of Henseler and Chin (2010) who found that the product-indicator approach and the orthogonalizing approach excel in terms of parameter recovery. A potential explanation for this divergence could be that, with standardized product terms of indicators, the orthogonalizing approach compensates the PLS-SEM's tendency to underestimate structural model parameters in common factor models by overestimating the interaction effect. In Model 1, the coefficient of the interaction term is closer to the true population value for the orthogonalizing approach compared to the other approaches. Yet, with respect to PLSc-SEM estimates and the second simulation model, we can see that this effect seems to be overestimated. Regarding, the product-indicator approach we particularly observe that it underperforms in models containing formatively specified constructs for the moderator and/or predictor construct (i.e., Model 2) whereas it is almost equivalent to the two-stage approach in purely reflective models (i.e., Model 1).

While our study offers important guidance regarding estimating moderating effects in PLS-SEM and PLSc-SEM, it has limitations that open opportunities for future research. First, the choice of design factors and factor levels limits the generalizability of any simulation study. We focused our simulation to a reduced set of factor combinations. Specifically, we used two different sample sizes of 100 and 500, one set of loadings, weights, and structural model coefficients. Future research could use a broader range of factor levels to assess the generalizability of our findings. However, our major concern was to show the influence of unstandardized data on the different approaches and to highlight the differences between the approaches. We do not expect the main findings to change much with the inclusion of additional factor level combinations.

Second, the advantageous two-stage approach is subject to collinearity. Even though collinearity between latent variables does not seem to be a critical issue in PLS-SEM applications, future research should attend to this potential problem and aim at developing the orthogonalizing two-stage approach. However, there are also criticism of such approaches, especially in the regression literature. For example, Echambadi, Arroniz, Reinartz, and Lee (2006) criticize the residual-centering approach by Lance (1988) in that it does not actually alleviate multicollinearity problems but distorts the interpretation of effects. Using a residual-centering or orthogonalizing approach, the p_1 and p_2 in Equation 1 cannot be interpreted as conditional (or simple) effects anymore as it is typically done in a moderation model (i.e., rearrangement to Equation 2 would not be possible). Instead they are very similar to the main effects from a model without interaction term. While this is sometimes stated as a benefit in the literature (e.g., Hair et al., 2017, Chapter 7; Henseler & Fassott, 2010), it can complicate the interpretation of effects or even render the effects invalid (Echambadi et al., 2006). Beyond these criticisms in the regression literature, the literature on covariance-based SEM has proposed further advancements of the orthogonalizing approach to product-indicators such as the double mean-centering approach (Lin, Wen, Marsh, & Lin, 2010) that future research should consider in a PLS-SEM context.

Third, we used a simulation model containing only reflective (common factor) constructs (Model 1) and one model containing both formative and reflective (composite and common factor) constructs, where the exogenous constructs are formative and the endogenous construct is reflective (Model 2). However, we did not use a model containing only formative measures or a mixed model that has formatively measured constructs in an endogenous position. The simple reason is a lack of available simulation procedures for composite model data that contains moderating effects. Future research should develop such an approach and test the performance of PLS-SEM on correctly specified composite population models that contain moderating effects.

Fourth, we did not assess the approaches' statistical power. Prior research found that the two-stage approach yields more power than the product-indicator and orthogonalizing approaches (Henseler & Chin, 2010; Henseler et al., 2012). Given that the two-stage approach produced lower standard deviations in our simulation than most of the other approaches (Tables 3 and 4), we assume that its advantage in terms of statistical power remains unchanged, making it the preferred option. However, future research should substantiate this assumption. Similarly, future research should investigate the approaches and data treatment options from an out-of-sample prediction perspective (Shmueli, Ray, Velasquez Estrada, & Chatla, 2016). In doing so, it would be particularly interesting to investigate the impact of different weighting schemes in the context of our research question. By default, the estimation of reflectively specified constructs draws on Mode A, whereas PLS-SEM uses Mode B for formatively specified constructs. However, Becker et al. (2013) show that this reflex-like use of Mode A and Mode B is not optimal in terms of out-of-sample prediction under all conditions.

