PLS-SEM: THE HOLY GRAIL FOR ADVANCED ANALYSIS

***LUCY MATTHEWS*, Middle Tennessee State University**
***JOE HAIR*, The University of South Alabama**
***RYAN MATTHEWS*, RLM Enterprises, LLC**

Advanced analytical techniques are reviewed for researchers wanting to expand their knowledge of how partial least squares structural equation modeling (PLS-SEM) facilitates better understanding of complex data relationships. We provide a brief overview of the differences between covariance-based structural equation modeling (CB-SEM) and PLS-SEM along with guidelines for the appropriate application of each. The focus is on mediation, moderation, multi-group analysis, and hierarchical component models. We also summarize several emerging analytical tools available with PLS-SEM. The ease of applying these advanced analytical techniques in many different research contexts makes PLS-SEM the "holy grail" for advanced analysis.

INTRODUCTION

For many researchers, keeping up with advanced methods can seem daunting. Learning new software along with the application and interpretation guidelines can sometimes be overwhelming. That is not the case, however, with partial least squares structural equation modeling (PLS-SEM), particularly using SmartPLS 3.0 (Ringle, Wende, & Becker, 2015). The recent rise in popularity of PLS-SEM can be attributed, at least in part, to the ease of understanding and applying the basic analytical tools of the method (Hair, Ringle, & Sarstedt, 2011). But learning to apply advanced methods such as mediation, moderation, multi-group analysis and more, is also relatively straightforward.

Most researchers are at least somewhat familiar with covariance-based structural modeling (CB-SEM), most often run with the AMOS or LISREL software. Few researchers are aware of and understand the fundamentals of variance -based structural modeling (PLS-SEM). The purpose of this paper is to introduce and provide an overview of the rapidly emerging method of variance-based structural equation modeling. In this paper, we first explain the differences in variance-based structural equation modeling (PLS-SEM) and the covariance-based CB-SEM method, and therefore the rationale for the selection of one approach over another. We then summarize several of the more advanced analytical tools available when applying PLS-SEM.

The Marketing Management Journal
Volume 28, Issue 1, Pages 1-13

PLS-SEM versus CB-SEM

When it comes to structural equations modeling (SEM), researchers have a choice of two methods: covariance-based SEM (CB-SEM; Jöreskog, 1978, 1993) and variance-based partial least squares (PLS-SEM; Lohmöller, 1989; Wold, 1982). A fundamental distinction between the two approaches is that CB-SEM is based on the common factor model, while PLS-SEM is based on the composite factor model (Hair, Hult, Ringle, & Sarstedt, 2017). With common factor models, the analysis is based only on the common variance in the data. Therefore, the solution begins by calculating the covariances between the variables in the study and only that common variance is used in the analysis (Hair, Matthews, Matthews, & Sarstedt, 2018; Sarstedt, Hair, Ringle, Theile, & Gudergan, 2016). With the composite factor model the constructs and their scores are represented by the total variance in the indicators, not just common variance that is the case with CB-SEM (Hair, Hult, Ringle, & Sarstedt, 2017). In addition, the statistical objectives are substantially different between the two methods. Using CB-SEM, the statistical objective is to estimate model parameters that minimize the differences between the observed sample covariance matrix, which is calculated before the theoretical model solution is obtained, with the covariance matrix that is estimated after the theoretical model solution is obtained (Hair, Sarstedt, Ringle, & Mena, 2012). If goodness of fit is (GOF) demonstrated, the theoretical structural model is confirmed. But if GOF is not possible the model is not confirmed. In contrast, when using PLS-SEM, the statistical objective is to maximize the variance explained in the dependent variable(s) (Hair, Sarstedt, Pieper, & Ringle, 2012a). Thus, the focus of PLS is on optimizing prediction of the endogenous constructs and not on fit, which is the focus of CB-SEM. Moreover, PLS-SEM is a variance-based approach and the analysis does not start or end with a covariance matrix. Thus, a Chi-square type of GOF is not possible.

Determining when the application of each of the methods is appropriate is straightforward. If the focus of the research is theory testing and confirmation (Sarstedt, Ringle, Henseler, & Hair, 2014), then CB-SEM is the appropriate method. But if prediction, theory development and explanation are the focus of the research, then PLS-SEM is the more appropriate method. PLS-SEM is somewhat similar, both conceptually and practically, to using multiple regression analysis (Hair et al., 2011). But unlike multiple regression, PLS-SEM can be applied to better understanding more complex structural measurement and path models. At the same time, PLS-SEM and CB-SEM are both appropriate for metric data and reflective measurement models. But PLS-SEM can easily be used with formative measurement models, non-metric data (e.g., ordinal & nominal), continuous moderators, higher order models, when latent variable scores are needed for further analysis, and with small sample sizes ($N \leq 100$) as well as large samples (Hair, Hollingsworth, Randolph, & Chong, 2017). Because it is nonparametric, PLS-SEM also has a wider application and greater flexibility in handling various modeling situations where it is difficult to meet rigorous assumptions, such as a normal distribution and homoscedasticity, that are typically required with more traditional multivariate statistics (Vinzi, Chin, Henseler, & Wang, 2010). Therefore, PLS-SEM is appropriate for exploratory research, theory development, and prediction. It can be executed on both small and large samples sizes, with reflective or formative measurement models, and does not assume the data has a normal distribution. Finally, the method can be used with metric and non-metric data, continuous moderators, secondary data, higher order models, multi-group analysis, invariance and unobserved heterogeneity, making PLS-SEM a “go to” methodology for researchers.

Mediation

When a third variable, called a mediator, intervenes between two other variables, the opportunity arises to examine mediation (Baron & Kenny, 1986). Specifically, a change in the exogenous construct produces a change in the mediator variable, which then produces a change in the endogenous construct, and the mediator variable dictates the nature of the relationship between the exogenous and endogenous constructs (Hair, Hult, et al., 2017). A crucial prerequisite for investigating mediating effects is strong apriori theoretical support (Preacher & Hayes, 2008).

