Intern. J. of Research in Marketing 26 (2009) 332–344

Contents lists available at ScienceDirect

Intern. J. of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

An empirical comparison of the efficacy of covariance-based and variance-based SEM

Werner Reinartz [a,b,*,1], Michael Haenlein [c,1,2], Jörg Henseler [d,e,1,3]

[a] University of Cologne, Germany
[b] INSEAD, Fontainebleau, France
[c] ESCP Europe, Paris, France
[d] Nijmegen School of Management, Radboud University, Nijmegen, The Netherlands
[e] ISEGI, New University of Lisbon, Portugal

ARTICLE INFO

Article history:
First received in 1, May 2007
and was under review for 7 months

Area editor: Bruce G.S. Hardie

Keywords:
Structural equation modeling
Maximum-likelihood
LISREL
PLS
Monte-Carlo simulation

ABSTRACT

Variance-based SEM, also known under the term partial least squares (PLS) analysis, is an approach that has gained increasing interest among marketing researchers in recent years. During the last 25 years, more than 30 articles have been published in leading marketing journals that have applied this approach instead of the more traditional alternative of covariance-based SEM (CBSEM). However, although an analysis of these previous publications shows that there seems to be at least an implicit agreement about the factors that should drive the choice between PLS analysis and CBSEM, no research has until now empirically compared the performance of these approaches given a set of different conditions. Our study addresses this open question by conducting a large-scale Monte-Carlo simulation. We show that justifying the choice of PLS due to a lack of assumptions regarding indicator distribution and measurement scale is often inappropriate, as CBSEM proves extremely robust with respect to violations of its underlying distributional assumptions. Additionally, CBSEM clearly outperforms PLS in terms of parameter consistency and is preferable in terms of parameter accuracy as long as the sample size exceeds a certain threshold (250 observations). Nevertheless, PLS analysis should be preferred when the emphasis is on prediction and theory development, as the statistical power of PLS is always larger than or equal to that of CBSEM; already, 100 observations can be sufficient to achieve acceptable levels of statistical power given a certain quality of the measurement model.

1. Introduction

Since Jöreskog's (1967) seminal work on maximum likelihood factor analysis and its later extensions to the estimation of structural equation systems (Jöreskog, 1973), structural equation modeling (SEM) has become one of the most important methods of empirical research, which has been applied in a multitude of areas including psychology (MacCallum & Austin, 2000), management research (Williams, Edwards, & Vandenberg, 2003), and marketing (Baumgartner & Homburg, 1996). For many researchers, applying SEM is equivalent to carrying out a maximum-likelihood, covariance-based analysis using, for example, the LISREL software (Jöreskog & Sörbom, 1982). Such covariance-based SEM (CBSEM) focuses on estimating a set of model parameters so that the theoretical covariance matrix implied by the system of structural equations is as close as possible to the empirical covariance matrix observed within the estimation sample. When carried out using maximum likelihood (ML) or generalized least squares (GLS), this estimation requires a set of assumptions to be fulfilled, such as the normal distribution of observed indicators and sufficient sample size. If these assumptions are violated, nontraditional alternatives to SEM, such as partial least squares (PLS, see, e.g., Rigdon, 2005; Wold, 1975), appear to be preferable options for researchers. Unlike CBSEM, a PLS analysis does not work with latent variables but rather with block variables, and estimates model parameters to maximize the variance explained for all endogenous constructs in the model through a series of ordinary least squares (OLS) regressions. It does not require any distributional assumptions to be fulfilled but results in inconsistent parameter estimates if the number of indicators per construct and the sample size are not infinitely large (Wold, 1975).

According to Fornell and Bookstein (1982), the different objectives of CBSEM and PLS may result in different parameter estimates for the same structural model in any given situation, which makes the choice between these two approaches "neither arbitrary nor straightforward." Previous research highlights three differences between CBSEM and PLS that can be used to guide this choice. First, parameter estimation in PLS is essentially carried out by a sequence of OLS regressions, which implies that no assumptions regarding the distribution or measurement scale of

* Corresponding author. Albertus Magnus Platz 1, D-50923 Cologne, Germany. Tel.: +49 221 470 5751; fax: +49 221 470 5191.
E-mail addresses: werner.reinartz@uni-koeln.de (W. Reinartz), haenlein@escpeurope.eu (M. Haenlein), j.henseler@fm.ru.nl (J. Henseler).
[1] All authors contributed equally to the research.
[2] Contact address: 79 Avenue de la République, F-75011 Paris, France. Tel.: +33 1 4923 2602; fax: +33 1 4923 2248.
[3] Contact address: Institute for Management Research, Thomas van Aquinostraat, 6525 GD Nijmegen, The Netherlands. Tel.: +31 24 361 1854; fax: +31 24 361 1933.

0167-8116/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijresmar.2009.08.001

Table 1
Articles published in the past 25 years using PLS and reasons provided for methodological choice.

Article	No assumptions about indicator distribution/ measurement scale	Suitability for small sample size	Focus on prediction and theory development	Suitability for unlimited number of formative indicators	Lack of improper solutions/factor indeterminacy
Fornell and Robinson (1983)	Yes	Yes			
Fornell, Robinson, and Wernerfelt (1985)	Yes				
Mayo and Qualls (1987)	Yes		Yes		
Qualls (1988)	Yes	Yes	Yes		
Zinkhan and Fornell (1989)	Yes		Yes	Yes	
Fornell, Lorange, and Roos (1990)	Yes		Yes		
Barclay (1991)			Yes		
Alpert, Kamins, and Graham (1992)				Yes	
Fornell (1992)	Yes		Yes		
Graham, Mintu, and Rodgers (1994)	Yes	Yes		Yes	
Green, Barclay, and Ryans (1995)	Yes	Yes	Yes		
Fornell et al. (1996)	Yes		Yes		
Smith and Barclay (1997)		Yes	Yes		
Dawes, Lee, and Dowling (1998)	Yes	Yes			
Sirohi, McLaughlin, and Wittink (1998)	Yes			Yes	Yes
Ahuja, Galletta, and Carley (2003)	Yes	Yes	Yes		
Arnett, Leverie, and Meiers (2003)	Yes	Yes		Yes	Yes
Vanhamme and Snelders (2003)	Yes	Yes			
White, Varadarajan, and Dacin (2003)	Yes		Yes	Yes	
Anderson, Fornell, and Mazvancheryl (2004)	Yes		Yes		
Cotte and Wood (2004)		Yes			
Dellande, Gilly, and Graham (2004)				Yes	
Gray and Meister (2004)			Yes		
Reinartz, Krafft, and Hoyer (2004)				Yes	Yes
Grégoire and Fisher (2005)			Yes		
Hennig-Thurau et al. (2006)	Yes				
Ulaga and Eggert (2006)			Yes	Yes	
Venkatesh and Agarwal (2006)	Yes			Yes	
Hennig-Thurau, Henning, and Sattler (2007)				Yes	
Mitchell and Nault (2007)		Yes			
McFarland, Bloodgood, and Payan (2008)				Yes	

observed indicators are required. In contrast, ML- or GLS-based CBSEM require normally distributed and interval-scaled variables (e.g., Dijkstra, 1983; Fornell & Bookstein, 1982). In addition, the use of OLS estimation also implies that PLS even works with small sample sizes, whereas ML- or GLS-based CBSEM usually require at least 200 observations to avoid non-convergence and improper solutions (Boomsma & Hoogland, 2001). Second, PLS focuses on maximizing the variance explained for all endogenous constructs in the model, whereas CBSEM determines the model parameters to reproduce an empirically observed covariance matrix. PLS is therefore better suited for situations in which the researcher wants to predict the latent variables in the model or identify relationships between them (e.g., in the early stages of theory development), while CBSEM should be the method of choice when the focus lies on confirming theoretically assumed relationships. Third, the PLS parameter estimation process continuously oscillates between estimating case values for the block variables and model parameters that depend on these case values. Block variables are hereby assumed to be a weighted average of all indicators that belong to the same construct. Because this basic approach is identical regardless of the type of operationalization used (reflective vs. formative), PLS can deal with an almost unlimited number of formative indicators. In contrast, CBSEM may result in implied covariances of zero among some indicators and/or equivalent models when formative measurements predominate (MacCallum & Browne, 1993). Furthermore, because all block variables are assumed to be linear combinations of their indicators, PLS does not suffer from improper solutions and factor indeterminacy, as sometimes occurs in the context of CBSEM (e.g., Bollen, 1987; Chen et al., 2001; Krijnen, Dijkstra, & Gill, 1998).

With respect to the use of CBSEM and PLS analysis in management research, the former approach easily dominates the latter. Yet, in recent years, interest in PLS has increased considerably, a phenomenon that we document in Table 1, in which we list all articles in eight leading marketing journals (*Advances in Consumer Research, International Journal of Research in Marketing, Journal of Consumer Research, Journal of Marketing, Journal of Marketing Research, Journal of Retailing, Management Science*, and *Marketing Science*) that have used PLS and been published in the past 25 years.[4] Two points emerge. First, it seems that PLS has prompted increasing interest among researchers in recent years. Of the 31 articles in Table 1, more than 50% (16) have appeared since 2003. Second, in each of these articles, one or several of the aforementioned differences between PLS and CBSEM are listed as reason(s) for the authors' methodological choices. Specifically, most articles mention the lack of assumptions regarding indicator distribution and measurement scales (19) for choosing PLS, followed by a focus on prediction and theory development (15) and the appropriateness of models with many formative indicators (12). The suitability of small sample sizes (11) and the nonexistence of improper solutions and factor indeterminacy (3) rank fourth and fifth, respectively. Thus, there seems to be at least an implicit agreement about the factors that should drive the choice between CBSEM and PLS. Yet, despite this agreement, there are to our knowledge no *quantitative* guidelines that help marketing researchers to make an unambiguous choice between these two approaches.