Fifths, future research should investigate how unstandardized PLS-SEM estimates can correctly be derived and utilized when estimating moderating effects. The procedure to unstandardize latent variable scores in the importance-performance matrix analysis (e.g., Hair, Hult, Ringle, & Sarstedt, 2017; Ringle & Sarstedt, 2016) could be a promising way to pursue this task. Finally, analyzing the efficacy of different interaction term generation and data treatment options may represent a fruitful avenue of future research when considering moderating effects and the use of alternative variance-based SEM algorithms (e.g., GSCA; Hwang, 2009).

Sixth, the PLSc-SEM approach requires additional explorations. The method builds on the PLS-SEM outcomes and uses the composite reliability ρ_A to adjust the estimated path coefficients to mimic common factor results. Further research should address how the indicators' error term correlations[5] affect the ρ_A computation and, thus, the PLSc-SEM results in general and in the context of moderator analyses.

REFERENCES

Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect Size and Power in Assessing Moderating Effects of Categorical Variables Using Multiple Regression: A 30-year Review. *Journal of Applied Psychology, 90*(1), 94.

Ali, F., Rasoolimanesh, S. M., Sarstedt, M., Ringle, C., & Ryu, K. (2017). An Assessment of the Use of Partial Least Squares Structural Equation Modeling (PLS-SEM) in Hospitality Research. *International Journal of Contemporary Hospitality Management*, 30(1), 514-538.

Baron, R. M., & Kenny, D. A. (1986). The Moderator–mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of personality and social psychology, 51*(6), 1173.

Becker, J.-M., Rai, A., & Rigdon, E. E. (2013). *Predictive Validity and Formative Measurement in Structural Equation Modeling: Embracing Practical Relevance.* Paper presented at the 2013 Proceedings of the International Conference on Information Systems, Milan.

Bentler, P. M., & Huang, W. (2014). On Components, Latent Variables, PLS and Simple Methods: Reactions to Rigdon's Rethinking of PLS. *Long Range Planning, 47*(3), 138-145.

[5] We like to thank an anonymous reviewer for raising this concern and suggesting this cautionary note.

Chin, W. W. (1998). The Partial Least Squares Approach to Structural Equation Modeling. In G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 295-358). Mahwah: Erlbaum.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and an Electronic-Mail Emotion/Adoption Study. *Information Systems Research*, *14*(2), 189-217.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2 ed.)*: Erlbaum Associates, Hillsdale.

Dijkstra, T. K. (2014). PLS' Janus Face – Response to Professor Rigdon's 'Rethinking Partial Least Squares Modeling: In Praise of Simple Methods'. *Long Range Planning*, *47*(3), 146-153.

Dijkstra, T. K., & Henseler, J. (2015a). Consistent and Asymptotically Normal PLS Estimators for Linear Structural Equations. *Computational Statistics & Data Analysis*, *81*(1), 10-23.

Dijkstra, T. K., & Henseler, J. (2015b). Consistent Partial Least Squares Path Modeling. *MIS Quarterly*, *39*(2), 297-316.

Dijkstra, T. K., & Schermelleh-Engel, K. (2014). Consistent Partial Least Squares for Nonlinear Structural Equation Models. *Psychometrika*, *79*(4), 585-604.

Echambadi, R., Arroniz, I., Reinartz, W., & Lee, J. (2006). Empirical Generalizations from Brand Extension Research: How Sure Are We? *International Journal of Research in Marketing*, *23*(3), 253-261.

Echambadi, R., & Hess, J. D. (2007). Mean-Centering Does Not Alleviate Collinearity Problems in Moderated Multiple Regression Models. *Marketing Science*, *26*(3), 438-445.

Goodhue, D., Lewis, W., & Thompson, R. (2012). Does PLS Have Advantages for Small Sample Size or Non-normal Data? *MIS Quarterly*, *36*(3), 981-1001.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (2 ed.). Thousand Oaks, CA: Sage.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2017). Mirror, Mirror on the Wall: A Comparative Evaluation of Composite-based Structural Equation Modeling Methods. *Journal of the Academy of Marketing Science*, *45*(5), 616-632.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2018). *Advanced Issues in Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.

Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York, NY: Guilford Press.

Henseler, J., & Chin, W. W. (2010). A Comparison of Approaches for the Analysis of Interaction Effects Between Latent Variables Using Partial Least Squares Path Modeling. *Structural Equation Modeling*, *17*(1), 82-109.

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Calantone, R. J. (2014). Common Beliefs and Reality about Partial Least Squares: Comments on Rönkkö & Evermann (2013). *Organizational Research Methods*, *17*(2), 182-209.

