The current issue and full text archive of this journal is available at
www.emeraldinsight.com/0955-534X.htm

Partial least squares structural equation modeling (PLS-SEM)

An emerging tool in business research

Joe F. Hair Jr
Department of Marketing & Professional Sales, Kennesaw State University, Kennesaw, Georgia, USA

Marko Sarstedt
Otto-von-Guericke-University Magdeburg, Magdeburg, Germany and University of Newcastle, Newcastle, Australia

Lucas Hopkins
Middle Georgia State College, Macon, Georgia, USA, and

Volker G. Kuppelwieser
NEOMA Business School, Mont-Saint-Aignan, France

Abstract

Purpose – The authors aim to present partial least squares (PLS) as an evolving approach to structural equation modeling (SEM), highlight its advantages and limitations and provide an overview of recent research on the method across various fields.

Design/methodology/approach – In this review article, the authors merge literatures from the marketing, management, and management information systems fields to present the state-of-the art of PLS-SEM research. Furthermore, the authors meta-analyze recent review studies to shed light on popular reasons for PLS-SEM usage.

Findings – PLS-SEM has experienced increasing dissemination in a variety of fields in recent years with nonnormal data, small sample sizes and the use of formative indicators being the most prominent reasons for its application. Recent methodological research has extended PLS-SEM's methodological toolbox to accommodate more complex model structures or handle data inadequacies such as heterogeneity.

Research limitations/implications – While research on the PLS-SEM method has gained momentum during the last decade, there are ample research opportunities on subjects such as mediation or multigroup analysis, which warrant further attention.

Originality/value – This article provides an introduction to PLS-SEM for researchers that have not yet been exposed to the method. The article is the first to meta-analyze reasons for PLS-SEM usage across the marketing, management, and management information systems fields. The cross-disciplinary review of recent research on the PLS-SEM method also makes this article useful for researchers interested in advanced concepts.

Keywords Structural equation modeling, Partial least squares, PLS-SEM

Paper type General review

Introduction

The popularity of structural equation modeling (SEM) has grown out of the need to test complete theories and concepts (Rigdon, 1998). Much of SEM's success can be attributed to the method's ability to evaluate the measurement of latent variables, while also testing relationships between latent variables (Babin *et al.*, 2008). Although the initial

European Business Review
Vol. 26 No. 2, 2014
pp. 106-121
© Emerald Group Publishing Limited
0955-534X
DOI 10.1108/EBR-10-2013-0128

application of this method embraced a covariance-based approach (CB-SEM), researchers also have the option of choosing the variance-based partial least squares technique (PLS-SEM).

Originally developed by Wold (1974, 1980, 1982), PLS is an SEM technique based on an iterative approach that maximizes the explained variance of endogenous constructs (Fornell and Bookstein, 1982). Unlike CB-SEM, which aims to confirm theories by determining how well a model can estimate a covariance matrix for the sample data, PLS-SEM operates much like a multiple regression analysis (Hair *et al.*, 2011). This characteristic makes PLS-SEM particularly valuable for exploratory research purposes:

> PLS is primarily intended for research contexts that are simultaneously data-rich and theory-skeletal. The model building is then an evolutionary process, a dialog between the investigator and the computer. In the process, the model extracts fresh knowledge from the data, thereby putting flesh on the theoretical bones. At each step PLS rests content with consistency of the unknowns (Lohmöller and Wold, 1980, p. 1).

While CB-SEM is the more popular method, PLS-SEM has recently received considerable attention in a variety of disciplines including marketing (Hair *et al.*, 2012b), strategic management (Hair *et al.*, 2012a), management information systems (Ringle *et al.*, 2012), operations management (Peng and Lai, 2012), and accounting (Lee *et al.*, 2011). Much of the increased usage of PLS-SEM can be credited to the method's ability to handle problematic modeling issues that routinely occur in the social sciences such as unusual data characteristics (e.g. nonnormal data) and highly complex models.

Given the popularity and expected continued growth of PLS-SEM, this paper aims to discuss the current state of PLS-SEM by first providing an overview of past studies that have summarized PLS-SEM usage. Next, we will explain the process and steps used to test a model using PLS-SEM. Finally, our paper concludes by exploring many of the advanced topics associated with the method.

Prior PLS-SEM review studies

The argument for PLS-SEM as a viable methodology is gaining acceptance throughout many business disciplines. Several scholars have published studies summarizing PLS-SEM usage within their respective fields. The studies summarize the application of PLS-SEM, including the year of publication, range of years covered by the review, number of articles analyzed, and the justifications given for using PLS-SEM. The articles also reported the top three reasons given for applying PLS-SEM, which included data distribution, sample size, and the use of formative indicators. Table I summarizes the information reported in these articles.

Overall the findings indicate a substantial increase in the use of PLS-SEM in recent years. Three of the studies explored the growth trend by conducting a time-series analysis using the number of PLS-SEM studies. Hair *et al.* (2012b) and Ringle *et al.* (2012) found that the use of PLS-SEM in the marketing and management information systems fields has accelerated over time. In the strategic management field, PLS-SEM usage has grown linearly as a function of time (Hair *et al.*, 2012a).

When to use PLS-SEM

PLS-SEM provides numerous advantages to researchers working with structural equation models. Given the popularity of CB-SEM, the use of PLS-SEM often requires

Business discipline	Authors	Time period	Number of studies	Top three reasons for PLS-SEM usage[a]
Marketing	Hair *et al.* (2012b)	1981-2010	204	Nonnormal data: 50 percent Small sample size: 46 percent Formative indicators: 33 percent
Strategic management	Hair *et al.* (2012a)	1981-2010	37	Nonnormal data: 59 percent Small sample size: 46 percent Formative indicators: 27 percent
Management information systems	Ringle *et al.* (2012)	1992-2011	65	Small sample size: 37 percent Nonnormal data: 34 percent Formative indicators: 31 percent
Productions and operations management	Peng and Lai (2012)	2000-2011	42	Small sample size: 33 percent Formative indicators: 19 percent Nonnormal data: 14 percent
Accounting	Lee *et al.* (2011)	2005-2011	20	Not analyzed

Notes: [a]Percent of studies providing the corresponding reason; not all articles provided justifications and some articles provided multiple reasons

Table I.
PLS-SEM review studies from business disciplines

additional discussion to explain the rationale behind the decision (Chin, 2010). As our meta-analysis of PLS-SEM review studies has shown, the most prominent justifications for using PLS-SEM are attributed to:

- nonnormal data;
- small sample sizes; and
- formatively measured constructs (Table I).