This lack of unambiguous quantitative guidelines is at least partly caused by the fact that previous simulation studies focusing on CBSEM and/or PLS frequently either include only one of these two approaches or only consider on a limited set of design factors. This can be seen in Table 2, where we provide an overview of the major simulation studies that have investigated the performance of CBSEM and/or PLS. Three results are particularly interesting. First, most studies, and especially the ones published by marketing scholars (e.g., Babakus, Ferguson, & Jöreskog, 1987; Gerbing & Anderson, 1985; Sharma,

[4] No articles using PLS appeared in Marketing Science, and only one appeared in the International Journal of Research in Marketing (Bagozzi, Yi, & Singh, 1991), the latter being methodological in nature. Table 1 includes only articles where a justification for the choice of PLS over CBSEM has been given; it excludes all articles that are purely methodological.

Table 2
Overview of Monte Carlo simulation studies focusing on CBSEM and/or PLS.

Article	Estimation technique		Design factor				Analysis focus		
	CBSEM	PLS	Sample size	Number of indicators	Indicator distribution	Indicator loadings	Parameter bias	Proper solutions	Statistical Power
Areskoug (1982)	Yes	Yes	Yes	Yes	No	No	Yes	No	No
Hui and Wold (1982)	No	Yes	Yes	Yes	No	No	Yes	NA	No
Gerbing and Anderson (1985)	Yes	No	Yes	Yes	No	Yes	Yes	No	No
Balderjahn (1986)	Yes	No	Yes	No	No	Yes	Yes	Yes	No
Babakus et al. (1987)	Yes	No	Yes	No	Yes	Yes	Yes	Yes	No
Sharma et al. (1989)	Yes	No	Yes	No	Yes	No	Yes	Yes	No
Marsh et al. (1998)	Yes	No	Yes	Yes	No	No	Yes	Yes	No
Cassel et al. (1999)	No	Yes	Yes	No	Yes	No	Yes	NA	No
Chin and Newsted (1999)	No	Yes	Yes	Yes	No	No	Yes	NA	No
Chen et al. (2001)	Yes	No	Yes	No	No	No	No	Yes	No
Goodhue et al. (2006)	Yes	Yes	Yes	No	No	No	Yes	No	Yes
Current study	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Does not include simulation studies that are focused on the relative performance of different fit indices (e.g., Bearden, Sharma, & Teel, 1982; Curran, West, & Finch, 1996; Hu & Bentler, 1998) or on the analysis of specific issues, such as the estimation of interaction effects (e.g., Chin et al., 2003), between-group differences (e.g., Qureshi & Compeau, 2009), measurement model misspecification (e.g., Jarvis, MacKenzie, & Podsakoff, 2003) and item parceling (e.g., Bandalos, 2002; Kim & Hagtvet, 2003; Nasser & Wisenbaker, 2003).

Durvasula, & Dillon, 1989), focus exclusively on the behavior of CBSEM estimates under various conditions. This is consistent with our previous observation that within the marketing literature, the use of CBSEM is far more frequent than the use of PLS, making a focus on CBSEM more appropriate, at least historically. Yet, while such studies provide interesting and relevant guidelines, they are only of limited usefulness when researchers want to compare the performance of CBSEM and PLS in different situations in order to choose the most appropriate approach for their research setting. Second, three studies investigate the performance of PLS (Cassel, Hackl, & Westlund, 1999; Chin & Newsted, 1999; Hui & Wold, 1982), but their focus is limited to a subset of two relevant design factors (sample size plus either number of indicators per construct or indicator distribution) and therefore does not allow one to balance competing objectives and requirements with regard to the choice between CBSEM and PLS. Third, only two studies (Areskoug, 1982; Goodhue, Lewis, & Thompson, 2006) include a simultaneous investigation of CBSEM and PLS. Yet, they equally only focus on a small subset of design factors and rely on relatively simple model structures that are not representative for the type of structural equation systems usually analyzed within the marketing discipline.

In summary, no previous research has empirically compared the performance of CBSEM and PLS along a large set of relevant design factors, which makes the relative performance of both approaches in many cases unclear. This lack of clear evidence makes it difficult for researchers to choose between CBSEM and PLS when some arguments favor one method whereas others suggest the other. Our study intends to provide a contribution in this area. Specifically, our objectives are twofold. First, we investigate the relative performance of ML-based CBSEM and PLS given a set of conditions, characterized by a full-factorial design of four factors that have previously been shown to have an impact on the performance of structural models[5]: number of indicators per construct, sample size, distribution, and indicator loadings. Second, we identify a set of rules that researchers can follow when choosing between ML-based CBSEM and PLS analysis. For the latter question, we focus on three different questions: First, does the approach converge to a proper solution? Second, what is the degree of parameter accuracy between the approaches and the relative importance of the different design factors in driving parameter accuracy? And finally, is the approach able to identify true relationships among the variables in the structural equation model—or, to put it differently, does it have low Type II error/high statistical power? We analyze these questions using a Monte Carlo simulation with 48,000 runs (240 scenarios with 200 replications each). For data generation, we use Mattson's method (Mattson, 1997; Reinartz, Echambadi, & Chin, 2002), which accounts substantially better for the non-normal distributions of latent variables than do traditional approaches recommended by, for example, Fleishman (1978) and Vale and Maurelli (1983).

Our results provide evidence that justifying the choice of PLS over ML-based CBSEM due to a lack of assumptions regarding indicator distribution is often inappropriate. Although PLS does not build on any distributional assumptions, ML-based CBSEM behaves robustly if those assumptions are violated, such that this difference seems to be irrelevant in many applications. Nevertheless, PLS is the preferable approach when researchers focus on prediction and theory development, as our simulations show that PLS requires only about half as many observations to reach a given level of statistical power as does ML-based CBSEM. Furthermore, choosing PLS over ML-based CBSEM when the sample size is limited appears sensible. The absolute relative error of parameters increases less quickly with a decrease in sample size for PLS than it does for ML-based CBSEM, and the negative effects of low sample sizes can easily be compensated for by increasing the number of indicators per construct or by using indicators with better psychometric properties (i.e., higher loadings). Finally, PLS should be the preferred approach when the researcher wants to avoid improper solutions, though we recognize that improper solutions are a relatively rare phenomenon in structural equation models with average complexity, affecting only a bit more than 1% of all our simulations.

2. Theoretical background

As stated in the previous section, the objective of our analysis is to compare the performance of ML-based CBSEM and PLS in a set of conditions, characterized by a full-factorial design of four factors (i.e., number of indicators per construct, sample size, distribution, and indicator loadings). Therefore, we first need to review prior studies that have investigated the behavior of either approach along these factors.

2.1. CBSEM

As noted previously, CBSEM and PLS analysis are essentially two different approaches to the same problem. Both start from the same set of theoretical and measurement equations but differ in how they approach the parameter estimation problem. Assume a structural equation model with a set of latent exogenous variables (ξ_i) measured by indicators x_i and associated measurement error δ_i, and a set of latent endogenous variables (η_j) measured by indicators y_j and

[5] Since our theoretical model includes only reflective indicators, our PLS analysis relies on the Mode-A estimation mode.

associated measurement error ε_j. If all latent variables in the model are assumed to be measured by reflective indicators, this structural equation model results in the following set of theoretical and measurement equations that describe the relationships of the structural and measurement model, respectively:

$$\eta = B\eta + \Gamma\xi + \zeta, \tag{1}$$

$$x = \Lambda_x\xi + \delta, \text{ and} \tag{2}$$

$$y = \Lambda_y\eta + \varepsilon. \tag{3}$$

Starting with this set of equations, covariance-based approaches such as LISREL estimate a vector of model parameters θ, so that the resulting covariance matrix predicted by the theoretical model $\Sigma = \Sigma(\theta)$ is as close as possible to the sample covariance matrix S. This estimation is usually conducted using maximum likelihood, with the likelihood function $F = \log|\Sigma| - \log|S| + tr\,(S\Sigma^{-1}) - k$, where $|A|$ denotes the determinant of A, $tr\,(A)$ is the sum of the diagonal elements of A, and k is the total number of manifest variables (indicators). As discussed, for example, by Long (1983), this likelihood function depends only on the vector of independent parameters θ, which consists of the free and constrained elements of Λ_x, Λ_y, B, and Γ, as well as Φ, Ψ, Θ_δ, and Θ_ε, which are the covariance matrices of ξ, ζ, δ, and ε, respectively. If determined using ML estimation, the estimated vector of the model parameters resulting from CBSEM is asymptotically efficient within the class of consistent estimators and can be considered optimal in that it is the most precise for large samples (Godambe, 1960).

2.1.1. Number of indicators per construct

As Long (1983) notes, CBSEM requires a minimum number of indicators to ensure model identification because the sample covariance matrix S must include at least as many non-redundant elements as the number of parameters to be estimated by the model. Baumgartner and Homburg (1996) go even further and state that every latent variable should be measured using at least three to four indicators to ensure meaningful results. Furthermore, the general consensus seems to be that an increase in the number of indicators is associated with positive effects. For example, Velicer and Fava (1987) show that more indicators decrease the risk of improper solutions, and Marsh, Hau, Balla, and Grayson (1998) suggest that more indicators per factor lead to more proper solutions, more accurate parameter estimates, and greater reliability. These findings, however, are true only up to a certain limit, because too many indicators lead to excessive power for the goodness-of-fit tests (MacCallum, Browne, & Sugawara, 1996), which in turn may significantly limit the usefulness of CBSEM (Haenlein & Kaplan, 2004).

2.1.2. Sample size

Sufficient sample size is necessary for both ML- and GLS-based CBSEM to ensure model identification because CBSEM requires the sample covariance matrix S to be positive-definite, which is only guaranteed when the sample size exceeds the number of indicators (Long, 1983). Additionally, a minimum sample size is required to generate results of sufficient accuracy due to the asymptotic property of ML estimation. Consistent with this thinking, Gerbing and Anderson (1985) show that the standard error of model estimates decreases with increasing sample size. As a rule of thumb, sample size should exceed 200 cases in most situations (Boomsma & Hoogland, 2001), and several strategies have been recommended if the available sample size falls below this threshold, including item parceling (e.g., Marsh et al., 1998; Nasser & Wisenbaker, 2003) or the use of alternative estimation techniques such as unweighted least squares (Balderjahn, 1986). Yet these strategies can be associated with significant risks (e.g., Kim & Hagtvet, 2003) or may not be applicable in all situations.