The foundation for mediation is a well-established theoretical relationship (c) between an exogenous (Y_1) construct and an endogenous (Y_3) construct (Figure 1) (Preacher & Hayes, 2008). Testing for mediation in a model requires a series of analyses beginning with testing the significance of the indirect effect (a b) via the mediator variable (Y_2) as seen in Figure 2. If the indirect effect is not significant, then Y_2 is not operating as a mediator in the relationship. However, if (a b) is significant, then the next test is to check the direct effect in the mediated model (c’). If c’ is not significant, indirect-only (full) mediation has occurred. This occurs when the indirect effect is significant, but not the direct effect in the mediated model. Alternatively, if c’ is significant, then partial mediation has occurred. When (a b c’) is positive, then complementary mediation has occurred, but if (a b c’) is negative then competitive mediation has occurred (Hair, Hult, et al., 2017).

FIGURE 1:
Direct Effect of Exogenous Variable on Endogenous Variable

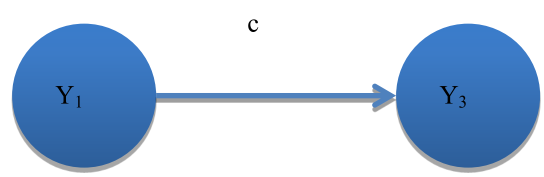
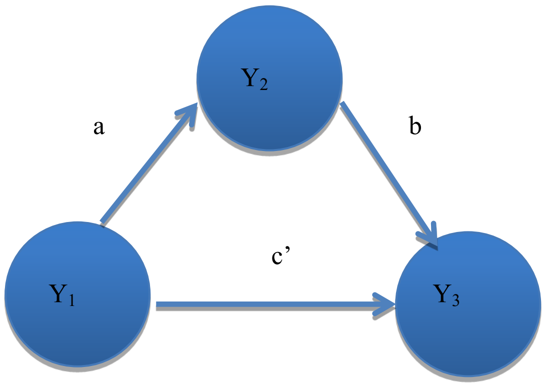


FIGURE 2:
Indirect Effect - Mediation Model



Mediation has traditionally been executed using multiple regression. Baron and Kenny (1986) and more recently the PROCESS approach by Preacher and Hayes (2008) both focus on multiple regression and examine significance in mediation using the Sobel test, which assumes the data are normally distributed. The advantages of using PLS-SEM for mediation are that bootstrapping makes no assumptions about the shape of the variables’ distribution or the sampling distribution of the statistics, and all the mediated relationships are tested simultaneously instead of separately, which reduces bias (Hair, Sarstedt, Hopkins, & Kuppelwieser, 2014). Finally, mediation testing using PLS-SEM can be applied with smaller sample sizes while yielding higher levels of statistical power compared to prior testing methods like the parametric Sobel (1982) test.

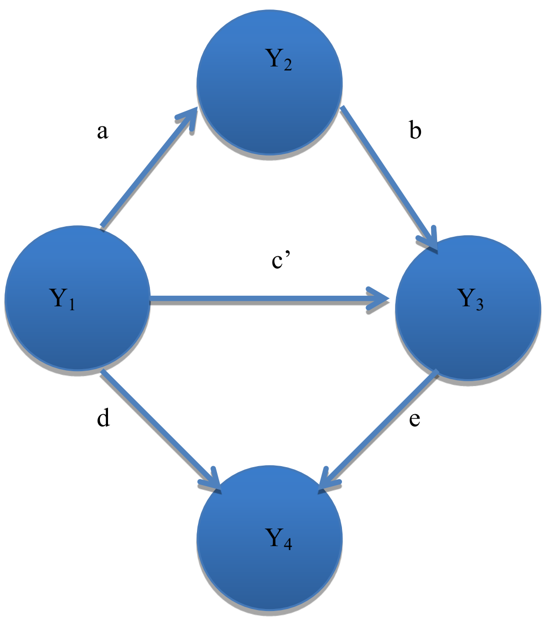
Path models that include a mediator are still required to meet the quality criteria of the measurement models. For formative measurement models, convergent validity (redundancy), collinearity between indicators, and significance/relevance of outer weights are required (Hair, Hult, et al., 2017). For reflective measurement models, the quality criteria include internal consistency reliability, convergent validity, and discriminant validity. It is important to also confirm that collinearity in the structural model is not at a critical level since biased path coefficients may incur. When high collinearity exists, the direct effect may suggest nonmediation via nonsignificance or an unexpected sign change may result in an erroneous differentiation between complementary and competitive mediation (Hair, Hult, et al., 2017). With complementary mediation, the mediated effect (a b) and direct effect (c’) both exist and point in the same direction (i.e., the signs are either both positive or both negative), while with competitive mediation, the mediated effect (a b) and direct effect (c’) are both present and point in opposite directions (i.e., a sign for one relationship is positive and the other relationship is negative) (Zhao, Lynch, & Chen, 2010).

The most common types of mediation are simple mediation analysis and multiple mediation analysis. Simple mediation analysis is when one mediator variable is specified in the structural model. Often times, however, exogenous constructs influence endogenous constructs through more than one mediating variable requiring multiple mediation analyses (Hair, Hult, et al., 2017). The mediators reveal the “real” relationship among the exogenous and endogenous constructs. Figure 3 depicts a model with two mediating variables, Y_2 and Y_4. The direct effect is measured by c’. But the indirect effect of Y_1 on Y_3 now includes the Y_4 mediator (d e) in addition to Y_2, and the total indirect effect of Y_1 on Y_3 is measured by the sum of the two indirect effects (i.e., a b + d e). Therefore, the total effect of Y_1 on Y_3 is the sum of the direct effect and the total indirect effect (i.e., c’ + a b + d e).

Analyzing all mediators concurrently allows for a more thorough understanding of the overall effect. If each mediator were analyzed in a simple mediation analysis (i.e., with a regression model where all relationships are tested separately), the indirect effect would likely be overstated due to the correlation of one mediator to another (Hair, Hult, et al., 2017). When PLS-SEM is used, multiple mediation in which all relationships (direct and indirect) are tested simultaneously is possible. Thus, with multiple mediation the impact of one or multiple mediators can be tested simultaneously, eliminating the overstatement of the correlation associated with each mediator.