These concepts are discussed below.

(1) Nonnormal data

Data collected for social science research often fails to follow a multivariate normal distribution. When attempting to evaluate a path model using CB-SEM, nonnormal data can lead to underestimated standard errors and inflated goodness-of-fit measures (Lei and Lomax, 2005). Fortunately, PLS-SEM is less stringent when working with nonnormal data because the PLS algorithm transforms nonnormal data in accordance with the central limit theorem (Beebe *et al.*, 1998; Cassel *et al.*, 1999). However, the caveat to PLS-SEM providing the end-all solution to models using nonnormal data is twofold. First, researchers should be aware that highly skewed data can reduce the statistical power of the analysis. More precisely, the evaluation of the model parameters' significances relies on standard errors from bootstrapping, which might be inflated when data are highly skewed (Hair *et al.*, 2014). Second, because CB-SEM has a variety of alternative estimation procedures, it may be problematic to assume that PLS-SEM is the automatic choice when considering data distribution (Hair *et al.*, 2012b).

(2) Small sample size

Sample size can affect several aspects of SEM including parameter estimates, model fit, and statistical power (Shah and Goldstein, 2006). However, different from CB-SEM, PLS-SEM can be utilized with much smaller sample sizes, even when models are highly complex. In these situations, PLS-SEM generally achieves higher levels of statistical power and demonstrates much better convergence behavior than CB-SEM (Henseler, 2010; Reinartz *et al.*, 2009). A popular heuristic states that the minimum sample size for a PLS model should be equal to the larger of the following:

- ten times the largest number of formative indicators used to measure one construct; or
- ten times the largest number of inner model paths directed at a particular construct in the inner model (Barclay *et al.*, 1995).

However, researchers should approach this guideline with caution, as misunderstandings have caused skepticism about the general uses of PLS-SEM (Hair *et al.*, 2014). As with any other model-based data analysis technique, researchers must consider sample size as it relates to the model complexity and data characteristics (Hair *et al.*, 2011). For example, while the rule of thumb put forth by Barclay *et al.* (1995) provides a rough estimate of minimum sample size, it fails to take into account the effect size, reliability, number of indicators, or other factors that are known to affect power (Henseler *et al.*, 2009).

(3) Formative indicators

The central difference between reflective and formative constructs is that formative measures represent instances in which the indicators cause the construct (i.e. the arrows point from the indicators to the construct), whereas reflective indicators are caused by the construct (i.e. the arrows point from the construct to the indicators). While both, PLS-SEM and CB-SEM can estimate models using formative indicators, PLS-SEM has received considerable support as the recommended method (Hair *et al.*, 2014). Because analyzing formative indicators with CB-SEM often leads to identification problems (Jarvis *et al.*, 2003), it is not uncommon for researchers to believe that PLS-SEM is the superior option. However, formative indicators should be approached with caution when using PLS-SEM. Researchers should be aware that the evaluation of formatively measured constructs relies on a totally different set of criteria compared to their reflective counterparts. Prior PLS-SEM review studies (Hair *et al.*, 2012a, b) have criticized the careless handling of formative indicators and researchers should apply the most recent set of evaluation criteria when examining the validity of formatively measured constructs (Hair *et al.*, 2014).

How to use PLS-SEM

When applying PLS-SEM, researchers need to follow a multi-stage process which involves the specification of the inner and outer models, data collection and examination, the actual model estimation, and the evaluation of results. In the following, this review centers around the three most salient steps:

(1) model specification;

(2) outer model evaluation; and

(3) inner model evaluation.

Hair *et al.* (2014) provide an in-depth introduction into each of the stages of PLS-SEM use.

(1) Model specification

The model specification stage deals with the set-up of the inner and outer models. The inner model, or structural model, displays the relationships between the constructs being evaluated. The outer models, also known as the measurement models, are used to evaluate the relationships between the indicator variables and their corresponding construct.

The first step in using PLS-SEM involves creating a path model that connects variables and constructs based on theory and logic (Hair *et al.*, 2014). In creating the path model such as that shown in Figure 1, it is important to distinguish the location of the constructs as well as the relationships between them. Constructs are considered either exogenous or endogenous. Whereas exogenous constructs act as independent variables and do not have an arrow pointing at them (Y_1, Y_2, and Y_3 in Figure 1), endogenous constructs are explained by other constructs (Y_4 and Y_5 in Figure 1). While often considered as the dependent variable within the relationship, endogenous constructs can also act as independent variables when they are placed between two constructs (Y_4 in Figure 1). When setting up the model, researchers need to be aware that in its basic form, the PLS-SEM algorithm can only handle models that have no circular relationship between the constructs. This requirement would be violated if we reversed the relationship $Y_2 \rightarrow Y_5$ in Figure 1. In this situation, Y_2 would predict Y_4, Y_4 would predict Y_5, and Y_5 would predict Y_2 again, yielding a circular loop (i.e. $Y_2 \rightarrow Y_4 \rightarrow Y_5 \rightarrow Y_2$).