Henseler, J., & Fassott, G. (2010). Testing Moderating Effects in PLS Path Models: An Illustration of Available Procedures. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of Partial Least Squares: Concepts, Methods and Applications (Springer Handbooks of Computational Statistics Series, vol. II)* (pp. 713-735). Heidelberg, Dordrecht, London, New York: Springer.

Henseler, J., Fassott, G., Dijkstra, T. K., & Wilson, B. (2012). Analysing Quadratic Effects of Formative Constructs by Means of Variance-based Structural Equation Modelling. *European Journal of Information Systems*, *21*(1), 99-112.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing Measurement Invariance of Composites Using Partial Least Squares. *International Marketing Review*, *33*(3), 405-431.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The Use of Partial Least Squares Path Modeling in International Marketing. In R. R. Sinkovics & P. N. Ghauri (Eds.), *Advances in International Marketing* (Vol. 20, pp. 277-320). Bingley: Emerald

Hwang, H. (2009). Regularized Generalized Structured Component Analysis. *Psychometrika, 74*(3), 517-530.

Hwang, H., & Takane, Y. (2004). Generalized Structured Component Analysis. *Psychometrika, 69*(1), 81-99.

Hwang, H., & Takane, Y. (2014). *Generalized Structured Component Analysis: A Component-Based Approach to Structural Equation Modeling*. New York: Chapman & Hall.

Jaccard, J., & Turrisi, R. (2003). *Interaction Effects in Multiple Regression*: Sage.

Jung, S., & Park, J. (2018). Consistent Partial Least Squares Path Modeling via Regularization. *Frontiers in Psychology, 9*(174).

Lance, C. E. (1988). Residual Centering, Exploratory and Confirmatory Moderator Analysis, and Decomposition of Effects in Path Models Containing Interactions. *Applied Psychological Measurement, 12*(2), 163-175.

Lin, G.-C., Wen, Z., Marsh, H. W., & Lin, H.-S. (2010). Structural Equation Models of Latent Interactions: Clarification of Orthogonalizing and Double-Mean-Centering Strategies. *Structural Equation Modeling: A Multidisciplinary Journal, 17*(3), 374-391.

Little, T. D., Bovaird, J. A., & Widaman, K. F. (2006). On the Merits of Orthogonalizing Powered and Product Terms: Implications for Modeling Interactions Among Latent Variables. *Structural Equation Modeling: A Multidisciplinary Journal, 13*(4), 497-519.

Lohmöller, J.-B. (1989). *Latent Variable Path Modeling with Partial Least Squares*. Heidelberg: Physica.

Matthews, L., Sarstedt, M., Hair, J. F., & Ringle, C. M. (2016). Identifying and Treating Unobserved Heterogeneity with FIMIX-PLS: Part II – A Case Study. *European Business Review, 28*(2), 208-224.

Mooi, E. A., Sarstedt, M., & Mooi-Reci, I. (2018). *Market Research: The Process, Data, and Methods Using Stata*. Heidelberg et al.: Springer.

Nitzl, C. (2016). The Use of Partial Least Squares Structural Equation Modelling (PLS-SEM) in Management Accounting Research: Directions for Future Theory Development. *Journal of Accounting Literature, 37*, 19-35.

R Core Team. (2014). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Reinartz, W., Haenlein, M., & Henseler, J. (2009). An Empirical Comparison of the Efficacy of Covariance-based and Variance-based SEM. *International Journal of Research in Marketing, 26*(4), 332-344.

Rigdon, E. E., Sarstedt, M., & Ringle, C. M. (2017). On Comparing Results from CB-SEM and PLS-SEM: Five Perspectives and Five Recommendations. Marketing ZFP - Journal of Research and Management, 39(3), 4-16.

Ringle, C. M., & Sarstedt, M. (2016). Gain More Insight from Your PLS-SEM Results: The Importance-Performance Map Analysis. *Industrial Management & Data Systems, 116*(9), 1865-1886.

Ringle, C. M., Sarstedt, M., Mitchell, R., & Gudergan, S. P. (2018). Partial Least Squares Structural Equation Modeling in HRM Research. *The International Journal of Human Resource Management*, forthcoming.