The steps for processing multiple mediation analyses are the same as in simple mediation analysis. The testing process begins with examining the significance of each indirect effect, and then the direct effect between the exogenous and endogenous constructs. To determine the total indirect effect manual calculations of the standard error for each specific indirect effect may be necessary. Using SmartPLS 3.0 (Ringle et al., 2015), however, this can be accomplished by simply obtaining the indirect effects results of the bootstrapping routine and using spreadsheet software such as Microsoft Excel. Finally, to calculate the *t* value of the specific indirect effect, divide the specific indirect effect by the standard error (Hair, Hult, et al., 2017).

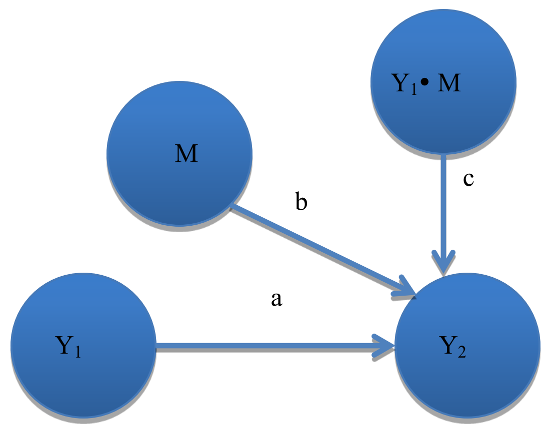
FIGURE 3:
Multiple Mediation Model



Moderation

Moderation (interaction effect) occurs when the relationship between two constructs varies depending on a third (moderator) variable (Henseler & Chin, 2010). The variation can influence the strength or direction of the relationship (Baron & Kenny, 1986). Moderator variables can be categorical (e.g., age, income, gender) and tested by means of group comparisons using either PLS-SEM or CB-SEM. Alternatively, with PLS-SEM moderators can also be a continuous variable (e.g., attitudes about satisfaction, loyalty, commitment, brand passion) typically measured using multi-item scales (note that continuous moderators cannot be used with CB-SEM). When including a moderator in the model, the variable will appear twice, once as the variable itself (main effect) and again as the interaction effect (a combination of the main effect and the moderator; see Figure 4). Unlike mediation where the exogenous construct acts as an antecedent to the mediator, in moderation the moderator variable and exogenous construct interact (Y_1*M) at the same level to impact the endogenous variable. This is a multiplicative relationship.

FIGURE 4:
Moderation Model



While several analytical procedures exist for estimating the measurement model with moderation (e.g., product indicator approach, orthogonalizing approach, and two-stage approach), the two-stage approach is typically recommended. The two-stage approach is able to handle both reflective and formative moderators and additional exogenous constructs in the path model. Moreover, compared to the other approaches, the two-stage approach exhibits higher levels of statistical power (Hair, Hult, et al., 2017).

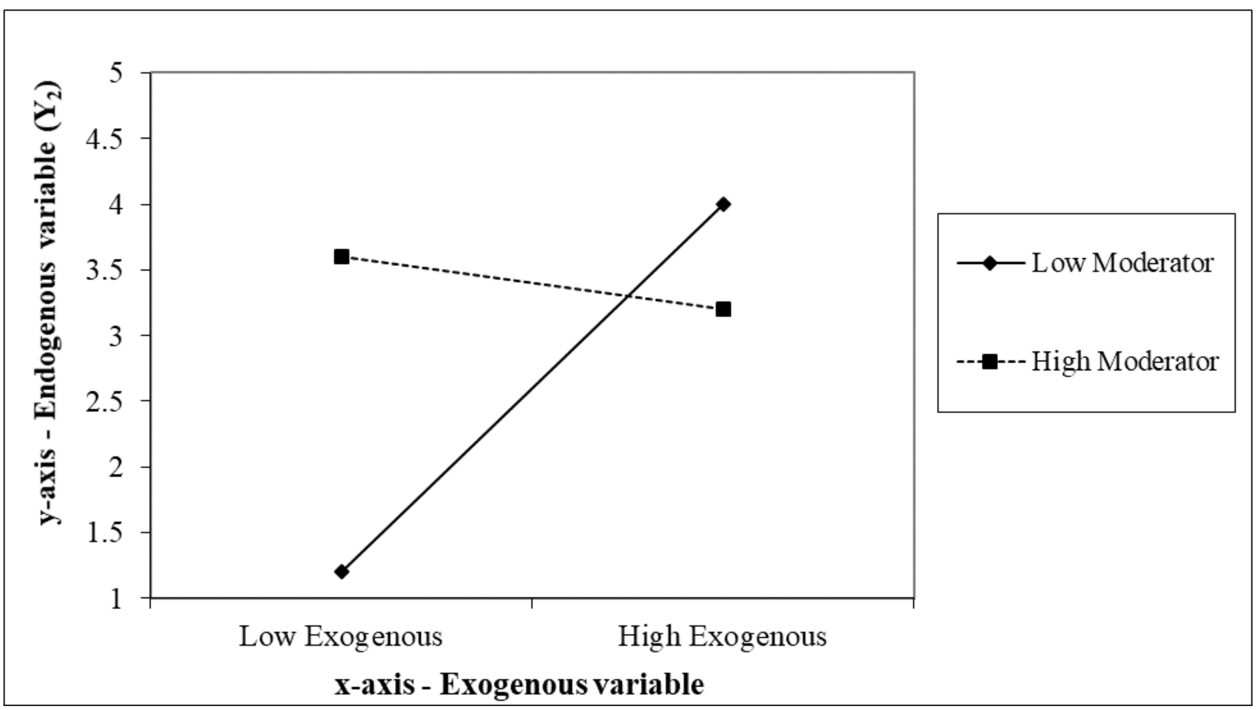
The two-stage approach begins by running the main effects model without the moderation interaction term in the model to estimate the latent variable scores (Henseler & Chin, 2010). In the second stage, the latent variable scores from stage one of the exogenous latent variable and the moderator variables are multiplied to create a single-item measure to represent the interaction term (Hair, Hult, et al., 2017). At the same time, all the other latent variables are represented by a single item measure (latent variable score) that was calculated in stage one. The moderator hypothesis is supported if the interaction effect (c) is significant (Hair, Sarstedt, Ringle, & Gudergan, 2018).