After the inner model is designed, the researcher must specify the outer models. This step requires the researcher to make several decisions such as whether to use a multi-item or single-item scale (Diamantopoulos *et al.*, 2012; Sarstedt and Wilczynski, 2009) or whether to specify the outer model in a reflective or formative manner (Diamantopoulos and Winklhofer, 2001; Gudergan *et al.*, 2008). The sound specification

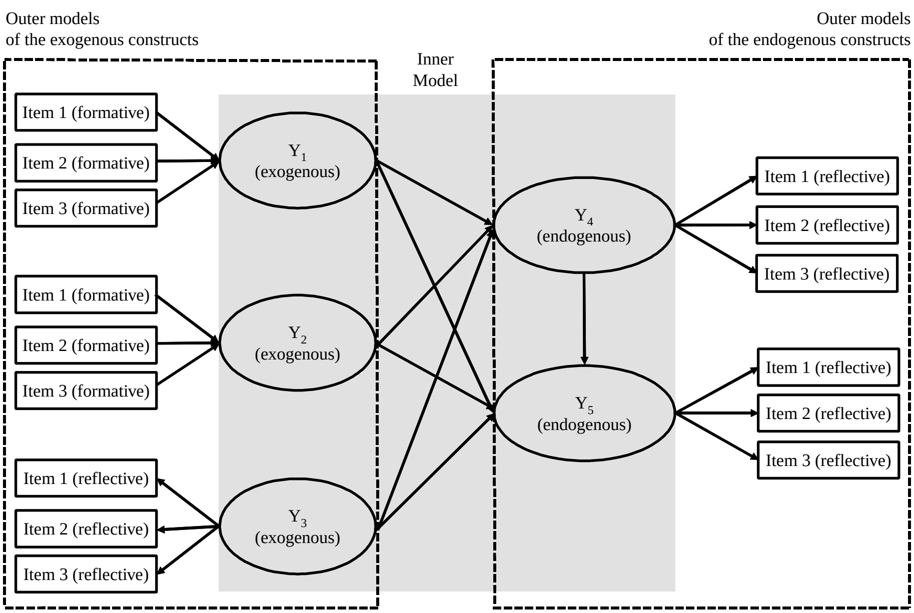

Figure 1.
A simple path model

of the outer models is crucial because the relationships hypothesized in the inner model are only as valid and reliable as the outer models. In Figure 1, Y_1 and Y_2 are measured formatively, while all other constructs have a reflective measurement specification. In this simple illustration, all constructs have an equal number of items. However, in applied research, the number of items per construct can be much higher, especially when formative measures are involved, as these – by definition – need to capture the entire domain of the construct (Diamantopoulos and Winklhofer, 2001; Diamantopoulos *et al.*, 2008).

(2) Outer model evaluation

Once the inner and outer models have been specified, the next step is running the PLS-SEM algorithm (for a description, see Henseler *et al.*, 2012) and, based on the results, evaluating the reliability and validity of the construct measures in the outer models. By starting with the assessment of the outer models, the researcher can trust that the constructs, which form the basis for the assessment of the inner model relationships, are accurately measured and represented. When evaluating the outer models, the researcher must distinguish between reflectively and formatively measured constructs (Ringle *et al.*, 2011; Sarstedt and Schloderer, 2010). The two approaches to measurement are based on different concepts and therefore require consideration of different evaluative measures.

(3) Reflective indicators

Reflective indicators constitute a representative set of all possible items within the conceptual domain of a construct (Diamantopoulos and Winklhofer, 2001). As a result, reflective items are interchangeable, highly correlated and capable of being omitted without changing the meaning of the construct. Reflective indicators are linked to a construct through loadings, which are the bivariate correlations between the indicator and the construct.

When assessing reflective outer models, researchers should verify both the reliability and validity. The first step is using composite reliability to evaluate the construct measures' internal consistency reliability. While traditionally assessed using Cronbach's α (Cronbach and Meehl, 1955), composite reliability provides a more appropriate measure of internal consistency reliability for at least two reasons. First, unlike Cronbach's α, composite reliability does not assume that all indicator loadings are equal in the population, which is in line with the working principle of the PLS-SEM algorithm that prioritizes the indicators based on their individual reliabilities during model estimation. Second, Cronbach's α is also sensitive to the number of items in the scale and generally tends to underestimate internal consistency reliability. By using composite reliability, PLS-SEM is able to accommodate different indicator reliabilities (i.e. differences in the indicator loadings), while also avoiding the underestimation associated with Cronbach's α.

The second step in evaluating reflective indicators is the assessment of validity. Validity is examined by noting a construct's convergent validity and discriminant validity. Support is provided for convergent validity when each item has outer loadings above 0.70 and when each construct's average variance extracted (AVE) is 0.50 or higher. The AVE is the grand mean value of the squared loadings of a set of indicators (Hair *et al.*, 2014) and is equivalent to the communality of a construct. Put succinctly, an AVE of 0.50 shows that the construct explains more than half of the variance of its indicators. Discriminant validity represents the extent to which the construct is empirically distinct from other constructs or, in other words, the construct measures what it is intended to measure. One method for assessing the existence of discriminant validity is the Fornell and Larcker (1981) criterion. This method states that the construct shares more variance with its indicators than with any other construct. To test this requirement, the AVE of each construct should be higher than the highest squared correlation with any other construct. The second option for verifying discriminant validity is examining the cross loadings of the indicators. This method, often considered more liberal (Henseler *et al.*, 2009), requires that the loadings of each indicator on its construct are higher than the cross loadings on other constructs.

Formative indicators. As indicated earlier, the principles underlying formative measurement are fundamentally different from the reflective type. Although PLS-SEM's ability to test models using formative indicators has attracted considerable attention across disciplines, many researchers applying the method disregard the specific steps that need to be followed when evaluating formative outer models (Hair *et al.*, 2012a, b).

First and foremost, the researcher needs to assess the content validity of the construct measures using expert assessment. Content validity evaluates the extent to which the indicators capture the major facets of the construct. Simply put, if an important item is omitted, the nature of the construct may be altered (Diamantopoulos *et al.*, 2008).

The empirical evaluation of formative outer models requires assessing convergent validity, or the extent to which a measure relates to other measures of the same phenomenon (Hair *et al.*, 2014). This assessment is done by means of a redundancy analysis in which each formatively measured construct is correlated with an alternative reflective or single-item measurement of the same construct. It is important to note that the redundancy analysis requires gathering data on the alternative measures at the same time as the original construct measures.