2.1.3. Distribution of indicators

As already highlighted by Jöreskog (1967), ML-based CBSEM requires that the observed variables have multinormal distribution. In reality, however, it is unlikely that empirical research will achieve this goal (Micceri, 1989). Therefore, several authors have investigated the behavior of ML-based CBSEM with non-normally distributed indicators, and it has been shown that in this case, standard errors in CBSEM tend to be inflated (Babakus, Ferguson, & Jöreskog, 1987). As with responses to the problem of limited sample size, item parceling (Bandalos, 2002) and alternative estimation techniques (Sharma, Durvasula, & Dillon, 1989) have been recommended as cures for non-normally distributed input data.

2.1.4. Indicator loadings

Badly operationalized constructs represent a problem for any type of empirical analysis, as they hinder the construction of theoretical knowledge. Therefore, a set of items used for construct operationalization should be both reliable and valid (Churchill, 1979). Construct reliability can be expressed as a function of indicator loadings, and higher average loadings coincide with higher reliability (Gerbing & Anderson, 1988). Because reliability pertains to the share of variance caused by (undesired) random error, high loadings are generally preferred over low ones. With respect to variability in the loadings of indicators that belong to the same construct, the case becomes less clear. Assuming constant average loadings (i.e., $\lambda_1 + \lambda_2 = 2\bar{\lambda}$ for two indicators), the average variance extracted (Fornell & Larcker, 1981), which is a measure of construct validity, will be minimal if the loadings are equal for all indicators of the same construct. Therefore, unequal loadings should be preferred over equal ones because they lead to higher validity. This statement also fits with the opinion that an overly high degree of item homogeneity should be avoided because it may indicate item redundancy (Boyle, 1991).

2.2. PLS

Developed by Herman Wold (who was Jöreskog's doctoral advisor), PLS analysis differs from CBSEM in that it works not with latent but with block variables, which are derived as weighted composites of their associated observed variables and are, hence, considered as observable themselves (Rigdon, 2005). The PLS estimation approach essentially consists of an iterative sequence of OLS regressions that starts with an outside approximation, during which the latent variables of the model are approximated by a linear combination of their indicators. For this process, a set of weights is determined in a manner similar to principal component analysis for reflective and regression analysis for formative indicators. In the next step, the inside approximation, alternative case values are determined as weighted means of those block variables that are adjacent within the structural model. Different ways to define adjacency associated with different weighting schemes are available (e.g., centroid, factor, path), but it has been shown that the choice among them has only a minor impact on the final result (Lohmöller, 1988). Using these new case values, the initial weights are modified, and the process of outside and inside approximation restarts and is repeated until the case values converge.

2.2.1. Number of indicators per construct and sample size

PLS analysis works not with latent variables but with block variables, which are defined as linear combinations of sets of indicators that usually involve measurement error. The block variables are therefore not free of error themselves. Hence, the scores determined for each block variable and each case, as well as the associated parameter estimates, must be considered inconsistent. They converge to their true population values only when both the number of indicators per construct and the sample size increase to infinity (Hui & Wold, 1982; Schneeweiss, 1993)—a property referred to in the literature as "consistency at large." In real-life situations, PLS therefore tends to underestimate the parameters of the structural model and overestimate those of the measurement model (Dijkstra, 1983). As with CBSEM, increasing sample size can be expected to decrease parameter variance. However, given a certain number of indicators, even an unlimited increase in sample size will not result in unbiased estimates, and given a certain sample size, any increase in the number of indicators per construct can only partially decrease the variation in parameter estimates. In turn, PLS analysis is particularly suited to cases in which CBSEM reaches its limits, such as when the number of indicators per latent variable becomes excessively large (as is the case, for example, in functional magnetic resonance imaging (fMRI) studies; see Haenlein & Kaplan, 2004) or when the sample size is small. For example, a Monte Carlo simulation carried out by Chin and Newsted (1999) shows that PLS can glean meaningful information from sample sizes as low as 20.

2.2.2. Distribution of indicators

As a limited-information approach, PLS only builds on mild statistical assumptions regarding the properties of the indicators and is therefore often described as a “soft modeling” technique to differentiate it from the “hard modeling” CBSEM approach. Specifically, PLS does not impose any requirements regarding the distribution or measurement scale of indicators used (Dijkstra, 1983). The only characteristic that must be fulfilled is that the systematic portion of all linear OLS regressions must be equal to the conditional expectation of the dependent variables (Wold, 1975). This condition, which is often referred to as a “predictor specification”, implies that the inner model is a causal chain system with uncorrelated residuals and that the residual that belongs to a given endogenous latent variable is uncorrelated with the corresponding predictor latent variables. The stability of PLS parameter estimates in the presence of non-normally distributed data has also been confirmed in a Monte Carlo simulation carried out by Cassel, Hackl, and Westlund (1999).

2.2.3. Indicator loadings

With respect to indicator loadings, the same points that we discussed with regard to CBSEM apply. Nevertheless, PLS can be expected to be more robust in the presence of inappropriately operationalized constructs, as the simultaneous estimation approach of CBSEM implies that one weak construct will likely influence all parameter estimates and latent variables estimates, while in PLS, such negative effects likely are limited to the construct itself and variables in its direct proximity.

3. Study design

Because ML-based CBSEM results in asymptotically efficient and optimal parameter estimates but relies on comparatively strong data assumptions, whereas PLS relies only on the mild condition of predictor specification but suffers from the problem of consistency at large, we argue that it is sensible to compare the relative efficacy of these two approaches within a set of conditions in which we expect one or the other approach to reach its limits. Such a comparison, which subsequently provides the basis for identifying a set of rules researchers can follow when choosing between ML-based CBSEM and PLS analysis, is the main objective of our manuscript.

3.1. Design factors

We define the number of indicators per construct on four levels (M = 2, 4, 6, 8), sample size on five levels (N = 100, 250, 500, 1000, 10,000), and the distribution of indicators on three levels (skewness/kurtosis = 0/0, 1/6, 2/12.8 for the independent latent variable). We specify the measurement model as depicted in Fig. 1. With respect to indicator loadings, we consider three different cases of equal standardized loadings (low: $\lambda_1 = \lambda_2 = .5$; medium: $\lambda_1 = \lambda_2 = .7$; high: $\lambda_1 = \lambda_2 = .9$), as well as one case of unequal standardized loadings ($\lambda_1 = .5$, $\lambda_2 = .9$).[6] These four design factors and their associated levels span a space of 240 scenarios ($4 \times 5 \times 3 \times 4$), for each of which we carried out 200 replications. These simulations build on the theoretical model visualized in Fig. 1, which mirrors the structure of a customer satisfaction index model (e.g., Fornell et al., 1996). We chose this type of model because it reflects the typical degree of complexity found for structural equation models within the marketing discipline. Additionally, there appears to be some debate about the preferable method of parameter estimation in this context. While the US customer satisfaction index literature has estimated the model using PLS (Fornell et al., 1996), some European modifications have applied CBSEM (e.g., Bruhn & Grund, 2000). Our population model consists of one exogenous (ξ) and five endogenous (η_1 to η_5) latent variables. The nine path coefficients γ_1 to γ_3 and β_1 to β_6 are assumed to have theoretical values of either 0.50, 0.30, or 0.15, to represent strong, medium, and weak population effect sizes, respectively (Cohen, 1988).

3.2. Data generation process

Generally, researchers can choose between two different methods of generating data for Monte Carlo simulations in the SEM context. The first method starts by calculating the covariance matrix of the observed indicators implied by the model and subsequently generates data from a multivariate distribution with the same covariance matrix. Fleishman (1978) and Vale and Maurelli (1983) have proposed approaches consistent with this method, which have been applied previously (e.g., Sharma, Durvasula, & Dillon, 1989). This technique is appropriate when the latent variables are assumed to be normally distributed, as the linearity inherent in the model implies that the indicators will also have a normal distribution. It becomes, however, less appropriate when this assumption is not met. In these situations, another technique proposed by Mattson (1997) and later applied by Reinartz, Echambadi, and Chin (2002) is preferable. This method generates data first for the latent variables within the structural model and subsequently for the observed indicators according to the relationships defined in the model. Mattson's approach has two advantages over the traditional technique described above. First, it is conceptually more satisfying because the data-generation process follows the theoretical model and the underlying relationships embedded in it. Second, it allows for complete control of the common and specific distributional characteristics of the latent and manifest variables. It takes account of the distributional characteristics of the latent independent variables and the latent dependent error terms, and ensures that the error terms influence only the distributional characteristics of the related indicators. Mattson's approach is currently the only one that enables researchers to control the skewness and kurtosis of both latent and observable variables simultaneously. To the best of our knowledge, this research represents the first time Mattson's approach has been applied in the context of a Monte Carlo simulation other than in the general analysis carried out by Reinartz, Echambadi, and Chin (2002).

3.3. Dependent variables

For each of the 48,000 replications, we calculate the relative error (RE) for the nine parameters of the structural model (γ_1 to γ_3 and β_1 to β_6), defined as:

$$RE(\hat{\theta}) = \frac{\hat{\theta} - \theta}{\theta} \tag{4}$$

[6] The error variance of each indicator can be determined as one minus the respective squared loading. The AVE is equal to the average of squared loadings and hence is 0.25 for low equal loadings, 0.49 for medium equal loadings, 0.81 for high equal loadings, and 0.53 for unequal loadings.