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). SmartPLS 3. Bönningstedt: SmartPLS. Retrieved from http://www.smartpls.com

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation Issues with PLS and CBSEM: Where the Bias Lies! *Journal of Business Research, 69*(10), 3998-4010. doi:10.1016/j.jbusres.2016.06.007

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2017). Partial Least Squares Structural Equation Modeling *Handbook of Market Research* (pp. 1-40): Springer.

Schwaiger, M. (2004). Components and Parameters of Corporate Reputation. *Schmalenbach business review, 56*(1), 46-71.

Shmueli, G., Ray, S., Velasquez Estrada, J. M., & Chatla, S. B. (2016). The Elephant in the Room: Predictive Performance of PLS Models. *Journal of Business Research, 69*(10), 4552–4564.

Venables , W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (4th ed.). New York: Springer.

Wold, H. O. A. (1982). Soft Modeling: The Basic Design and Some Extensions. In K. G. Jöreskog & H. O. A. Wold (Eds.), *Systems Under Indirect Observations: Part II* (pp. 1-54). Amsterdam: North-Holland.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *Journal of Consumer Research, 37*(3), 197-206

ACKNOWLEDGEMENT

This article uses the statistical software SmartPLS 3 (http://www.smartpls.com). Jan-Michael Becker and Christian M. Ringle acknowledges a financial interest in SmartPLS.

APPENDIX A – TERMINOLOGY

Prior literature on PLS-SEM, also with regards to interactions terms in the moderator analysis (e.g., Henseler & Fassott, 2010), refers to reflective and formatively measured constructs. This terminology has been further clarified by Sarstedt et al. (2016). Theoretically established latent variables build on operational definitions to establish their reflective or formative conceptualization. Covariance-based SEM uses common factors as proxies of reflective constructs, whereas PLS-SEM estimates composites which serve as proxies of reflectively and formatively established constructs. PLSc-SEM uses the composites obtained by PLS-SEM to mimic common factors as proxies of reflectively established constructs.

PLS-SEM traditionally allows to choose between two different ways of estimating the composite weights: correlation weights (Mode A) and regression weights (Mode B). According to Becker et al. (2013), correlation weights (Mode A) estimations in PLS-SEM circumvent collinearity issues and perform particularly well with decent sample sizes. Composites obtained by regressions weights (Mode B) perform better when sample sizes are very large and the model exhibits high R^2 levels.

In line with these explications, this research uses correlation weights (Mode A) to obtain composites as proxies of reflectively conceptualized constructs. Thereby, we bypass potential collinearity problems. These collinearity problems occur by design as reflective indicators are assumed to correlate highly. In contrast, both correlation weights (Mode A) and regression weights (Mode B) are suitable candidates to obtaining proxies of formatively conceptualized constructs. Here, we follow the traditional approach in PLS-SEM and use regression weights (Mode B) to estimate composites that become proxies of formatively established constructs (Chin, 1998; Hair, Hult, Ringle, & Sarstedt, 2017; Henseler, Ringle, & Sinkovics, 2009). This decision is justified as our simulation does not introduce collinearity between the indicators of formative measurement models. In addition, our simulations utilize relatively large sample sizes (100 and 500).

The question may arise whether the interaction term is a common factor (reflective) or a composite (formative).[6] Conceptually, it's none of both but an artificial modeling element to express the moderating effect. The focal constructs in our models are Y_1, Y_2, and M. All of these are assumed to have theoretical importance and we try to estimate effects between these theoretical entities. Yet, the theoretical test space does not contain any interaction constructs. It is simply a statistical vehicle to estimate the proposed effects. Hence, the question whether an interaction term is common factor (reflective) or a composite (formative) is not relevant. Consequently, researchers do not assess the interaction terms and its measurement model based on the conventional evaluation criteria (Hair et al. 2017, Chapter 7). Statistically, however, the interaction term for the product -indicator approach and the orthogonalizing approach is a composite as PLS-SEM is a method that uses composites to represent constructs in the path model and to estimate their (structural) model relations. Therefore, the researcher also has to specify whether to use correlation weights (Mode A) or regression weights (Mode B) to estimate the indicator weights for the product-indicators or residual-indicators. We follow prior literature (i.e., Chin et al., 2003; Henseler & Chin, 2010) and use Mode A. For the two-stage approach it might be debatable whether to name the interaction term construct scores a composite as a single indicator variable is used so that constructs scores and indicator are equal.

[6] We like to thank an anonymous reviewer for raising this question.