Next, the outer model indicators on each construct must be tested for collinearity. As with multiple regression (Mooi and Sarstedt, 2011), high collinearity between two or more formative indicators can seriously bias the results. More precisely, the weights linking the formative indicators with the constructs (which represent each indicators' contribution to the construct, controlling for the influence of all other indicators of the same construct) could be reversed and their significance underestimated as a result of increased standard errors.

Finally, researchers should evaluate the significance and relevance of each formative indicator. Since PLS-SEM does not assume a normal distribution, the researcher must apply the bootstrapping routine to determine the level of significance of each indicator weight. Bootstrapping is a resampling technique that draws a large number of subsamples from the original data (with replacement) and estimates models for each subsample. This way, the researcher obtains a large number (typically 5,000 or more) of model estimates, which can be used to compute a standard error of each model parameter. Drawing on the standard error, the significance of each parameter can be determined, using *t*-values. The assessment of the relevance of the indicators involves comparing the weights of the indicators to determine their relative contribution to forming the construct (Hair *et al.*, 2014). In specific instances (i.e. when the indicator weight is not significant), the researcher also needs to evaluate the bivariate correlation (loading) between the (nonsignificant) indicator and the construct in order to decide whether to exclude the indicator from the outer model (Hair *et al.*, 2014).

However, eliminating formative indicators from the model should generally be the exception, as formative measurement theory requires that the measures fully capture the entire domain of a construct. In short, omitting an indicator is equivalent to omitting a part of the construct.

Inner model evaluation. Once the reliability and validity of the outer models is established, several steps need to be taken to evaluate the hypothesized relationships within the inner model. This aspect of PLS-SEM is different from CB-SEM in that the model uses the sample data to obtain parameters that best predict the endogenous constructs, as opposed to estimating parameters that minimize the difference between the observed sample covariance matrix and the covariance matrix estimated by the model. As a result, PLS-SEM does not have a standard goodness-of-fit statistic and prior efforts to establishing a corresponding statistic have proven highly problematic (Henseler and Sarstedt, 2013). Instead, the assessment of the model's quality is based on its ability to predict the endogenous constructs. The following criteria facilitate this assessment: Coefficient of determination (R^2), cross-validated redundancy (Q^2), path coefficients, and the effect size (f^2). Prior to this assessment, the researcher needs to test the inner model for potential collinearity issues. As the inner model estimates result from sets of regression analyzes, their values and significances can be subject to biases if constructs are highly correlated (for a discussion and demonstration, see Hair *et al.*, 2014). While the Fornell-Larcker criterion usually discloses collinearity problems in the inner model earlier in the model evaluation process, this is not the case when formatively measured constructs are involved. The reason is that the AVE – which forms the basis for the Fornell-Larcker assessment – is not a meaningful measure for formative indicators. Therefore, collinearity assessment in the inner model is of pivotal importance when the model includes formatively measured constructs.

Coefficient of determination (R^2). The R^2 is a measure of the model's predictive accuracy. Another way to view R^2 is that it represents the exogenous variable's combined effect on the endogenous variable(s). This effect ranges from 0 to 1 with 1 representing complete predictive accuracy. Because R^2 is embraced by a variety of disciplines, scholars must rely on a "rough" rule of thumb regarding an acceptable R^2, with 0.75, 0.50, 0.25, respectively, describing substantial, moderate, or weak levels of predictive accuracy (Hair *et al.*, 2011; Henseler *et al.*, 2009). Though R^2 is a valuable tool in assessing the quality of a PLS model, too much reliance on R^2 can prove problematic. Specifically, if researchers attempt to compare models with different specifications of the same endogenous constructs, reliance only on R^2 may result in the researcher selecting a less efficient model. For example, the R^2 will increase even if a nonsignificant yet slightly correlated construct is added to the model. As a result, if the researcher's only goal is to improve the R^2, the researcher would benefit from adding additional exogenous constructs even if the relationships are not meaningful. Rather, the decision for a model should be based on the adjusted R^2, which penalizes increasing model complexity by reducing the (adjusted) R^2 when additional constructs are added to the model.

Cross-validated redundancy (Q^2). The Q^2 is a means for assessing the inner model's predictive relevance. The measure builds on a sample re-use technique, which omits a part of the data matrix, estimates the model parameters and predicts the omitted part using the estimates. The smaller the difference between predicted and original values the greater the Q^2 and thus the model's predictive accuracy. Specifically, a Q^2 value larger than zero for a particular endogenous construct indicates the path model's predictive relevance for this particular construct. It should, however, be noted that while comparing the Q^2 value to zero is indicative of whether an endogenous construct can be predicted, it does not say anything about the quality of the prediction (Rigdon, 2014; Sarstedt *et al.*, 2014).

Path coefficients. After running a PLS model, estimates are provided for the path coefficients, which represent the hypothesized relationships linking the constructs. Path coefficient values are standardized on a range from -1 to $+1$, with coefficients closer to $+1$ representing strong positive relationships and coefficients closer to -1 indicating strong negative relationships. Although values close to $+1$ or -1 are almost always statistically significant, a standard error must be obtained using bootstrapping to test for significance (Helm *et al.*, 2009). After verifying whether the relationships are significant, the researcher should consider the relevance of significant relationships. In short, are the sizes of the structural coefficients meaningful? As stated by Hair *et al.* (2014), many studies overlook this step and merely rely on the significance of effects. If this important step is omitted, researchers may focus on a relationship that, although significant, may be too small to merit managerial attention.