The results indicate that the value of c (interaction effect) represents the strength of the relationship between Y_1 and Y_2 when the moderator variable M has a value of zero (Hair, Hult, et al., 2017). However, since many scales either do not include a value of zero or a value of zero does not make sense, standardization is often necessary. Standardization facilitates interpretation as well as reduces collinearity among the exogenous construct, the moderator, and the interaction term. To standardize, the variable's mean is subtracted from each observation and divided by the variable's standard error (Sarstedt & Mooi, 2014). When using SmartPLS 3, the software executes many types moderation, standardizes when necessary, and produces a simple slope analysis for interpreting moderation results.

To assess the impact on the R^2 value when the interaction effect is omitted from the model, the f^2 effect size is examined. The f^2 measures the extent to which the endogenous latent variable is explained by the moderation. The f^2 effect sizes of 0.02, 0.15, and 0.35 suggest small, medium, and large effect sizes, respectively (Cohen, 1988).

Interpreting and drawing conclusions from the moderation results can be difficult. Slope plots are typically used as a visual illustration to gain a better understanding of the moderation effect. Figure 5 displays a two-way interaction of the relationship between Y_1 and Y_2. The horizontal x-axis represents the exogenous construct (Y_1) and the vertical y-axis represents the endogenous construct (Y_2). The two lines illustrate the relationship between Y_1 and Y_2 for both low and high levels of the moderator construct (M). The low level of M (solid line) is one standard deviation unit below the average, while the high level of M (dotted line) is one standard deviation unit above the average. There is a negative moderating effect of -0.80 between the interaction term and the endogenous construct. The high moderator's slope is relatively flat but decreases slightly as

FIGURE 5:
Graphical Illustration of Moderation Effect



the exogenous variable changes from low to high. Thus, the relationship between Y_1 and Y_2 becomes weaker with high levels of the moderator construct. But for low levels of the moderator variable, the slope is quite steep and the relationship between Y_1 and Y_2 becomes much stronger with high levels of M. To facilitate interpretation of the interaction, SmartPLS 3 computes and the output displays a simple slope plot.

Multi-Group

Multi-group analysis (MGA) or between-group analysis is a means of testing apriori defined groups to determine if there are significant differences in group-specific parameter estimates (e.g., outer weights, outer loadings and path coefficients) obtained when using PLS -SEM (Hair, Hult, Ringle, & Sarstedt, 2014; Henseler & Chin, 2010). By applying MGA, researchers are able to test for differences between two identical models for different apriori specified groups within the data set. In contrast to standard approaches to testing moderation, which examine a single structural relationship at a time, MGA via PLS-SEM is an efficient way to assess moderation across multiple relationships (Hair, Sarstedt, Ringle, et al., 2012).

This type of analysis enables researchers to identify differences between the structural paths of multiple groups. For example, Matthews (2017) illustrated how skill discrepancy partially mediates the relationship between autonomy and cognitive engagement for male salespeople, but not for female salespeople. PLS-MGA also facilitates a more accurate and comprehensive assessment of group differences and strategy implementation based on more specific outcomes for the heterogeneous groups in the data (Matthews, 2017). Finally, the differences identified can be used to highlight the potential error if these subpopulations are considered a single homogeneous group (Schlagel & Sarstedt, 2016).

The first step in conducting MGA involves generating data groups that are based on the categorical variable of interest (e.g., gender, country of origin, age, or income). Once the data is subdivided, it is necessary to ensure that the sample sizes of the new subgroups are large enough (See Hair, Hult, et al., 2014). Additionally, the subgroups should be similar in size to avoid introducing error (Becker, Rai, Ringle, & Volckner, 2013).This involves coding the master data file into subgroups that can be executed with PLS-SEM (Matthews, 2017).

While a number of approaches can be used to compare the path coefficients of the group SEMs, most researchers recommend the permutation test (Hair, Sarstedt, Ringle, & Gudergan, 2018). Permutation is non-parametric and more conservative than the parametric test, and controls well for Type 1 error (Henseler, Ringle, & Sarstedt, 2016). To execute the permutation test, the correlations between the composite scores using the weights obtained from the first group are computed against the composite scores using the weights obtained from the second group during each permutation run (Henseler et al., 2016).

The focus of multi-group analysis is to examine the path coefficients of the theoretical models for the two groups to determine if they are significantly different. This begins by running the model for each group separately, using the guidelines set out for evaluation of a measurement model (Hair, Hult, et al., 2014, 2017). This determines if there are group specific differences. Then, it is necessary to determine if the difference between the two groups is significant, which can be accomplished by running the Permutation Test. A permutation p-value of less than or equal to 0.10 indicates a significant difference between the two groups being compared (Matthews, 2017).