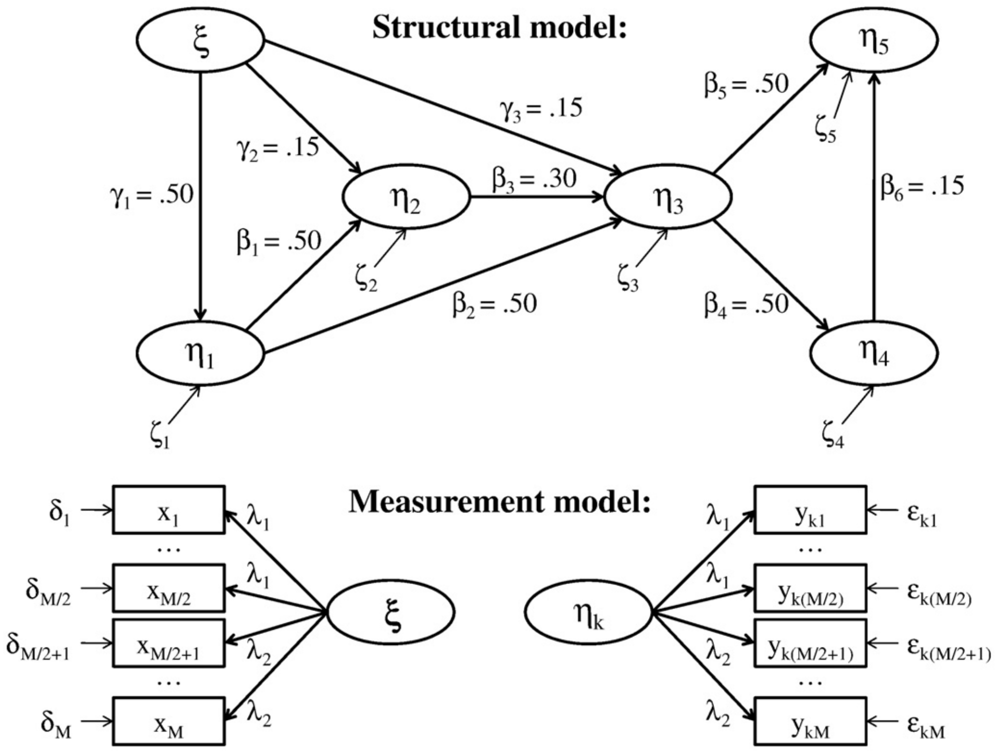

Fig. 1. Theoretical model.

where θ represents the theoretical value assumed for the respective parameter and $\hat{\theta}$ is equal to the estimated value of the same parameter in a given replication.[7] All simulations have been conducted within the *R* computing environment, Version 2.7.0 (R Development Core Team, 2008) using the SEM package (Fox, 2006) and a proprietary implementation of the PLS algorithm in the form as described by Tenenhaus et al. (2005).[8]

4. Analysis and results

The objectives of our Monte Carlo simulation are threefold. First, we are interested in the conditions that must be fulfilled so that ML-based CBSEM converges to a proper solution.[9] Second, we want to compare ML-based CBSEM and PLS with respect to their parameter accuracy and identify the relative importance of different design factors in driving parameter error. Third, we intend to identify the statistical power of ML-based CBSEM and PLS—that is, their ability to detect true relationships among latent variables.

4.1. Proper solutions in ML-based CBSEM

In line with a previous research, we define improper solutions as those estimates that would be impossible (or implausible) for the corresponding parameters (e.g., Bollen, 1987; Chen et al., 2001) and consider an ML-based CBSEM solution as proper if the iteration process converges to some solution and all variance estimates for that solution are positive. Table 3 shows the frequency of occurrence of proper solutions, convergence problems, and inadmissible solutions by design factor. Overall, 98.9% of our simulations resulted in proper solutions for ML-based CBSEM. For almost all scenarios, it is possible to achieve a proper solution with a probability greater than 0.90. Only in the worst case, i.e., 2 indicators per construct, low equal loadings, and 100 observations, is the probability of achieving a proper solution substantially lower, about 0.53. In order to investigate the extent to which the design factors included in our Monte Carlo simulation influence the probability of achieving a proper solution, we conducted a logistic regression analysis in which we modeled the properness of a solution as a function of indicator loadings, the logarithm of the sample size, the number of indicators, and their distribution.[10] Our model significantly explains the occurrence of proper solutions and results in acceptable pseudo-R^2 statistics (Nagelkerke's R^2: 0.5123; McFadden's R^2: 0.4973). Based on this analysis, the occurrence of proper solutions is significantly influenced by the sample size, indicator loadings, and number of indicators, but not by the distribution of indicators. Improper solutions are more likely in case of smaller sample size and low equal indicator loadings. Moreover, the influence of the number of indicators is nonlinear. While decreasing the number of indicators from four to two makes improper solutions significantly more likely, an increase in indicators from four to six or eight does not result in significantly more proper solutions.

Previous research (Boomsma & Hoogland, 2001) has investigated the minimum sample size necessary to achieve a proper solution as a function of the number of indicators per construct but has not considered indicator loadings. The well-known rule of thumb that ML-based CBSEM requires at least 200 observations to avoid problems of non-convergence and improper solutions emerged from this prior work. On the basis of our findings, we confirm that this rule is true on average but that wide variations depend on indicator loadings (see Table 4). Based on our analysis, the minimum sample size ranges from as low as 100 (medium or high equal loadings) to a maximum of 500 (low equal loadings and two indicators per construct). Note that if the number of indicators is low, it does make a difference whether there are medium equal or medium unequal loadings. In this case, researchers are well advised to take indicator loadings into account

[7] For some analyses on an aggregate level, we also used the absolute relative error (ARE), equal to the absolute value of RE, in order to avoid a canceling-out of positive and negative errors.

[8] The respective *R* codes used for data generation and the estimation of the PLS model are available from the third author upon request.

[9] PLS, as a limited information approach that works with block instead of latent variables, does not suffer from the problem of improper solutions.

[10] Details on this analysis are available upon request.

Table 3
Occurrence of proper solutions, convergence problems and inadmissible solutions by design factor.

Design factor	Factor level	Frequency of occurrence		
		Proper solution	Gradient not close to zero	At least one negative variance
Number of indicators per latent variable	2	11,635	236	129
	4	11,952	39	9
	6	11,939	51	10
	8	11,953	38	9
Indicator loadings	Low equal	11,525	339	136
	Medium equal	11,987	5	8
	High equal	11,987	9	4
	Unequal	11,980	11	9
Skewness and kurtosis	None	15,824	127	49
	Moderate	15,825	118	57
	High	15,830	119	51
Number of observations	100	9149	322	129
	250	9536	36	28
	500	9596	4	0
	1000	9600	0	0
	10,000	9598	2	0

when evaluating whether their sample size is sufficient for ML-based CBSEM.

4.2. Overall comparison of parameter accuracy in ML-based CBSEM and PLS

To explore the overall performance of ML-based CBSEM and PLS in terms of parameter accuracy, we compared the theoretical and estimated values for the nine parameters of the structural model across the 240 scenarios analyzed and for an "ideal" case. As can be seen in Table 5, on average across all 240 scenarios, parameter estimates do not differ significantly from their theoretical values for either ML-based CBSEM (p-values between 0.3963 and 0.5621) or PLS (p-values between 0.1906 and 0.3449). Nevertheless, ML-based CBSEM emerges as the more precise estimation method, as the mean parameter estimates are much closer to their theoretical values for CBSEM than for PLS (absolute difference 0.00–1.03% for CBSEM, 6.10–19.99% for PLS). Therefore, if consistency matters, ML-based CBSEM should be preferred over PLS.

To further clarify the relative performance of ML-based CBSEM and PLS in terms of parameter bias, we also compared the absolute relative error (ARE) for all parameters in an ideal scenario—i.e., the combination of different design factors for which the highest level of parameter accuracy can be expected from a theoretical perspective. Using our review of prior research, we define this scenario as the case with the maximum number of indicators per construct ($M=8$), maximum (asymptotic) sample size ($N=10{,}000$), normally distributed indicators (skewness = kurtosis = 0), and high equal loadings ($\lambda_1=\lambda_2=0.9$). In such conditions, estimates are virtually identical to their theoretical values for ML-based CBSEM, which suggests p-values between 0.745 and 0.957, an absolute difference between the theoretical and estimated parameter values of less than 0.04%, and an ARE between 0.011 and 0.053. In other words, under optimal conditions, parameter estimates obtained by ML-based CBSEM can be considered as accurate. For PLS, the same is not true. Although in general, the ARE is similar to the one for ML-based CBSEM (between 0.020 and 0.053), the ARE of strong effects is more than twice as large as in the case of ML-based CBSEM. Moreover, the difference between the theoretical and estimated parameter values is significant (p-values < .05 for eight of the nine effects) and substantial (between 0.58% and 3.23%). Thus, even in an ideal case, PLS path coefficients are biased and differ from the true parameters of the structural model. Our analysis indicates that, based on an overall comparison, ML-based CBSEM dominates PLS in terms of parameter accuracy.

Table 4
Minimum sample size necessary to achieve a proper solution with probability greater than 0.975.

Number of indicators per construct	Psychometric properties of indicators				Previous research (Boomsma & Hoogland, 2001)
	Equal loadings			Unequal loadings	
	Low	Medium	High		
2	500	100	100	250	NA
4	250	100	100	100	200+
6	250	100	100	100	50+
8	250	100	100	100	50+

4.3. Relative importance of different design factors in driving parameter accuracy

After having compared the overall performance of ML-based CBSEM and PLS in terms of parameter accuracy, we now analyze the relative importance of different design factors in driving parameter error (i.e., bias and variation). In order to avoid the problem of accumulated α errors that would result from a large number of individual comparisons, we compute the mean absolute relative error (MARE) defined as the mean ARE across all parameter estimates for each replication:

$$MARE = \frac{1}{t}\sum_{j=1}^{t}\left|\frac{\hat{\theta}_j-\theta_j}{\theta_j}\right|, \tag{5}$$

where t equals the number of parameters (here: 9), θ_j represents the theoretical value assumed for the respective parameter, and $\hat{\theta}_j$ represents the estimated value of the same parameter in any given replication.