Effect size (f^2). The effect size for each path model can be determined by calculating Cohen's f^2. The f^2 is computed by noting the change in R^2 when a specific construct is eliminated from the model. To calculate the f^2, the researcher must estimate two PLS path models. The first path model should be the full model as specified by the hypotheses, yielding the R^2 of the full model (i.e. $R^2_{included}$). The second model should be identical except that a selected exogenous construct is eliminated from the model, yielding the R^2 of the reduced model (i.e. $R^2_{excluded}$). Based on the f^2 value, the effect size of the omitted construct for a particular endogenous construct can be determined such that 0.02, 0.15, and 0.35 represent small, medium, and large effects, respectively, (Cohen, 1988). That is, if an exogenous construct strongly contributes to explaining an endogenous construct, the difference between $R^2_{included}$ and $R^2_{excluded}$ will be high, leading to a high f^2 value. The effect size can be calculated using the formula below:

$$f^2 = \frac{R^2_{included} - R^2_{excluded}}{1 - R^2_{included}}$$

Advanced topics

The growing application of PLS-SEM is accompanied by broad range of methodological research that extends the method's toolbox. Several of these extensions deal with approaches to allow researchers specifying more complex model set-ups. In its simplest form, a PLS path model considers direct relationships between (sets of) constructs. However, more complex model set-ups are easily conceivable such as the estimation of moderating effects, mediating effects, or hierarchical component models.

Furthermore, methodological advances deal with the issue of heterogeneous data structures, which threaten the validity of the results. One stream of research in this field deals with multigroup analysis techniques to assess whether parameters (usually path coefficients) differ significantly across two or more groups of data. A second stream deals with the treatment of unobserved heterogeneity (i.e. heterogeneity that cannot be attributed to a single observable variable such as demographic variables) by means of latent class techniques. In the following, we give a brief description of recently discussed topics.

Moderation

Moderation occurs when the effect of an exogenous construct on an endogenous construct depends on the values of another variable, which influences (i.e. moderates) the relationship. For example, in their analysis of the relationship between dynamic capabilities and organizational performance, Wilden *et al.* (2013) demonstrate that the performance effect is contingent on the competitive intensity faced by firms as well as the firm's organizational structure. Research has brought forward several approaches for estimating moderating effects in PLS-SEM, which Henseler and Fassott (2010) and Rigdon *et al.* (2010) review. Henseler and Chin (2010) evaluate different approaches to moderation in PLS-SEM in terms of their applicability to reflective and formative measures, statistical power or predictive power.

Mediation

Mediation represents a situation in which a mediator variable to some extent absorbs the effect of an exogenous on an endogenous construct in the PLS path model. For example, in their study on the performance of consulting teams, Klarner *et al.* (2013) show that the relationship between consulting teams' competencies and their performance is sequentially mediated by client communication and team adaptability. As such, their analysis – opposed to a simple evaluation of direct effects – provides a more appropriate picture of management consulting team performance. Several authors have criticized the far-reaching neglect of explicitly examining mediating effects in PLS path models, which can easily lead to erroneous conclusions when interpreting model estimates (Hair *et al.*, 2013, 2012a, b). A potential reason for this neglect might be that there is still some ambiguity on how to evaluate mediating effects in PLS-SEM. Hair *et al.* (2014) provide an initial illustration on how to analyze mediating effects but more research is needed to provide guidance regarding the evaluation of more complex effects such as mediated moderation or moderated mediation.

Hierarchical component models

In some instances, the constructs researchers wish to examine are quite complex and can also be operationalized at higher levels of abstraction. For example, in their study on management consulting team performance, Klarner *et al.* (2013) conceptualize task competence as a two-dimensional construct with the dimensions relating to the team's generic and specific competencies. That is, instead of modeling these two competence types on a single construct layer, the authors summarize them as two lower-order components related to a single multidimensional higher-order construct. This modeling approach leads to more theoretical parsimony, reduces model complexity and can avert confounding effects in multidimensional model structures, such as multicollinearity (Kuppelwieser and Sarstedt, 2014; Ringle *et al.*, 2012). Theoretically, this process can be extended to any number of multiple layers, but researchers usually restrict their modeling approach to two layers. Wilson and Henseler (2007) as well as Becker *et al.* (2012) provide a review and evaluation of different approaches to modeling higher-order constructs using PLS-SEM. Hair *et al.* (2014) offer a tutorial on how to set-up and evaluate hierarchical component models.

Multigroup analysis

Multigroup analysis is a type of moderator analysis where the moderator variable is categorical (usually with two categories) and is assumed to potentially affect all relationships in the inner model. For example, using multigroup analysis, Elbanna *et al.* (2013) demonstrate that the role of intuition in strategic decision-making differs significantly in situations of low vs high environmental hostility. Research has brought forward several multigroup analysis approaches, which build on standard independent samples *t*-test (Keil *et al.*, 2000), permutation procedures (Chin, 2003; Chin and Dibbern, 2010), or bootstrap confidence intervals (Sarstedt *et al.*, 2011a, b). Sarstedt *et al.* (2011a, b) review the different approaches and propose an omnibus test of differences between more than two groups of data, which translates the standard *F*-test for use with PLS-SEM. As there are no concrete guidelines on when to use each approach, future research should empirically compare them by means of a large-scale simulation study.

Latent class techniques

When estimating PLS path models, situations arise in which differences related to unobserved heterogeneity prevent the model from being accurately estimated. Since researchers never know if unobserved heterogeneity is causing estimation problems, they need to apply complementary techniques for response-based segmentation (latent class techniques) that allow for identifying and treating unobserved heterogeneity. Recent research has brought forward a variety of latent class techniques which generalize, for example, finite mixture (Hahn *et al.*, 2002; Sarstedt *et al.*, 2011a, b), genetic algorithm (Ringle *et al.*, 2013a, b), or hill-climbing approaches (Becker *et al.*, 2013; Esposito *et al.*, 2008) to PLS-SEM. Sarstedt (2008) provides an early review of latent class techniques. In light of the considerable biases that result from neglecting unobserved heterogeneity (Ringle *et al.*, 2010; Rigdon *et al.*, 2011; Sarstedt and Ringle, 2010), recent research has called for the routine application of latent class techniques for evaluating the PLS path models (Becker *et al.*, 2013; Hair *et al.*, 2012b; Rigdon *et al.*, 2010).