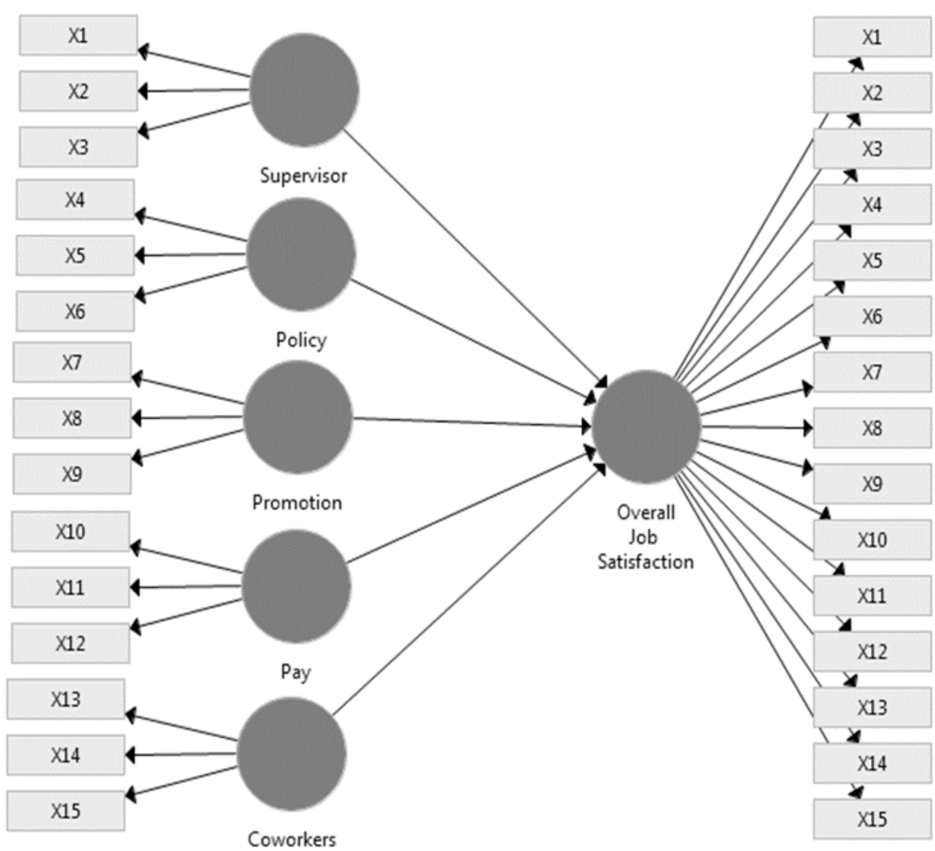
MGA allows researchers to determine significant differences among observed characteristics. These differences may not be evident in aggregate data since significant positive and negative group-specific results may offset one another resulting in non-significant relationships (Hair, Hult, et al., 2017). Although the path coefficients for the subdivided groups will often display numerical differences, MGA assists in identifying when those differences are statistically significant.

Hierarchical Component Models

Hierarchical component models (HCM), or higher order models, involve testing measurement structures that contain two layers of constructs, the higher order component (HOC) of the model and the lower order component (LOC). An example of a HCM is when lower order components of job satisfaction are specified as multi-faceted and the higher order component is a single overall construct of job satisfaction. Figure 6 displays a theoretical HCM for job satisfaction that includes the LOCs (lower order components) and the HOC (higher order component). In the figure, the LOCs represent the first-order multi-item constructs for job satisfaction, and the HOC is the second-order overall (combined) construct for job satisfaction. Researchers may find the use of a HCM helpful when trying to reduce the number of relationships in the structural model, making the model more parsimonious and easier to understand. In addition, introducing a HCM into a structural model can reduce multicollinearity among first-order constructs, or formative indicators that exhibit high levels of collinearity. In either situation, the use of HCMs should be supported by theory. Note that higher order models can also be used with CB-SEM, but the assumptions are much more restrictive and therefore limit their application with that method.

There are four main types of HCMs (Jarvis, MacKenzie, & Podsakoff, 2003; Wetzels, Odekerken-Schroder, & van Oppen, 2009) used in PLS-SEM models (Ringle, Sarstedt, & Straub, 2012). The HCM model begins with the lower-order components (LOCs), which are used to make up the higher-order component (HOC) (Hair, Hult, et al., 2017). Each model is

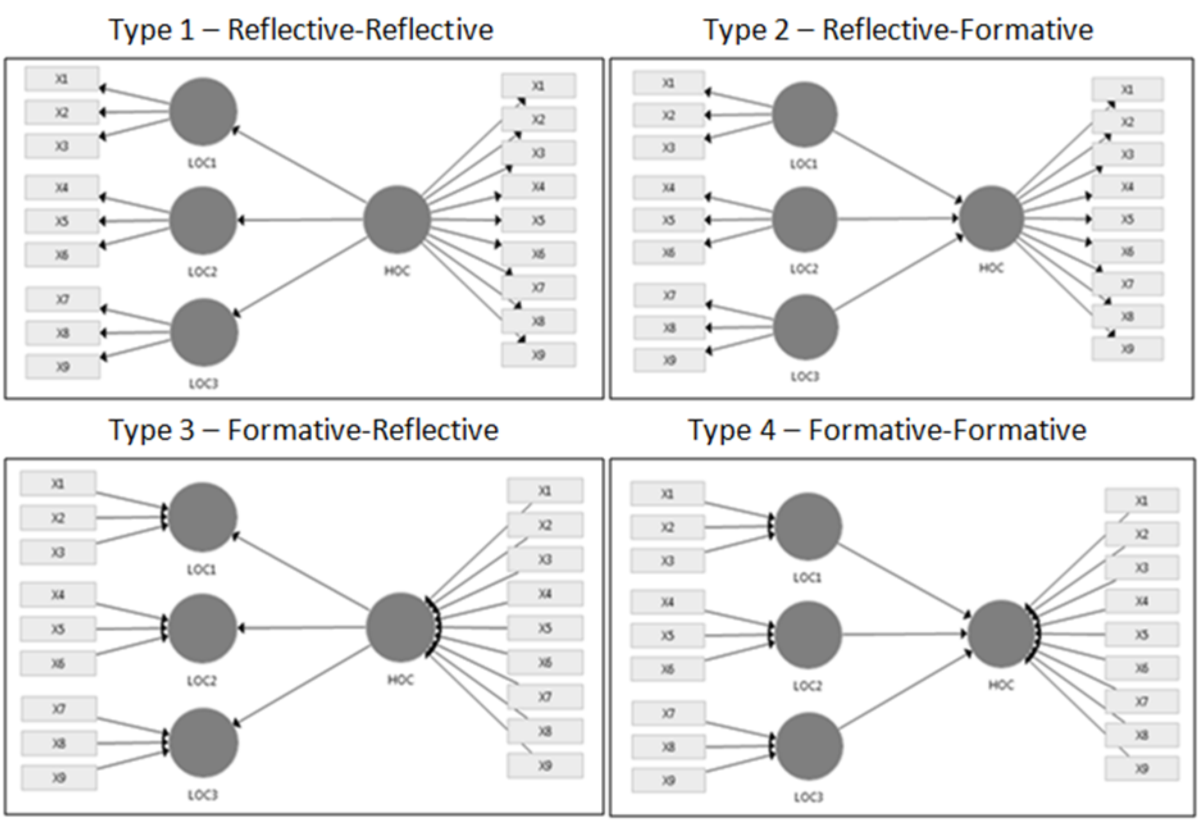
FIGURE 6:
Hierarchical Component Model for Multi-faceted Job Satisfaction