In Table 6, we provide the ANCOVA results for a mixed-effects model explaining parameter accuracy operationalized as $\log_{10}$(MARE) as a function of the estimation method (ML-based CBSEM vs. PLS), the four design factors and their interactions.[11] This model shows that parameter accuracy is virtually unaffected by non-normality of the data. Neither the main effect nor the moderating effects of the distribution of indicators is significant. As the distribution of indicators has therefore neither a between-subjects nor a within-subjects effect, we can conclude that the accuracy of both ML-based CBSEM and PLS is independent of the distribution of indicators. All other design factors require a more differentiated assessment.

The between-subjects effects explain the variance in parameter error that both ML-based CBSEM and PLS share. Here, we note that the interaction of sample size × indicator loadings is the only relevant interaction effect (partial η^2: 0.0555). All other interaction effects are either not significant or not substantial (i.e., they have a partial η^2 clearly below 0.02). Regarding the main effects, sample size has the strongest impact on parameter accuracy, contributing by far the most to explaining the variance in MARE (partial η^2: 0.7980). The main effects of indicator loadings and the number of indicators are significant but not substantial.

The within-subjects effects describe the differences in accuracy between ML-based CBSEM and PLS. The most relevant highest-order within-subjects effects are the two interaction effects method × loadings × sample size and method × number of indicators × sample size. Both interaction effects subsume several highly significant lower-order effects. Although the main effect of the estimation method itself is strong and significant (partial η^2: 0.6511), suggesting that the two methods, ML-based CBSEM and PLS, differ strongly in parameter accuracy, these substantial interactions prevent a straight preference for either method. Fig. 2 displays the estimated marginal means of MARE along both interactions. The clear crossover interactions of

[11] We take $\log_{10}$(MARE) instead of MARE to avoid floor effects.

Table 5
Theoretical versus estimated parameter values (mean and standard deviation for parameter estimates and ARE).

Parameter		Theoretical value	ML-based CBSEM				PLS			
			Parameter estimate		ARE		Parameter estimate		ARE	
			μ	σ	μ	σ	μ	σ	μ	σ
Average across 240 scenarios	γ_1	0.5000	0.5005	0.0638	0.0795	0.0999	0.4079	0.0854	0.1950	0.1583
	γ_2	0.1500	0.1506	0.1015	0.3912	0.5522	0.1591	0.0550	0.2603	0.2650
	γ_3	0.1500	0.1485	0.1005	0.3443	0.5746	0.1594	0.0468	0.2252	0.2251
	β_1	0.5000	0.4990	0.0910	0.1057	0.1482	0.4012	0.0871	0.2093	0.1601
	β_2	0.5000	0.5027	0.1188	0.1193	0.2055	0.4083	0.0814	0.1945	0.1493
	β_3	0.3000	0.2991	0.1013	0.1844	0.2829	0.2750	0.0556	0.1468	0.1407
	β_4	0.5000	0.4986	0.0594	0.0769	0.0905	0.4062	0.0862	0.1983	0.1599
	β_5	0.5000	0.5000	0.0906	0.1052	0.1475	0.4001	0.0891	0.2130	0.1623
	β_6	0.1500	0.1488	0.0996	0.3825	0.5429	0.1603	0.0561	0.2613	0.2761
Ideal scenario	γ_1	0.5000	0.4997	0.0074	0.0120	0.0086	0.4855	0.0072	0.0292	0.0142
	γ_2	0.1500	0.1504	0.0096	0.0504	0.0396	0.1538	0.0091	0.0525	0.0399
	γ_3	0.1500	0.1498	0.0075	0.0417	0.0277	0.1534	0.0072	0.0431	0.0307
	β_1	0.5000	0.4998	0.0085	0.0138	0.0100	0.4839	0.0081	0.0323	0.0162
	β_2	0.5000	0.5004	0.0083	0.0128	0.0105	0.4851	0.0078	0.0301	0.0149
	β_3	0.3000	0.2997	0.0077	0.0203	0.0158	0.2982	0.0073	0.0201	0.0149
	β_4	0.5000	0.4993	0.0068	0.0105	0.0088	0.4851	0.0067	0.0299	0.0130
	β_5	0.5000	0.5004	0.0076	0.0118	0.0095	0.4844	0.0073	0.0314	0.0142
	β_6	0.1500	0.1495	0.0099	0.0531	0.0391	0.1530	0.0094	0.0525	0.0401

Notes: In the first case (average across 240 scenarios), mean and standard deviation refers to the parameter estimates and the absolute relative errors across the 240 scenarios; in the second case (ideal scenario), they refer to the parameter estimates and the absolute relative errors across the 200 runs within the ideal scenario (10,000 observations, 8 indicators, equally high loadings, 0 skewness and kurtosis). Means and standard deviations are calculated across all Monte Carlo runs for which maximum likelihood based structural equation modeling provided a proper solution.

MARE imply that the priority of methods alters with an increase in sample size. For small sample sizes, PLS tends to feature a higher level of accuracy than ML-based CBSEM, while the opposite is true for medium-sized and large samples. Besides these effects, we also identify moderate interaction effects of sample size, method and indicator loadings (Partial η^2: 0.2010) and as method and loadings (Partial η^2: 0.1528), indicating that the two methods are not equally sensitive to the psychometric properties of indicators. In sum, we can conclude that ML-based CBSEM clearly outperforms PLS in terms of consistency. While ML-based CBSEM is able to recover the population parameters on average, PLS path coefficients systematically deviate from the true parameter values. Moreover, ML-based CBSEM is

Table 6
ANCOVA explaining $\log_{10}$(MARE) by method (ML-based CBSEM/PLS) and design factor.

	Effect	F	df	Sig.	Partial η^2
Between-subjects effects	Intercept	5883.3361	1	0.0000	0.1105
	# of indicators	57.8819	3	0.0000	0.0037
	Distributions	1.6795	2	0.1865	0.0001
	Loadings	188.3461	3	0.0000	0.0118
	$\log_{10}(N)$	187,146.9764	1	0.0000	0.7980
	# of indicators × distributions	1.0163	6	0.4123	0.0001
	# of indicators × loadings	67.4141	9	0.0000	0.0126
	# of indicators × $\log_{10}(N)$	229.1549	3	0.0000	0.0143
	Distributions × loadings	0.0338	6	0.9998	0.0000
	Distributions × $\log_{10}(N)$	1.1896	2	0.3043	0.0001
	Loadings × $\log_{10}(N)$	927.3734	3	0.0000	0.0555
	# of indicators × distributions × loadings	0.0653	18	1.0000	0.0000
	# of indicators × distributions × $\log_{10}(N)$	0.9267	6	0.4743	0.0001
	# of indicators × loadings × $\log_{10}(N)$	2.3772	9	0.0000	0.0039
	Distributions × loadings × $\log_{10}(N)$	0.0269	6	0.9999	0.0000
	# of indicators × distributions × loadings × $\log_{10}(N)$	0.0484	18	1.0000	0.0000
	Error		47,383		
Within-Subjects Effects	Method	88,431.9109	1	0.0000	0.6511
	Method × # of indicators	1054.4940	3	0.0000	0.0626
	Method × distributions	0.4973	2	0.6082	0.0000
	Method × loadings	2849.0200	3	0.0000	0.1528
	Method × $\log_{10}(N)$	127,735.4391	1	0.0000	0.7294
	Method × # of indicators × distributions	0.9632	6	0.4484	0.0001
	Method × # of indicators × loadings	16.5847	9	0.0000	0.0031
	Method × # of indicators × $\log_{10}(N)$	1261.7954	3	0.0000	0.0740
	Method × distributions × loadings	0.1234	6	0.9936	0.0000
	Method × distributions × $\log_{10}(N)$	0.2625	2	0.7691	0.0000
	Method × loadings × $\log_{10}(N)$	3974.2111	3	0.0000	0.2010
	Method × # of indicators × distributions × loadings	0.1636	18	1.0000	0.0001
	Method × # of indicators × distributions × $\log_{10}(N)$	1.0405	6	0.3965	0.0001
	Method × # of indicators × loadings × $\log_{10}(N)$	33.9025	9	0.0000	0.0064
	Method × distributions × loadings × $\log_{10}(N)$	0.1246	6	0.9934	0.0000
	Method × # of indicators × distributions × loadings × $\log_{10}(N)$	0.1340	18	1.0000	0.0001
	Error(method)		47,383		

Notes: Includes only cases in which CBSEM resulted in a proper solution.