Discussion

SEM has become the dominant analytical tool for testing cause-effect-relationships models with latent variables. When the goal of the analysis is to gain substantial knowledge about the drivers of, for example, customer satisfaction, brand image or corporate reputation, SEM is the technique of choice. For many researchers, SEM is equivalent to carrying out CB-SEM. While researchers have a basic understanding of CB-SEM, most of them have limited familiarity with the other useful approach – PLS-SEM.

Does lack of familiarity with PLS-SEM imply the loss of opportunities? It certainly does! Broadly speaking, the use of empirical methods in business applications has two objectives: prediction and explanation (Sarstedt *et al.*, 2014). Application of CB-SEM typically overlooks a key objective of empirical studies, which is prediction. The solution to this inherent weakness is the use of PLS-SEM, which has the overriding objective of predicting the dependent latent variables.

Compared to CB-SEM, PLS-SEM offers other significant advantages. Many empirical analysts pay lip service to distributional assumptions of the variables used in the analysis. In fact, most empirical business and social sciences data is characterized by nonnormal data. Consequently, CB-SEM applications that use the maximum likelihood algorithm – which most do – overlook the inherent violations of this technique’s requirements. Since PLS-SEM does not require these restrictive distributional assumptions, it is often a more viable approach than CB-SEM.

Another major advantage of PLS-SEM is that it permits the use of formative measures, which differ considerably from the reflective measures. Formatively measured constructs are particularly useful for studies that aim at explaining and predicting key constructs such as the sources of competitive advantage or corporate success (Albers, 2010). While CB-SEM can principally handle formative measures, their inclusion requires imposing considerable constraints on the model (Diamantopoulos and Riefler, 2011) or using a MIMIC approach, which is often questioned by SEM scholars.

PLS-SEM is subject to some constraints, however, related to the assessment of model fit (as commonly done in CB-SEM) and consistency of the parameter estimates. Recent research advances the basic PLS-SEM algorithm to improve its statistical properties, for example in terms of providing consistent parameter estimates. Dijkstra and Hensler's (2014) extension of PLS-SEM provides consistent parameter estimates and introduces the option of testing the path model's goodness-of-fit while maintaining the strengths of the method. Dijkstra and Schermelleh-Engel (2014) extend this approach to non-linear structural equation models. Further efforts at extending non-linear structural equation models have been made by Bentler and Huang (2014).

To summarize, depending on the specific empirical context and objectives of the study, PLS-SEM's distinctive methodological features make it a particularly valuable and potentially better-suited alternative to the more popular CB-SEM approaches in practical applications. Generally, however, neither method is superior to the other overall. Rather, the selection of the proper method depends on the objective of the study (Rigdon, 2012; Sarstedt *et al.*, 2014).

References

Albers, S. (2010), "PLS and success factor studies in marketing", in Esposito, V.V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares: Concepts, Methods and Applications in Marketing and Related Fields*, Springer, Berlin, pp. 409-425.

Babin, B.J., Hair, J.F. and Boles, J.S. (2008), "Publishing research in marketing journals using structural equations modeling", *Journal of Marketing Theory and Practice*, Vol. 16 No. 4, pp. 279-285.

Barclay, D.W., Higgins, C.A. and Thompson, R. (1995), "The partial least squares approach to causal modeling: personal computer adoption and use as illustration", *Technology Studies*, Vol. 2 No. 2, pp. 285-309.

Becker, J.-M., Klein, K. and Wetzels, M. (2012), "Hierarchical latent variable models in PLS-SEM: guidelines for using reflective-formative type models", *Long Range Planning*, Vol. 45 Nos 5/6, pp. 359-394.

Becker, J.M., Rai, A., Ringle, C.M. and Völckner, F. (2013), "Discovering unobserved heterogeneity in structural equation models to avert validity threats", *MIS Quarterly*, Vol. 37 No. 3, pp. 665-694.

Beebe, K.R., Pell, R.J. and Seasholtz, M.B. (1998), *Chemometrics: A Practical Guide*, Wiley, New York, NY.

Bentler, P.M. and Huang, W. (2014), "On components, latent variables, PLS and simple methods: reactions to Ridgon's rethinking of PLS", *Long Range Planning* (in press).

Cassel, C., Hackl, P. and Westlund, A.H. (1999), "Robustness of partial least-squares method for estimating latent variable quality structures", *Journal of Applied Statistics*, Vol. 26 No. 4, pp. 435-446.

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

Chin, W.W. (2003), "A permutation procedure for multi-group comparison of PLS models", in Vilares, M., Tenenhaus, M., Coelho, P., Esposito, V.V. and Morineau, A. (Eds), *PLS and Related Methods: Proceedings of the International Symposium PLS'03*, Decisia, Lisbon, pp. 33-43.

Chin, W.W. (2010), "How to write up and report PLS analyses", in Esposito, V.V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares: Concepts, Methods and Applications in Marketing and Related Fields*, Springer, Berlin, pp. 655-690.

Chin, W.W. and Dibbern, J. (2010), "An introduction to a permutation based procedure for multigroup PLS analysis: results of tests of differences on simulated data and a cross cultural analysis of the sourcing of information system services between Germany and the USA", in Esposito, V.V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares: Concepts, Methods and Applications*, Springer, Berlin, pp. 171-193.

Cronbach, L.J. and Meehl, P.E. (1955), "Construct validity in psychological tests", *Psychological Bulletin*, Vol. 52 No. 4, pp. 281-302.

Diamantopoulos, A. and Riefler, P. (2011), "Using formative measures in international marketing models: a cautionary tale using consumer animosity as an example", in Sarstedt, M., Schwaiger, M. and Taylor, C.R. (Eds), *Advances in International Marketing*, Emerald, Bradford, pp. 11-30.

Diamantopoulos, A. and Winklhofer, H.M. (2001), "Index construction with formative indicators: an alternative to scale development", *Journal of Marketing Research*, Vol. 38 No. 2, pp. 269-277.