characterized by the different relationships between the LOCs and the HOC as well as the indicators with each construct. The first type is the Reflective-Reflective, where the indicator measures for the first order components are reflective and the measures from the LOC to the HOC are reflective (Figure 7). Type two is Reflective-Formative. For this model, the LOC indicators are reflective but the LOC to the HOC is formative. Type three is Formative-Reflective, such that the first-order indicators are measured formatively and the HOC from the LOC is reflective. The final type (type four), is Formative-Formative where the indicators of the first-order are measured formatively and the measures from the LOC to the HOC are formative.

When creating the HOC in PLS-SEM, all the indicators from the LOC are assigned to the HOC using a repeated indicators approach (Hair, Hult, et al., 2017). Therefore, the indicators for the HOC x_1 to x_9 are the same as the underlying components LOC_1, LOC_2, and LOC_3 in the measurement model. However, some issues arise using the repeated indicator approach when the model is formative-formative (type 4) or reflective-formative (type 2). In this situation, when the relationship from the LOC to the HOC is formative, almost all of the HOC variance is explained by the LOC (R^2 close to 1.0). This can be an issue if there are other relationships pointing to the HOC, as they will have a very small and insignificant impact. Therefore, for type 2 and type 4 models, a two-stage HCM analysis should be used (Hair, Hult, et al., 2017). Similar to the two-stage approach for moderation, in stage one the repeated indicator approach is used to obtain the latent variable scores for the LOCs. Then in the second stage, the LOC constructs and first-order indicators used for the HOC are replaced with latent variable scores for the LOC from stage one (Figure 8). The two-stage HCM analysis

FIGURE 7:
Four Types of Hierarchical Component Models



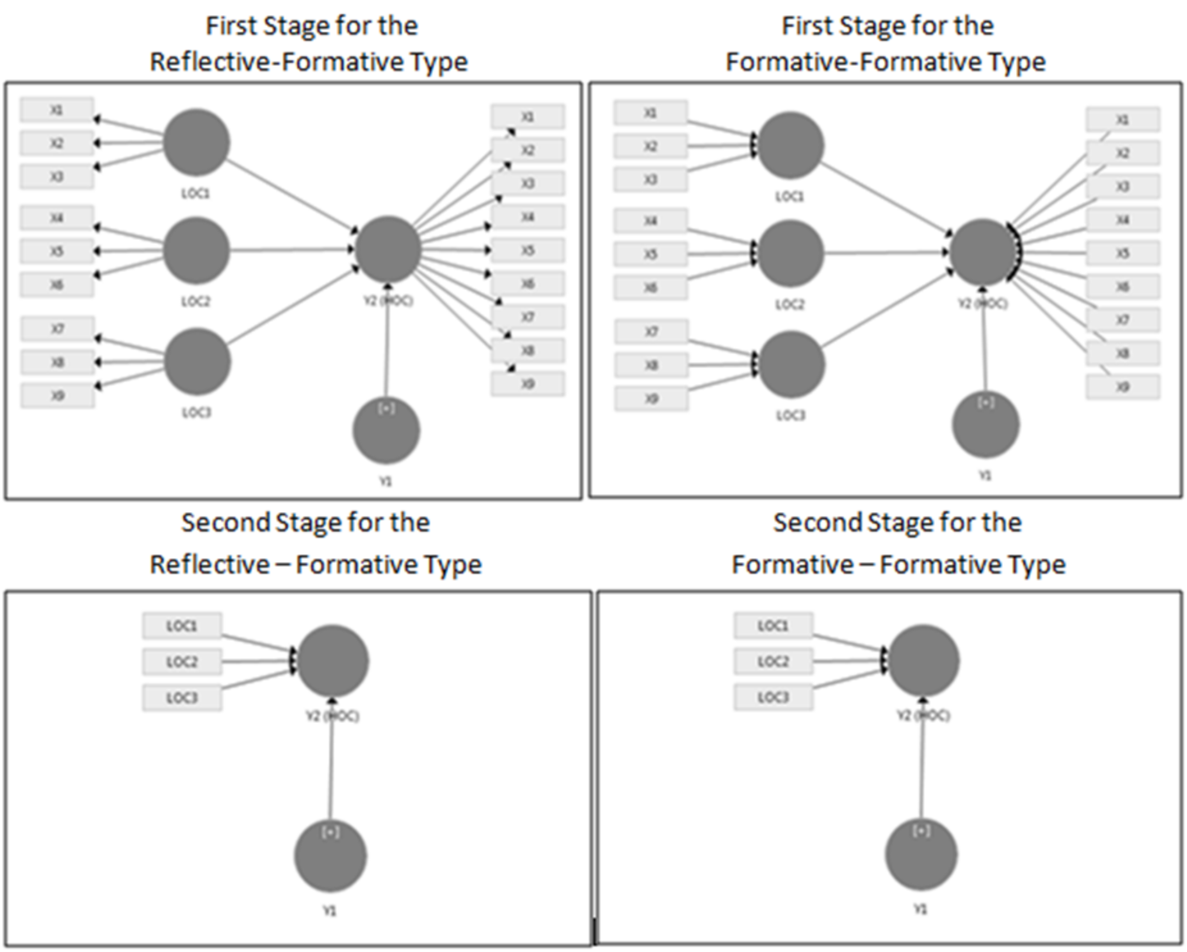
allows other latent variables outside of the HOC to explain some of the variance.