APPENDIX B – DATA GENERATION R CODE

R Code (Model 1)

```
library(MASS)
iterations <- 1000

correlation.matrix <- diag(12)
sample.size <- 500
mean.vector <- c(3,4,rep(0,10))
loading <- c(0.7, 0.8, 0.9)
loadings2 <- (1-loading^2)/loading^2

Lambda <- matrix(0, ncol=3, nrow=9)
Lambda[1:3,1]<-1
Lambda[4:6,2]<-1
Lambda[7:9,3]<-1

bi <- matrix(0,iterations,4)
bim <-matrix(0,iterations,4)
bis <- matrix(0,iterations,3)

for(i in 1:iterations) {
sim.data      <-      data.frame(mvrnorm(sample.size,      mean.vector,
correlation.matrix, empirical=FALSE))

A <- sim.data[,1]
B <- sim.data[,2]
E <- sim.data[,3]

F2 <- 5 + 6*A + 3*B + 3*A*B + sqrt(99)*E
# Note: mean-centered = 18*A + 12*B + 3*A*B
# Note: standardized = 0.75*A + 0.50*B +0.125*A*B

bi[i,] <- coef(lm(F2~A+B+A*B))
bim[i,]  <-  coef(lm(F2~scale(A,scale=FALSE)  +  scale(B,scale=FALSE)  +
scale(A,scale=FALSE)*scale(B,scale=FALSE)))
bis[i,]   <-   coef(lm(scale(F2)~0   +   scale(A)   +   scale(B)   +
scale(A)*scale(B)))

varmatrix2 <- diag(c(rep(var(A),3), rep(var(B),3), rep(var(F2),3)))

dat <- cbind(A,B,F2)  %*% t(Lambda) + as.matrix(sim.data[,4:12]) %*%
(sqrt(loadings2 * varmatrix2))

write.table(dat,  "Simulation_Interaction.csv",  sep=";",  append=TRUE,
row.names=FALSE, col.names=FALSE)
}

round(colMeans(bi),3) # Normal Interaction Model
round(colMeans(bim),3) # Mean-Centered Interaction Model
round(colMeans(bis),3) # Standardized Interaction Model
round(apply(bi,2,sd),3)
round(apply(bim,2,sd),3)
round(apply(bis,2,sd),3)
```

R Code (Model 2)

```
library(MASS)
iterations <- 1000

d <- c(0.25,0.4,0.1,0.25,0.2,0.35,0.2,0.25,1,1,1,1)
d <- 1/d
correlation.matrix <- diag(d)
sample.size <- 500
mean.vector <- c(4,2,4.5,3,3,4,5,4,rep(0,4))
loading <- c(0.7, 0.8, 0.9)
loadings2 <- (1-loading^2)/loading^2

bi <- matrix(0,iterations,4)
bim <-matrix(0,iterations,4)
bis <- matrix(0,iterations,3)

for(i in 1:iterations) {
sim.data     <-     data.frame(mvrnorm(sample.size,     mean.vector,
correlation.matrix, empirical=FALSE))

A  <-  0.25*sim.data[,1]  +  0.4*sim.data[,2]  +  0.1*sim.data[,3]  +
0.25*sim.data[,4]
B  <-  0.2*sim.data[,5]  +  0.35*sim.data[,6]  +  0.2*sim.data[,7]  +
0.25*sim.data[,8]
E <- sim.data[,9]

F2 <- 5 + 6*A + 3*B + 3*A*B + sqrt(100)*E
# mean-centered = 18*A + 12*B + 6*A*B
# standardized = 0.75*A + 0.50*B +0.125*A*B

bi[i,] <- coef(lm(F2~A+B+A*B))
bim[i,]  <-  coef(lm(F2~scale(A,scale=FALSE)  +  scale(B,scale=FALSE)  +
scale(A,scale=FALSE)*scale(B,scale=FALSE)))
bis[i,]   <-   coef(lm(scale(F2)~0   +   scale(A)   +   scale(B)   +
scale(A)*scale(B)))

varmatrix2 <- diag(rep(var(F2),3))

dat    <-    cbind(sim.data[,1:8],(F2        %*%    t(rep(1,3))    +
as.matrix(sim.data[,10:12]) %*% (sqrt(loadings2 * varmatrix2))))

write.table(dat,      "Simulation_Interaction_Form.csv",      sep=";",
append=TRUE, row.names=FALSE, col.names=FALSE)
}

round(colMeans(bi),3) # Normal Interaction Model
round(colMeans(bim),3) # Mean-Centered Interaction Model
round(colMeans(bis),3) # Standardized Interaction Model
round(apply(bi,2,sd),3)
round(apply(bim,2,sd),3)
round(apply(bis,2,sd),3)
```