Diamantopoulos, A., Riefler, P. and Roth, K.P. (2008), "Advancing formative measurement models", *Journal of Business Research*, Vol. 61 No. 12, pp. 1203-1218.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P. and Kaiser, S. (2012), "Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective", *Journal of the Academy of Marketing Science*, Vol. 40 No. 3, pp. 434-449.

Dijkstra, T.K. and Hensler, J. (2014), "Consistent and asymptotically normal PLS estimators for linear structural equations", working paper, Economics and Business, University of Groningen, Groningen, June 17.

Dijkstra, T.K. and Schermelleh-Engel, K. (2014), "Consistent partial least squares for nonlinear structural equation models", *Psychometrika* (in press).

Elbanna, S., Child, J. and Dayan, M. (2013), "A model of antecedents and consequences of intuition in strategic decision-making: evidence from Egypt", *Long Range Planning*, Vol. 46 Nos 1/2, pp. 149-176.

Esposito, V.V., Trinchera, L., Squillacciotti, S. and Tenenhaus, M. (2008), "REBUS-PLS: a response-based procedure for detecting unit segments in PLS path modeling", *Applied Stochastic Models in Business and Industry*, Vol. 24 No. 5, pp. 439-458.

Fornell, C.G. and Bookstein, F.L. (1982), "Two structural equation models: LISREL and PLS applied to consumer exit-voice theory", *Journal of Marketing Research*, Vol. 19 No. 4, pp. 440-452.

Fornell, C.G. and Larcker, D.F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, Vol. 18 No. 1, pp. 39-50.

Gudergan, S.P., Ringle, C.M., Wende, S. and Will, A. (2008), "Confirmatory tetrad analysis in PLS path modeling", *Journal of Business Research*, Vol. 61 No. 12, pp. 1238-1249.

Hahn, C., Johnson, M.D., Herrmann, A. and Huber, F. (2002), “Capturing customer heterogeneity using a finite mixture PLS approach”, *Schmalenbach Business Review*, Vol. 54 No. 3, pp. 243-269.

Hair, J.F., Ringle, C.M. and Sarstedt, M. (2011), “PLS-SEM: indeed a silver bullet”, *Journal of Marketing Theory and Practice*, Vol. 19 No. 2, pp. 139-151.

Hair, J.F., Ringle, C.M. and Sarstedt, M. (2013), “Partial least squares structural equation modeling: rigorous applications, better results and higher acceptance”, *Long Range Planning*, Vol. 46 Nos 1/2, pp. 1-12.

Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2014), *A Primer on Partial Least Squares Structural Equation Modeling*, Sage, Thousand Oaks, CA.

Hair, J.F., Sarstedt, M., Pieper, T.M. and Ringle, C.M. (2012a), “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*, Vol. 5 Nos 5/6, pp. 320-340.

Hair, J.F., Sarstedt, M., Ringle, C.M. and Mena, J.A. (2012b), “An assessment of the use of partial least squares structural equation modeling in marketing research”, *Journal of the Academy of Marketing Science*, Vol. 40 No. 3, pp. 414-433.

Helm, S., Eggert, A. and Garnefeld, I. (2009), “Modelling the impact of corporate reputation on customer satisfaction and loyalty using PLS”, in Esposito, V.V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares: Concepts, Methods, and Applications*, Springer, Berlin.

Henseler, J. (2010), “On the convergence of the partial least squares path modeling algorithm”, *Computational Statistics*, Vol. 25 No. 1, pp. 107-120.

Henseler, J. and 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*, Vol. 17 No. 1, pp. 82-109.

Henseler, J. and Fassott, G. (2010), “Testing moderating effects in PLS path models: an illustration of available procedures”, in Esposito, V.V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares: Concepts, Methods and Applications*, Springer, Berlin, pp. 713-735.

Henseler, J. and Sarstedt, M. (2013), “Goodness-of-fit indices for partial least squares path modeling”, *Computational Statistics*, Vol. 28 No. 2, pp. 565-580.

Henseler, J., Ringle, C.M. and Sarstedt, M. (2012), “Using partial least squares path modeling in advertising research: basic concepts and recent issues”, *Handbook of Research on International Advertising*, Edward Elgar Publishing, Cheltenham.

Henseler, J., Ringle, C.M. and Sinkovics, R.R. (2009), “The use of partial least squares path modeling in international marketing”, *Advances in International Marketing*, Vol. 20, pp. 277-320.

Jarvis, C.B., Mackenzie, S.B., Podsakoff, P.M., Mick, D.G. and Bearden, W.O. (2003), “A critical review of construct indicators and measurement model misspecification in marketing and consumer research”, *Journal of Consumer Research*, Vol. 30 No. 2, pp. 199-218.

Keil, M., Saarinen, T., Tan, B.C.Y., Tuunainen, V., Wassenaar, A. and Wei, K.K. (2000), “A cross-cultural study on escalation of commitment behavior in software projects”, *Management Information Systems Quarterly*, Vol. 24 No. 2, pp. 299-325.

Klarner, P., Sarstedt, M., Hoeck, M. and Ringle, C.M. (2013), “Disentangling the effects of team competences, team adaptability, and client communication on the performance of management consulting teams”, *Long Range Planning*, Vol. 46 No. 3, pp. 258-286.

Kuppelwieser, V. and Sarstedt, M. (2014), "Applying the future time perspective scale to advertising research", *International Journal of Advertising* (in press).

Lee, L., Petter, S., Fayard, D. and Robinson, S. (2011), "On the use of partial least squares path modeling in accounting research", *International Journal of Accounting Information Systems*, Vol. 12 No. 4, pp. 305-328.

Lei, M. and Lomax, R.G. (2005), "The effect of varying degrees of nonnormality in structural equation modeling", *Structural Equation Modeling*, Vol. 12 No. 1, pp. 1-27.

Lohmöller, J.B. and Wold, H. (1980), "Three-mode path models with latent variables and partial least squares (PLS) parameter estimation", paper presented at European Meeting of the Psychometric Society, Groningen, The Netherlands.

Mooi, E.A. and Sarstedt, M. (2011), *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics*, Springer, Berlin.