When using HCM it is important that a similar number of indicators are used for all the LOCs. Otherwise, the relationship between the HOC and the LOC can be biased due to the disproportionate number of indicators (Hair, Hult, et al., 2017). Note that the number of indicators on the LOCs does not have to be equal (as shown in Figures 6 & 7), but should be comparable. Additionally, for the inner PLS path model, not all algorithmic weighting schemes apply when estimating HCMs in PLS-SEM. In particular, the centroid weighting scheme should not be used (Hair, Sarstedt, Ringle, et al., 2012). Prior research using and explaining HCM models can further assist in understanding and explaining the outcomes of this technique (e.g., Becker, Klein, & Wetzels, 2012; Kuppelwieser & Sarstedt, 2014; Ringle et al., 2012).

Other Advanced Topics

In addition to the topics addressed above, researchers have further opportunities to improve their analysis and understanding of theoretical relationships. Measurement model invariance, which tests datasets for differences in measurement model estimates, is a useful tool that should be combined with multigroup analysis. By employing the measurement invariance of composite models (MICOM) procedure (Henseler et al., 2016), configural and compositional invariance can be

FIGURE 8:
Two-Stage Approach for HCM Analysis



established. Doing so ensures that variations in the path relationships between latent variables is a result of the true differences in the structural relationships, and is not the result of different meanings in the groups' responses attributed to the phenomena being measured (Hair, Hult, et al., 2017; Henseler et al., 2016). Failure to establish data equivalence using MICOM may potentially result in measurement error and thus misleading results (Hult et al., 2008), reduce the overall power of the statistical tests, and influence the precision of the estimators (Hair, Hult, et al., 2017).

The importance-performance map analysis (IPMA), or importance-performance matrix analysis, displays the structural model total effects on a specific endogenous construct. The total effects of the predecessor variables are used to assess each exogenous construct's importance in shaping the endogenous construct. The average latent variable scores of the exogenous constructs measure their performance (Hair, Hult, et al., 2017) using a rescaling technique (Höck, Ringle, & Sarstedt, 2010). Combined, researchers can identify constructs with relatively high importance (strong total effect) and low performance (low average latent variable scores) as areas for further research. IPMA can be conducted at the indicator level as well to identify and improve upon those indicators that are most relevant.

Finally, rather than using a priori characteristics to partition datasets into groups, as was described in multi-group analysis, tools like finite mixture PLS (FIMIX-PLS, Sarstedt, Becker, Ringle, & Schwaiger, 2011) or prediction-oriented segmentation (FIMIX-POS, Sarstedt, Ringle, & Hair, 2017) can be used to uncover unobserved heterogeneity. Since sources of heterogeneity in the data aren't always known a priori, identifying and treating unobserved heterogeneity allows researchers to feel confident about analyzing data at an aggregate level (Hair, Sarstedt, Matthews, & Ringle, 2016). Examples of FIMIX-PLS are available to aid researchers in the application to their own dataset (Matthews, Sarstedt, Hair, & Ringle, 2016; Sarstedt, Schwaiger, & Ringle, 2009). Failure to consider heterogeneity may also result in invalid outcomes (Becker et al., 2013).

Summary

When analyzing research that requires advanced analytical approaches, it is important to understand the differences between CB-SEM and PLS-SEM, as well as other multivariate analysis methods. Because PLS-SEM has a much greater capacity for handling a variety of modeling issues and does not impose restrictive assumptions (Vinzi et al., 2010), the use of PLS -SEM is highlighted. For mediation, the advantage of using PLS-SEM is the lack of restrictive distribution assumptions, the flexibility to execute with both formative and reflective measurement models, and the ability to yield higher levels of statistical power with smaller sample sizes, while overcoming the limitations of multiple regression approaches (Hair, Sarstedt, et al., 2014). The two-stage approach for moderation using PLS-SEM exhibits high levels of statistical power and is capable of handling both reflective and formative moderators when the structural model includes other exogenous constructs. Multi-group analysis via permutation in PLS-SEM enables researchers to easily identify heterogeneous groups in the data to more accurately assess group differences (Matthews, 2017). Finally, higher order component models (HCMs) can be applied with PLS-SEM to obtain more accurate solutions for structural models exhibiting high multicollinearity.

Beyond these advanced analysis approaches, PLS-SEM can: (1) establish data equivalence via the three stage MICOM process (Henseler et al., 2016) to minimize measurement error, (2) identify the importance and performance of antecedent constructs to target areas for further research (Hair, Hult, et al., 2017), and (3) uncover unobserved heterogeneity so structural and measurement models can be examined either at the individual group level or the aggregate level (Matthews et al., 2016; Hair et al. 2016). Therefore, when facing complex research models, even though there are a variety of multivariate methods available, the numerous flexible analysis options, limited assumptions, and user-friendliness of PLS-SEM make it the "holy grail" for advanced methods.

REFERENCES

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), 1173-1182.

Becker, J.-M., Klein, K., & Wetzels, M. (2012). Formative Hierarchical Latent Variable Models in PLS-SEM: Recommendations and Guidelines. *Long Range Planning, 45*(5-6), 359-394.

Becker, J.-M., Rai, A., Ringle, C. M., & Volckner, F. (2013). Discovering unobserved heterogeneity in structural equation models to avert validity threats. *MIS Quarterly, 37*(3), 665-694.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Mahwah, NJ: Lawrence Erlbaum.

Hair, J. F., Hollingsworth, C. L., Randolph, A. B., & Chong, A. (2017). An Updated and Expanded Assessment of PLS-SEM in Information Systems Research. *Industrial Management & Data Systems* (3), 442-458.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2014). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Los Angeles, London, New Delhi, Singapore, Washington DC: SAGE.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (2nd ed.). Los Angeles, London, New Delhi, Singapore, Washington DC: SAGE.

Hair, J. F., Matthews, L. M., Matthews, R. L., & Sarstedt, M. (2017). PLS-SEM or CB-SEM: Updated Guidelines on Which Method to Use. *International Journal of Multivariate Data Analysis*,*1*(2), 107-123.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *The Journal of Marketing Theory and Practice, 19*(2), 139-152.