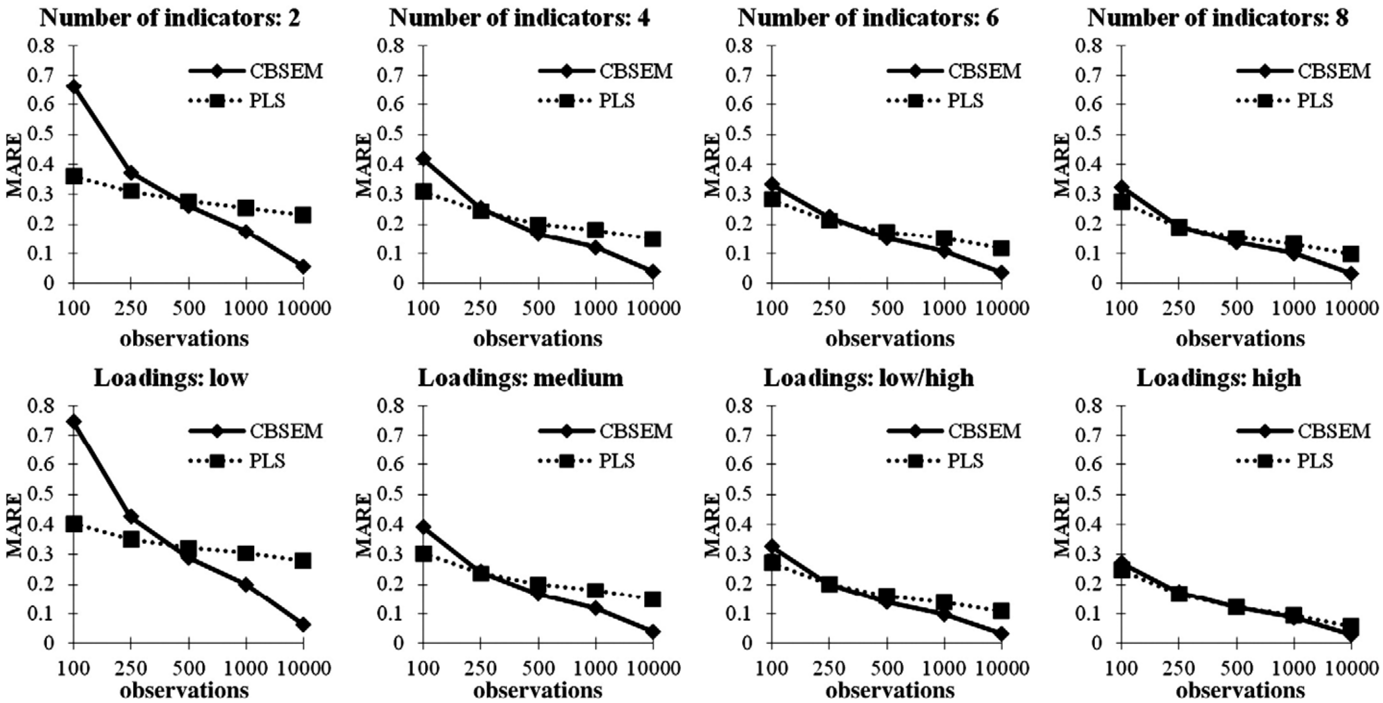

Fig. 2. Mean absolute relative error (MARE) of CBSEM and PLS for different numbers of indicators and loading patterns.

preferable in terms of parameter accuracy as long as the sample size exceeds a certain threshold. Below this threshold (about 250 observations in our case), PLS provides estimates with a lower MARE.

4.4. Statistical power of ML-based CBSEM and PLS

The statistical power of a significance test refers to the probability of rejecting a false H_0, given a certain population effect size, sample size, and significance criterion. If β is the probability of a Type II error (i.e., failure to reject a false H_0), power can be expressed as $1-\beta$ (e.g., Cohen, 1992). Sufficient statistical power is crucial, especially in the early stages of theory development, when the focus lies on identifying potentially significant relationships that could exist rather than confirming the significance of relationships whose existence can be assumed based on ample prior research. One of the main reasons provided for the methodological choice of PLS rather than ML-based CBSEM is its focus on prediction and theory development (see Table 1). Therefore, an implicit understanding seems to exist that statistical power can be expected to be higher in PLS than in ML-based CBSEM.

To verify this implicit belief, we determined the share of (proper) solutions for ML-based CBSEM and PLS in which the relationships between the latent variables specified in our structural model (i.e., the nine path coefficients β_1 to β_6 and γ_1 to γ_3) have not been rejected. While CBSEM instantly provides *t*-statistics for the path coefficients that can be used to perform such a significance test for the parameter estimates, we used bootstrapping (without sign correction) with 200 resamples to obtain standard errors for the PLS path coefficient estimates. We analyzed the statistical power of the two methods on an aggregated level and determined the frequency with which each method detects a significant ($p<0.05$) effect for all 240 scenarios. We hereby distinguish between three groups of effects, depending on the population effect size: strong effects (γ_1, β_1, β_2, β_4, β_5), medium effects (β_3), and weak effects (γ_2, γ_3, β_6).

Table 7 compares the statistical power of ML-based CBSEM and PLS for low, medium and high population effect sizes based on three design factors (i.e., sample size, indicator loadings and number of indicators).[12] As can be seen, the statistical power of PLS is always larger than or equal to that of ML-based CBSEM. To put it differently, the minimum sample size necessary to achieve a given level of statistical power in PLS is always less than or equal to the size required for ML-based CBSEM, and in many cases, ML-based CBSEM needs twice as much information as PLS to avoid Type II error. To the best of our knowledge, ours is the first quantitative study that confirms the widespread belief that PLS is preferable to ML-based CBSEM when the research focus lies in identifying relationships (i.e., prediction and theory development) instead of confirming them. Table 7 is one of the few published power tables for PLS, next to Chin and Newsted (1999) and Chin, Marcolin, and Newsted (2003). It is an essential tool for researchers who want to determine the statistical power of their estimation method given a particular population effect size, sample size and measurement model quality.

5. Discussion

In our Introduction, we recognized the increasing interest researchers in marketing have paid to PLS in recent years. Nevertheless, there appears to be an implicit agreement regarding the factors that should drive the methodological choice between the more traditional ML-based CBSEM and PLS—but no research has until now compared the performance of the two approaches in different scenarios. We therefore conducted a set of Monte Carlo simulations to address this issue. These simulations rely on 240 scenarios, defined according to a full-factorial design of four design factors (number of

[12] With respect to our analysis of power, one could argue that our comparison of ML-based CBSEM with PLS is inappropriate because the two methods use different ways of determining the standard error of the estimates: parametric assumptions in the case of ML-based CBSEM and bootstrap in the case of PLS. However, comparing the two approaches using the same method in determining the standard error of the estimates seems inappropriate. On the one hand, it has been shown that parametric assumptions always lead to higher statistical power than does bootstrapping in the context of ML-based CBSEM (Nevitt & Hancock, 2001). Additionally, applied research only rarely relies on bootstrapping in the context of ML-based CBSEM, except for the calculation of goodness-of-fit measures (Bollen & Stine, 1993). On the other hand, using parametric assumptions in the context of PLS leads to inflated Type I errors (Goodhue et al., 2006).

Table 7
Statistical power ($\alpha = 0.05$) of CBSEM and PLS.

Design factors			Low effect size ($\beta = .15$)		Medium effect size ($\beta = .30$)		High effect size ($\beta = .50$)	
Sample size	Loadings	Indicators	CBSEM	PLS	CBSEM	PLS	CBSEM	PLS
100	Low	2	0.01	0.24	0.01	0.47	0.16	0.59
	(.5/.5)	4	0.05	0.36	0.09	0.61	0.47	0.86
		6	0.07	0.43	0.31	0.79	0.69	0.94
		8	0.09	0.47	0.36	0.85	0.80	0.96
	Moderate	2	0.09	0.38	0.33	0.76	0.75	0.93
	Equal	4	0.18	0.44	0.51	0.85	0.96	0.99
	(.7/.7)	6	0.24	0.47	0.80	0.92	0.98	0.99
		8	0.27	0.45	0.81	0.95	0.99	0.99
	Moderate	2	0.09	0.41	0.39	0.84	0.83	0.98
	Unequal	4	0.29	0.44	0.80	0.94	1.00	1.00
	(.5/.9)l	6	0.33	0.48	0.90	0.94	1.00	0.99
		8	0.36	0.45	0.93	0.95	1.00	1.00
	High	2	0.28	0.40	0.84	0.93	0.99	1.00
	(.9/.9)	4	0.31	0.41	0.91	0.96	1.00	1.00
		6	0.40	0.46	0.94	0.95	1.00	0.99
		8	0.39	0.43	0.95	0.96	1.00	1.00
250	Low	2	0.05	0.54	0.16	0.90	0.54	0.90
	(.5/.5)	4	0.20	0.74	0.51	0.96	0.95	0.99
		6	0.23	0.76	0.80	1.00	0.99	1.00
		8	0.34	0.84	0.88	1.00	1.00	1.00
	Moderate	2	0.31	0.76	0.78	0.99	0.99	1.00
	Equal	4	0.45	0.81	0.95	1.00	1.00	1.00
	(.7/.7)	6	0.48	0.77	0.99	1.00	1.00	1.00
		8	0.60	0.83	1.00	1.00	1.00	1.00
	Moderate	2	0.36	0.80	0.88	1.00	1.00	1.00
	Unequal	4	0.62	0.82	1.00	1.00	1.00	1.00
	(.5/.9)l	6	0.63	0.77	1.00	1.00	1.00	1.00
		8	0.72	0.83	1.00	1.00	1.00	1.00
	High	2	0.62	0.80	1.00	1.00	1.00	1.00
	(.9/.9)	4	0.72	0.80	1.00	1.00	1.00	1.00
		6	0.68	0.75	1.00	1.00	1.00	1.00
		8	0.76	0.81	1.00	1.00	1.00	1.00
500	Low	2	0.13	0.78	0.31	0.97	0.86	1.00
	(.5/.5)	4	0.33	0.94	0.77	1.00	1.00	1.00
		6	0.48	0.97	0.96	1.00	1.00	1.00
		8	0.58	0.98	0.99	1.00	1.00	1.00
	Moderate	2	0.46	0.96	0.91	1.00	1.00	1.00
	Equal	4	0.73	0.98	1.00	1.00	1.00	1.00
	(.7/.7)	6	0.79	0.97	1.00	1.00	1.00	1.00
		8	0.87	0.98	1.00	1.00	1.00	1.00
	Moderate	2	0.64	0.97	1.00	1.00	1.00	1.00
	Unequal	4	0.90	0.98	1.00	1.00	1.00	1.00
	(.5/.9)l	6	0.91	0.97	1.00	1.00	1.00	1.00
		8	0.93	0.98	1.00	1.00	1.00	1.00
	High	2	0.85	0.97	1.00	1.00	1.00	1.00
	(.9/.9)	4	0.94	0.98	1.00	1.00	1.00	1.00
		6	0.93	0.95	1.00	1.00	1.00	1.00
		8	0.96	0.97	1.00	1.00	1.00	1.00
1000	Low	2	0.29	0.95	0.67	1.00	0.98	1.00
	(.5/.5)	4	0.59	1.00	0.99	1.00	1.00	1.00
		6	0.80	1.00	1.00	1.00	1.00	1.00
		8	0.87	1.00	1.00	1.00	1.00	1.00
	Moderate	2	0.79	1.00	1.00	1.00	1.00	1.00
	Equal	4	0.95	1.00	1.00	1.00	1.00	1.00
	(.7/.7)	6	0.98	1.00	1.00	1.00	1.00	1.00
		8	0.99	1.00	1.00	1.00	1.00	1.00
	Moderate	2	0.94	1.00	1.00	1.00	1.00	1.00
	Unequal	4	0.99	1.00	1.00	1.00	1.00	1.00
	(.5/.9)l	6	0.99	1.00	1.00	1.00	1.00	1.00
		8	1.00	1.00	1.00	1.00	1.00	1.00
	High	2	0.99	1.00	1.00	1.00	1.00	1.00
	(.9/.9)	4	1.00	1.00	1.00	1.00	1.00	1.00
		6	1.00	1.00	1.00	1.00	1.00	1.00
		8	1.00	1.00	1.00	1.00	1.00	1.00

Notes: Includes only cases in which CBSEM resulted in a proper solution. In case with 10,000 observations, a statistical power of one was obtained for all conditions.

indicators per construct, sample size, distribution, and indicator loadings) and Mattson's (1997) approach to data generation. Specifically, our analysis has attempted to answer three research questions. 1) Which conditions need to be fulfilled so that ML-based CBSEM converges to a proper solution? 2) What is the difference in the parameter bias between the two approaches and the relative importance of different design factors in driving parameter accuracy? 3) What is the ability of ML-based CBSEM versus PLS to detect true relationships among latent variables?