APPENDIX C – ILLUSTRATIONS OF SIMULATION PATH MODELS

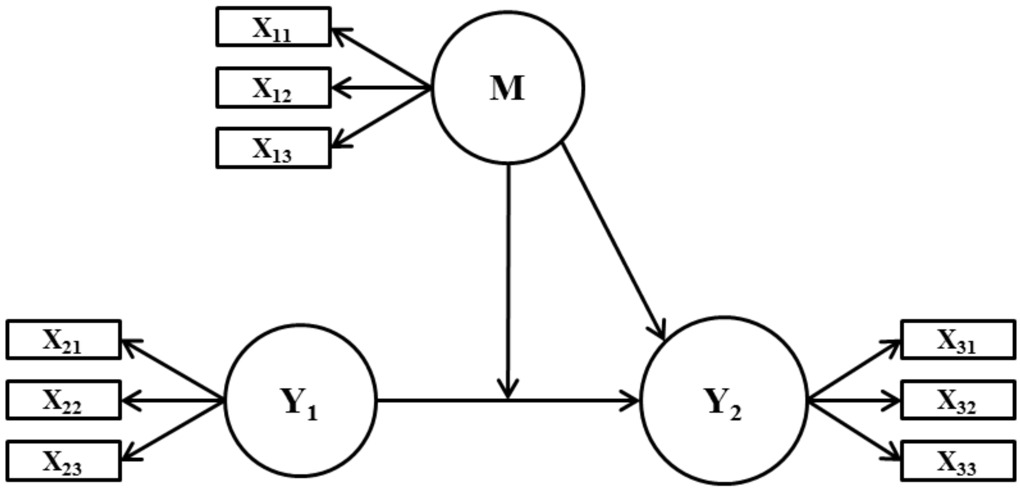

Figure C1. Model 1 (all constructs are reflective)

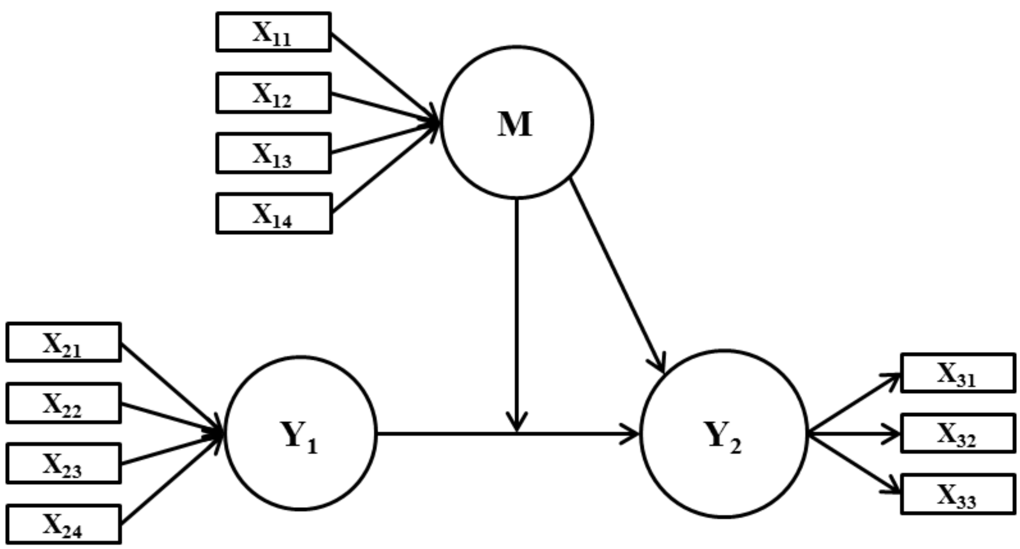

Figure C2. Model 2
(exogenous construct and moderator are formative, endogenous construct is reflective)

APPENDIX D – ADDITIONAL SIMULATION RESULTS FOR SAMPLE SIZE OF 100

Table B1. Mean regression estimates for Model 1 (sample size: 100)