Hair, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review, 26*(2), 106-121.

Hair, J. F., Sarstedt, M., Matthews, L. M., & Ringle, C. M. (2016). Identifying and Treating Unobserved Heterogeneity with FIMIX-PLS: Part I - Method. *European Business Review, 28*(1) 63-76.

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012). The Use of Partial Least Squares Structural Equation Modeling in Strategic Management Research: A Review of Past Practices and Recommendations for Future Applications. *Long Range Planning, 45*(5-6), 320-340.

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

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An Assessment of the use of Partial Least Squares Structural Equation Modeling in Marketing Research. *Journal of the Academy of Marketing Science, 40*(3), 414-433.

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: A Multidisciplinary Journal, 17*(1), 82-109.

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

Höck, C., Ringle, C. M., & Sarstedt, M. (2010). Management of Multi-purpose Stadiums: Importance and Performance Measurement of Service Interfaces. *International Journal of Services Technology and Management, 14*(2-3), 188-207.

Hult, G. T. M., Ketchen Jr, D. J., Griffith, D. A., Finnegan, C. A., Gonzalez-Padron, T., Harmancioglu, N., & et al. (2008). Data Equivalence in Cross-Cultural International Business Research: Assessment and Guidelines. *Journal of International Business Studies, 39*(6), 1027-1044.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research. *Journal of Consumer Research, 30* (2), 199-218.

Jöreskog, K. G. (1978). Structural Analysis of Covariance and Correlation Matrices. *Psychometrika, 43*(4), 443-477.

Jöreskog, K. G. (1993). Testing Structural Equation Models. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 294-316). Newbury Park: Sage.

Kuppelwieser, V., & Sarstedt, M. (2014). Applying the Future Time Perspective Scale to Advertising Research. *International Journal of Advertising, 33*(1), 113-136.

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

Matthews, L. M. (2017). Applying Multi-group Analysis in PLS-SEM: A Step-by-Step Process. In H. Latan & R. Noonan (Eds.), *Partial Least Squares Structural Equation Modeling - Basic Concepts, Methodological Issues and Applications*. Berlin: Springer International Publishing AG.

Matthews, L. M., 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.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and Resampling Strategies for Assessing and Comparing Indirect Effects in Multiple Mediator Models. *Behavior Research Methods, 40*(3), 879-891.

Ringle, C. M., Sarstedt, M., & Straub, D. W. (2012). A Critical Look at the use of PLS-SEM in MIS Quarterly. *MIS Quarterly, 36* (1), iii-xiv.

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

Sarstedt, M., Becker, J.-M., Ringle, C. M., & Schwaiger, M. (2011). Uncovering and treating unobserved heterogeneity with FIMIX-PLS: which model selection criterion provides an appropriate number of segments? *Schmalenbach Business Review, 63*, 34-62.

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

Sarstedt, M., & Mooi, E. A. (2014). *A Concise Guide to Market Research: The Process, Data, and Methods using IBM SPSS Statistics* (2nd ed.). Berlin: Springer.

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2017). Treating Unobserved Heterogeneity in PLS-SEM: A Multi-method Approach. In H. Latan & R. Noonan (Eds.), *Partial Least Squares Structural Equation Modeling - Basic Concepts, Methodological Issues and Applications*. Berlin: Springer International Publishing AG, 197-213.

Sarstedt, M., Ringle, C. M., Henseler, J., & Hair, J. F. (2014). On the Emancipation of PLS-SEM: A Commentary on Rigdon (2012). *Long Range Planning, 47*(3), 154-160.

Sarstedt, M., Schwaiger, M., & Ringle, C. M. (2009). Do We Fully Understand the Critical Success Factors of Customer Satisfaction with Industrial Goods? Extending Festge and Schwaiger's Model to Account for Unobserved Heterogeneity. *Journal of Business Market Management, 3*(3), 185-206.

Schlagel, C., & Sarstedt, M. (2016). Assessing the Measurement Invariance of the Four-Dimensional Cultural Intelligence Scale Across Countries: A Composite Model Approach. *European Management Journal, 34* (6), 633-649.

Sobel, M. E. (1982). Asymptotic Confidence Intervals for Indirect Effects in Structural Equations Models. *Sociological Methodology, 13*, 290-312.

Vinzi, V. E., Chin, W. W., Henseler, J., & Wang, H. (2010). Editorial: Perspectives on Partial Least Squares. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of Partial Least Squares: Concepts, Methods and Applications*. Heidelberg Dordrecht London New York: Springer.

Wetzels, M., Odekerken-Schroder, G., & van Oppen, C. (2009). Using PLS Path Modeling for Assessing Hierarchical Construct Models: Guidelines and Empirical Illustration. *MIS Quarterly, 33*, 177-195.

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

GLOSSARY

Bootstrapping: a resampling technique that draws a specified (large) number of subsamples from the original data and using replacement, estimates models for each subsample. It is used to assess statistical significance without relying on distributional assumptions to determine standard errors of coefficients.

Collinearity: when two variables are highly correlated.

Formative measurement model: a type of measurement model setup in which the direction of the arrows is from the indicator variables to the construct, thus indicating an assumption that the indicator variables cause the measurement of the construct.

Orthogonalizing approach: an approach to model the interaction term when including a moderator variable in the model. This creates an interaction term with orthogonals indicators. In the moderator model, these orthogonal indictors are not correlated with independent variable indicators and the moderator variable indicators.

Product indicator approach: an approach to model the interaction term when including a moderator variable in the model. This approach involves multiplying the indicators of the moderator with the indicators of the exogenous latent variable to establish a measurement model of the interaction term. The approach is only applicable when both moderator and exogenous latent variables are reflective.

Reflective measurement: a type of measurement model setup in which measures the direction of the arrow is from the construct to the indicator variables, thus the measures represent the effects (or manifestations) of an underlying construct. Causality is from the construct to its measures (indicators).