5.1. Theoretical implications

On the basis of our results, we can evaluate four of the five main reasons provided for a methodological choice between PLS and ML-based CBSEM, as cited in Table 1, and develop a set of recommendations for that choice. Because our population model only includes constructs measured using reflective indicators, we cannot make any statement with respect to PLS performance in cases when formative measures predominate. In addition, we do not discuss reasons that may favor a particular method other than those that appear in Table 1. Such reasons might include, for example, the availability of tests to judge overall model fit and the suitability of an approach to dealing with multi-level structures, growth modeling, mixtures and/or equality constraints. Nevertheless, our recommendations should be useful for practicing researchers, among whom there seems to be high heterogeneity in terms of the reasoning for choosing one method over another but no systematic quantitative and empirical assessment to help rationalize that choice.

5.1.1. When assumptions regarding indicator distribution are not met

As highlighted above, most authors cite a lack of assumptions regarding indicator distribution and measurement scale as their main reason for choosing PLS over ML-based CBSEM. Our results indicate that such a justification is often inappropriate, as ML-based CBSEM proves extremely robust with respect to violations of its underlying distributional assumptions. The distribution of indicators impacts neither the share of proper solutions for ML-based CBSEM nor parameter accuracy in any significant and substantial manner, even in extreme cases of skewness and kurtosis. Although PLS does not build on any distributional assumptions, ML-based CBSEM behaves so robustly in the case of their violation that justifying the choice of one approach over the other on the basis of this factor alone is not sufficient.

5.1.2. When the focus is on prediction and theory development

15 of the 30 articles listed in Table 1 justify the use of PLS based on a focus on prediction and theory development vs. empirical confirmation of theoretically indicated relationships. Our comparison of the statistical power of ML-based CBSEM and PLS clearly supports this statement. The statistical power of PLS is always larger than or equal to that of ML-based CBSEM, and in many cases, PLS requires only half as much information as ML-based CBSEM. With a reasonable measurement model (e.g., four indicators per construct with at least medium loadings), PLS can achieve a statistical power of 0.80 for medium population effect sizes with a sample size as small as 100 and for weak population effect sizes with about 250 observations. To achieve similar results, ML-based CBSEM requires 250 and 1000 observations, respectively. However, in these circumstances, PLS estimates must be expected to be inaccurate by roughly 25% (ARE for $M = 4$, $N = 100$; medium equal loadings are 0.3035 for small effects and 0.2412 for medium effects). Although this level of bias is sufficiently low to reject the null hypothesis that the parameter value is zero, it may cast doubt on the actual parameter estimate obtained, which should be interpreted with caution.

5.1.3. When sample size is small

The third most cited reason for using PLS is its suitability for small sample sizes. Our simulations show that PLS can be a very sensible methodological choice if sample size is restricted, since already 100 observations can be sufficient to achieve acceptable levels of statistical power, given a certain quality of the measurement model. Although parameter estimates may be inaccurate in this case, ARE depends much less on sample size within PLS than it does within ML-based CBSEM. Whereas sample size is by far the most important factor driving parameter accuracy in ML-based CBSEM, it plays a less important role in PLS. Additionally, low sample size in PLS can easily be compensated for by improving the number of indicators or by choosing indicators with higher loadings. It can be derived from Fig. 2 that PLS should be the method of choice for all situations in which the number of observations is lower than 250 (400 observations in the case of less reliable measurement models, i.e., low loadings and/or few indicators), while ML-based CBSEM should be chosen otherwise. In the case of 100 observations, and if constructs are measured with at least six indicators with at least medium loadings, the ARE falls between 0.2420 and 0.2747. In contrast, ML-based CBSEM shows an ARE between 0.2557 and 0.3178 in the same circumstances. This advantage of PLS is particularly relevant when researchers plan to use SEM in cases where the sample sizes required for ML-based CBSEM is not available. For example, Green, Barclay, and Ryans (1995) investigate the impact of entry strategy on long-term performance in the business word processor and graphics markets, where only 39 and 44 companies entered the market in the analysis period. However, in such situations, researchers must pay particular attention to including a sufficient number of indicators per construct. For example, of the eleven articles in Table 1 that cite suitability for small sample sizes as a reason for choosing PLS, eight are based on sample sizes less than or equal to 100. Of these eight, only one (Qualls, 1988) uses constructs operationalized with at least six indicators each. On the basis of our results, we encourage researchers to include the number of indicators per construct as a factor when choosing between PLS and ML-based CBSEM, especially in the presence of limited sample sizes.

5.1.4. To avoid improper solutions

Several authors (Arnett, Laverie, & Meiers, 2003; Reinartz, Krafft, & Hoyer, 2004; Sirohi, McLaughlin, & Wittink, 1998) have chosen PLS because it does not suffer from identification and convergence problems. While this reasoning is true theoretically, we observe that improper solutions are a relatively rare phenomenon that affects only 1.1% of our simulations. However, especially when indicator loadings are low, ML-based CBSEM can require significant sample sizes of more than 500 observations to avoid them.

5.2. Limitations and areas for further research

As with those of any Monte Carlo simulation, our findings are valid only within the boundaries of the scenarios we investigate, and they only apply to the theoretical model on which we base our simulations (Fig. 1). Furthermore, we assume that all indicators in our model are continuous, which rarely occurs in real life. Such increased precision of information regarding the latent constructs is likely to influence our results. In general, however, our approach is substantially more complex than those applied by other researchers in similar situations, specifically due to the use of Mattson's (1997) method of data generation, which gives us confidence in the external validity of our results. Regarding areas of further research, we believe that an extension of our study to misspecified models (see Hu & Bentler, 1998 for a similar analysis in the context of fit indices) and second-order factor specifications could be very interesting. Furthermore, questions surrounding PLS regarding prediction and theory development, as well as its suitability for an unlimited number of formative indicators, deserve deeper investigation. With respect to the first point, for example, it would be very interesting to analyze differences in factor scores derived using PLS and ML-based CBSEM in more detail. Theoretically, the focus on maximizing explained variance, which lies at the heart of PLS, should lead to better predictions than the estimation approach that underlies CBSEM. In turn, many authors tend to choose PLS over CBSEM, especially when factor scores are of particular interest, such as in the context of index construction (e.g., Arnett, Laverie, & Meiers, 2003; Fornell et al., 1996). Tenenhaus et al. (2005) suggest, however, that the differences in factor scores between ML-based CBSEM and PLS are less a question of the estimation procedure than one of the specific way in which factor scores are calculated for both approaches. In the specific example they analyze, factor scores that follow the logic of PLS for their calculation but use CBSEM estimates as input parameters lead to results that are highly correlated with traditional PLS factor scores. The question of whether their finding is idiosyncratic to the example they investigate or can be generalized to a broader setting seems highly relevant in this context.

With respect to the suitability of PLS for models with many formative indicators, MacCallum and Browne (1993) highlight several issues that may occur when formative indicators are predominant in ML-based CBSEM. Therefore, most recommendations involve combining reflective and formative indicators in the form of a MIMIC model (Jöreskog and Goldberger, 1975) to avoid such problems. To our knowledge however, no study has compared the relative advantages of this approach to the use of formative indicators only or the performance of PLS and ML-based CBSEM in both cases with Monte Carlo simulations. This lack may be partly caused by the fact that the simulation of formative indicators is a nontrivial issue. All approaches currently used to generate artificial data in the SEM context build on the assumption of reflective measurement. However, formative indicators can be expected to grow in importance because of their high degree of suitability for modeling managerial constructs (Jarvis, MacKenzie, & Podsakoff, 2003). Therefore, we recommend a focus on theoretical ways to simulate formative indicators in the context of SEM, probably building on Mattson's (1997) approach and investigating whether the use of a logical flow from latent constructs to indicators to generate artificial data might be extended to formative measurements.

References

Ahuja, M. K., Galletta, D. F., & Carley, K. M. (2003). Individual centrality and performance in virtual R&D groups: An empirical study. *Management Science*, *49*(1), 21–38.

Alpert, F. H., Kamins, M. A., & Graham, J. L. (1992). An examination of reseller buyer attitudes toward order of brand entry. *Journal of Marketing*, *56*(3), 25–37.

Anderson, E. W., Fornell, C., & Mazvancheryl, S. K. (2004). Customer satisfaction and shareholder value. *Journal of Marketing*, *68*(4), 172–185.

Areskoug, B. (1982). The first canonical correlation: Theoretical PLS analysis and simulation experiments. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction*. Amsterdam: North Holland.