	α **Mean** (Sd)	β_1 **Mean** (Sd)	β_2 **Mean** (Sd)	β_3 **Mean** (Sd)
Unstandardized	**4.72** (14.139)	**6.07** (4.523)	**3.05** (3.390)	**2.99** (1.080)
Mean-Centered	**71.08** (2.412)	**18.04** (1.071)	**12.03** (1.039)	**2.99** (1.080)
Standardized		**.750** (.050)	**.499** (.053)	**.122** (.044)

SD = standard deviation

Table B2. Mean regression estimates for Model 2 (sample size: 100)

	α **Mean** (Sd)	β_1 **Mean** (Sd)	β_2 **Mean** (Sd)	β_3 **Mean** (Sd)
Unstandardized	**4.48** (13.710)	**6.13** (4.311)	**3.14** (3.401)	**2.96** (1.070)
Mean-Centered	**71.03** (2.420)	**17.98** (1.104)	**12.04** (1.028)	**2.96** (1.070)
Standardized		**.749** (.050)	**.502** (.052)	**.122** (.044)

SD = standard deviation

Table B3. Simulation results of Model 1 (sample size: 100)

Population parameters: p_1 = .750; p_2 = .500; p_3 = .125

	PLS-SEM Mean	PLS-SEM SD	PLSc-SEM Mean	PLSc-SEM SD		PLS-SEM Mean	PLS-SEM SD	PLSc-SEM Mean	PLSc-SEM SD
Product-indicator (Mean-Centered)					**Orthogonalizing (Mean-Centered)**				
p_1	.635	.054	.744	.070	p_1	.644	.053	.750	.068
p_2	.421	.062	.488	.080	p_2	.425	.063	.491	.079
p_3	.063	.059	.064	.060	p_3	.081	.062	.087	.068
Product-indicator (Unstandardized)					**Orthogonalizing (Unstandardized)**				
p_1	.439	.254	-.264	3.088	p_1	.644	.053	.750	.068
p_2	.271	.200	-.288	3.043	p_2	.425	.063	.491	.079
p_3	.041	.051	.168	.614	p_3	.081	.062	.087	.068
Product-indicator (Standardized)					**Orthogonalizing (Standardized)**				
p_1	.636	.054	.743	.072	p_1	.644	.053	.750	.068
p_2	.421	.062	.488	.081	p_2	.425	.063	.491	.079
p_3	.097	.089	.117	.108	p_3	.121	.093	.156	.124
Two-Stage (*)					**Main Effects Only**				
p_1	.644	.054	.753	.071	p_1	.644	.053	.750	.068
p_2	.426	.062	.494	.081	p_2	.425	.063	.491	.079
p_3	.097	.070	.125	.097					

SD = standard deviation

Table B4. Simulation results of Model 2 (sample size: 100)

Population parameters: p_1 = .750; p_2 = .500; p_3 = .125

	PLS-SEM		PLSc-SEM			PLS-SEM		PLSc-SEM	
	Mean	**SD**	**Mean**	**SD**		**Mean**	**SD**	**Mean**	**SD**
Product-indicator (Mean-Centered)					**Orthogonalizing (Mean-Centered)**				
p_1	.645	.052	.698	.056	p_1	.676	.051	.732	.056
p_2	.426	.061	.461	.066	p_2	.447	.061	.484	.066
p_3	.021	.025	.023	.027	p_3	.048	.021	.052	.022
Product-indicator (Unstandardized)					**Orthogonalizing (Unstandardized)**				
p_1	.470	.130	.508	.141	p_1	.676	.051	.732	.056
p_2	.287	.116	.310	.125	p_2	.447	.061	.484	.066
p_3	.025	.015	.027	.017	p_3	.048	.021	.052	.022
Product-indicator (Standardized)					**Orthogonalizing (Standardized)**				
p_1	.643	.052	.695	.057	p_1	.676	.051	.732	.056
p_2	.425	.060	.459	.066	p_2	.447	.061	.484	.066
p_3	.095	.113	.103	.122	p_3	.216	.087	.234	.094
Two-Stage (*)					**Main Effects Only**				
p_1	.676	.052	.731	.057	p_1	.676	.051	.731	.056
p_2	.447	.061	.483	.066	p_2	.448	.061	.484	.066
p_3	.101	.057	.109	.062					

SD = standard deviation