Peng, D.X. and Lai, F. (2012), "Using partial least squares in operations management research: a practical guideline and summary of past research", *Journal of Operations Management*, Vol. 30 No. 6, pp. 467-480.

Reinartz, W., Haenlein, M. and Henseler, J. (2009), "An empirical comparison of the efficacy of covariance-based and variance-based SEM", *International Journal of Research in Marketing*, Vol. 26 No. 4, pp. 332-344.

Rigdon, E.E. (1998), "Structural equations modeling", in Marcoulides, G.A. (Ed.), *Modern Methods for Business Research*, Lawrence-Erlbaum Associates, Mahwah, NJ, pp. 251-294.

Rigdon, E.E. (2012), "Rethinking partial least squares path modeling: in praise of simple methods", *Long Range Planning*, Vol. 45 Nos 5/6, pp. 341-358.

Rigdon, E.E. (2014), "Rethinking partial least squares path modeling: breaking chains and forging ahead", *Long Range Planning* (in press).

Rigdon, E.E., Ringle, C.M. and Sarstedt, M. (2010), "Structural modeling of heterogeneous data with partial least squares", in Malhotra, N.K. (Ed.), *Review of Marketing Research*, Emerald Group Publishing Limited, Bradford, pp. 255-296.

Rigdon, E.E., Ringle, C.M., Sarstedt, M. and Gudergan, S.P. (2011), "Assessing heterogeneity in customer satisfaction studies: across industry similarities and within industry differences", in Sarstedt, M., Schwaiger, M. and Taylor, C.R. (Eds), *Measurement and Research Methods in International Marketing*, Emerald Group Publishing Limited, Bradford, pp. 169-194 (*Advances in International Marketing*, Vol. 22).

Ringle, C.M., Sarstedt, M. and Mooi, E.A. (2010), "Response-based segmentation using finite mixture partial least squares", *Data Mining*, Springer, New York, NY, pp. 19-49.

Ringle, C.M., Sarstedt, M. and Schlittgen, R. (2013a), "Genetic algorithm segmentation in partial least squares structural equation modeling", *OR Spectrum*, pp. 1-26.

Ringle, C.M., Sarstedt, M. and Straub, D.W. (2012), "A critical look at the use of PLS-SEM in MIS Quarterly", *MIS Quarterly*, Vol. 36 No. 1, pp. 3-14.

Ringle, C.M., Sarstedt, M. and Zimmermann, L. (2011), "Customer satisfaction with commercial airlines: the role of perceived safety and purpose of travel", *The Journal of Marketing Theory and Practice*, Vol. 19 No. 4, pp. 459-472.

Ringle, C.M., Sarstedt, M., Schlittgen, R. and Taylor, C.R. (2013b), "PLS path modeling and evolutionary segmentation", *Journal of Business Research*, Vol. 66 No. 9, pp. 1318-1324.

Sarstedt, M. (2008), "A review of recent approaches for capturing heterogeneity in partial least squares path modeling", *Journal of Modeling in Management*, Vol. 3 No. 2, pp. 140-161.

Sarstedt, M. and Ringle, C.M. (2010), "Treating unobserved heterogeneity in PLS path modeling: a comparison of FIMIX-PLS with different data analysis strategies", *Journal of Applied Statistics*, Vol. 37 No. 8, pp. 1299-1318.

Sarstedt, M. and Schloderer, M.P. (2010), "Developing a measurement approach for reputation of non-profit organizations", *International Journal Nonprofit Voluntary Sector Marketing*, Vol. 15 No. 3, pp. 276-299.

Sarstedt, M. and Wilczynski, P. (2009), "More for less? A comparison of single-item and multi-item measures", *Die Betriebswirtschaft*, Vol. 69 No. 2, pp. 211-227.

Sarstedt, M., Henseler, J. and Ringle, C.M. (2011a), "Multigroup analysis in partial least squares (PLS) path modeling: alternative methods and empirical results", *Advances in International Marketing*, Vol. 22, pp. 195-218.

Sarstedt, M., Becker, J.M., Ringle, C. and Schwaiger, M. (2011b), "Uncovering and treating unobserved heterogeneity with FIMIX-PLS: which model selection criterion provides an appropriate number of segments?", *Schmalenbach Business Review*, Vol. 63 No. 1, pp. 34-62.

Sarstedt, M., Ringle, C.M., Henseler, J. and Hair, J.F. (2014), "On the emancipation of PLS-SEM", *Long Range Planning* (in press).

Shah, R. and Goldstein, S.M. (2006), "Use of structural equation modeling in operations management research: looking back and forward", *Journal of Operations Management*, Vol. 24 No. 2, pp. 148-169.

Wilden, R., Gudergan, S.P., Nielsen, B.B. and Lings, I. (2013), "Dynamic capabilities and performance: strategy, structure and environment", *Long Range Planning*, Vol. 46 Nos 1/2, pp. 72-96.

Wilson, B. and Henseler, J. (2007), "Modeling reflective higher-order constructs using three approaches with PLS path modeling: a Monte Carlo comparison", *Australian and New Zealand Marketing Academy Conference, Otago, Australia*, pp. 791-800.

Wold, H. (1974), "Causal flows with latent variables: partings of ways in the light of NIPALS modelling", *European Economic Review*, Vol. 5 No. 1, pp. 67-86.

Wold, H. (1980), "Model construction and evaluation when theoretical knowledge is scarce: theory and application of PLS", in Kmenta, J. and Ramsey, J.B. (Eds), *Evaluation of Econometric Models*, Academic Press, New York, NY.

Wold, H. (1982), "Soft modeling: the basic design and some extensions", in Jöreskog, K.G. and Wold, H. (Eds), *Systems Under Indirect Observations: Part II*, North-Holland, Amsterdam.

Corresponding author
Marko Sarstedt can be contacted at: marko.sarstedt@ovgu.de

To purchase reprints of this article please e-mail: **reprints@emeraldinsight.com**
Or visit our web site for further details: **www.emeraldinsight.com/reprints**