Arnett, D. B., Laverie, D. A., & Meiers, A. (2003). Developing parsimonious retailer equity indexes using partial least squares analysis: A method and applications. *Journal of Retailing*, *79*(3), 161–170.

Babakus, E., Ferguson, C. E., & Jöreskog, K. G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research*, *24*(2), 222–228.

Bagozzi, R. P., Yi, Y., & Singh, S. (1991). On the use of structural equation models in experimental designs: Two extensions. *International Journal of Research in Marketing*, *8*(2), 125–140.

Balderjahn, I. (1986). The robustness of LISREL unweighted least squares estimation against small sample size in confirmatory factor analysis models. In W. Gaul & M. Schader (Eds.), *Classification as a tool of research*. Amsterdam: North-Holland publishing company.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling*, *9*(1), 78–102.

Barclay, D. W. (1991). Interdepartmental conflict in organizational buying: The impact of the organizational context. *Journal of Marketing Research*, *28*(2), 145–159.

Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: A review. *International Journal of Research in Marketing*, *13*(2), 139–161.

Bearden, W. O., Sharma, S., & Teel, J. E. (1982). Sample size effects on Chi square and other statistics used in evaluating causal models. *Journal of Marketing Research*, *19*(4), 425–430.

Bollen, K. A. (1987). Outliers and improper solutions: A confirmatory factor analysis example. *Sociological Methods and Research*, *15*(4), 375–384.

Bollen, K. A., & Stine, R. A. (1993). Bootstrapping goodness-of-fit measures in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models*. Newbury Park,CA: Sage.

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation models: Present and future*. Chicago: Scientific Software International.

Boyle, G. J. (1991). Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Personality and Individual Differences*, *12*(3), 291–294.

Bruhn, M., & Grund, M. A. (2000). Theory, development and implementation of national customer satisfaction indices: The Swiss index of customer satisfaction (SWICS). *Total Quality Management*, *11*(7), 1017–1028.

Cassel, C. M., Hackl, P., & Westlund, A. H. (1999). Robustness of partial least-squares method for estimating latent variable quality structures. *Journal of Applied Statistics*, *26*(4), 435–446.

Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., & Kirby, J. B. (2001). Improper solutions in structural equation models: Causes, consequences and strategies. *Sociological Methods and Research*, *29*(4), 468–508.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modelling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, *14*(2), 189–217.

Chin, W. W., & Newsted, P. R. (1999). Structural equation modelling analysis with small samples using partial least squares. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research*. Thousand Oaks: SAGE Publications.

Churchill, G. A., Jr. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, *16*(1), 64–73.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, New Jersey: Lawrence Erlbaum associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Cotte, J., & Wood, S. L. (2004). Families and innovative consumer behavior: A triadic analysis of sibling and parental influence. *Journal of Consumer Research*, *31*(1), 78–86.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29.

Dawes, P. L., Lee, D. Y., & Dowling, G. R. (1998). Information control and influence in emergent buying centers. *Journal of Marketing*, *62*(3), 55–68.

Dellande, S., Gilly, M. C., & Graham, J. L. (2004). Gaining compliance and losing weight: The role of the service provider in health care services. *Journal of Marketing*, *68*(3), 78–91.

Dijkstra, T. (1983). Some comments on maximum likelihood and partial least squares methods. *Journal of Econometrics*, *22*(1/2), 67–90.

Fleishman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika*, *43*(4), 521–532.

Fornell, C. (1992). A national customer satisfaction barometer: The Swedish experience. *Journal of Marketing*, *56*(1), 6–21.

Fornell, C., & Bookstein, F. L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research*, *19*(4), 440–452.

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American customer satisfaction index: Nature, purpose, and findings. *Journal of Marketing*, *60*(4), 7–18.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Fornell, C., Lorange, P., & Roos, J. (1990). The cooperative venture formation process: A latent variable structural modeling approach. *Management Science*, *36*(10), 1246–1255.

Fornell, C., & Robinson, W. T. (1983). Industrial organization and consumer satisfaction/dissatisfaction. *Journal of Consumer Research*, *9*(4), 403–412.

Fornell, C., Robinson, W. T., & Wernerfelt, B. (1985). Consumption experience and sales promotion expenditure. *Management Science*, *31*(9), 1084–1105.

Fox, J. (2006). Structural equation modeling with the SEM package in R. *Structural Equation Modeling*, *13*(3), 465–486.

Gerbing, D. W., & Anderson, J. C. (1985). The effects of sampling error and model characteristics on parameter estimation for maximum likelihood confirmatory factor analysis. *Multivariate Behavioral Research*, *20*(3), 255–271.

Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, *25*(2), 186–192.

Godambe, V. P. (1960). An optimum property of regular maximum likelihood estimation. *The Annals of Mathematical Statistics*, *31*(4), 1208–1211.

Goodhue, D., Lewis, W., & Thompson, R. (2006). *PLS, small sample size and statistical power in MIS research, in 39th Hawaii International Conference on System Sciences (HICSS-39)*. Maui, Hawaii: IEEE Computer Society.

Graham, J. L., Mintu, A. T., & Rodgers, W. (1994). Explorations of negotiation behaviors in ten foreign cultures using a model developed in the United States. *Management Science*, *40*(1), 72–95.

Gray, P. H., & Meister, D. B. (2004). Knowledge sourcing effectiveness. *Management Science*, *50*(6), 821–834.

Green, D. H., Barclay, D. W., & Ryans, A. B. (1995). Entry strategy and long-term performance: Conceptualization and empirical examination. *Journal of Marketing*, *59*(4), 1–16.

Grégoire, Y., & Fisher, R. J. (2005). The effects of prior relationships on consumer retaliation. *Advances in Consumer Research*, *32*, 98–99.

Haenlein, M., & Kaplan, A. M. (2004). A beginner's guide to partial least squares analysis. *Understanding Statistics*, *3*(4), 283–297.

Hennig-Thurau, T., Groth, M., Paul, M., & Gremler, D. D. (2006). Are all smiles created equal? How emotional contagion and emotional labor affect service relationships. *Journal of Marketing*, *70*(3), 58–73.

Hennig-Thurau, T., Henning, V., & Sattler, H. (2007). Consumer file sharing of motion pictures. *Journal of Marketing*, *71*(4), 1–18.

Hu, L. -T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparametrized model misspecification. *Psychological Methods*, *3*(4), 424–453.

Hui, B. S., & Wold, H. (1982). Consistency and consistency at large of partial least squares estimates. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation, Part II* Amsterdam: North Holland.

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. (1967). Some contributions to Maximum Likelihood factor analysis. *Psychometrika*, *32*(4), 443–482.

Jöreskog, K. G. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences*. New York: Seminar Press.

Jöreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association*, *70*(351), 631–639.

Jöreskog, K. G., & Sörbom, D. (1982). Recent developments in structural equation modeling. *Journal of Marketing Research*, *19*(4), 404–416.

Kim, S., & Hagtvet, K. A. (2003). The impact of misspecified item parceling on representing latent variables in covariance structure modeling: A simulation study. *Structural Equation Modeling*, *10*(1), 101–127.

Krijnen, W. P., Dijkstra, T. K., & Gill, R. D. (1998). Conditions for factor (in)determinacy in factor analysis. *Psychometrika*, *63*(4), 359–367.

Lohmöller, J. -B. (1988). The PLS program system: Latent variables path analysis with partial least squares estimation. *Multivariate Behavioral Research*, *23*(1), 125–127.

Long, J. S. (1983). *Covariance structure models: An introduction to LISREL.* Beverly Hills: Sage Publications.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modelling in psychological research. *Annual Review of Psychology*, *51*(1), 201–226.

MacCallum, R. C., & Browne, M. W. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin*, *114*(3), 533–541.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130–149.

Marsh, H. W., Hau, K. -T., Balla, J. R., & Grayson, D. (1998). Is more ever too much?: The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, *33*(2), 181–220.

Mattson, S. (1997). How to generate non-normal data for simulation of structural equation models. *Multivariate Behavioral Research*, *32*(4), 355–373.

Mayo, M. C., & Qualls, W. J. (1987). Household durable goods acquisition behavior: A longitudinal study. *Advances in Consumer Research*, *14*(1), 463–467.

McFarland, R. G., Bloodgood, J. M., & Payan, J. M. (2008). Supply chain contagion. *Journal of Marketing*, *72*(2), 63–79.

Micceri, T. (1989). The unicorn, the normal curve and other improbable creatures. *Psychological Bulletin*, *105*(1), 156–166.

Mitchell, V. L., & Nault, B. R. (2007). Cooperative planning, uncertainty, and managerial control in concurrent design. *Management Science*, *53*(3), 375–389.

Nasser, F., & Wisenbaker, J. (2003). A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. *Educational and Psychological Measurement*, *63*(5), 729–757.

Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling*, *8*(3), 353–377.

Qualls, W. J. (1988). Toward understanding the dynamic of household decision conflict behavior. *Advances in Consumer Research*, *15*(1), 442–448.

Qureshi, I., & Compeau, D. (2009). Assessing between-group differences in information systems research: A comparison of covariance- and component-based SEM. *MIS Quarterly*, *33*(1), 197–214.

R Development Core Team. (2008). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Reinartz, W. J., Echambadi, R., & Chin, W. W. (2002). Generating non-normal data for simulation of structural equation models using Mattson's method. *Multivariate Behavioral Research*, *37*(2), 227–244.

Reinartz, W., Krafft, M., & Hoyer, W. D. (2004). The customer relationship management process: Its measurement and impact on performance. *Journal of Marketing Research*, *41*(3), 293–305.

Rigdon, E. E. (2005). Structural equation modeling: Nontraditional alternatives. In B. Everitt & D. Howell (Eds.), *Encyclopedia of statistics in behavioral science* New York: John Wiley & Sons, Inc.

Schneeweiss, H. (1993). Consistency at large in models with latent variables. In K. Haagen, D. J. Bartholomew, & M. Deistler (Eds.), *Statistical modeling and latent variables*. Amsterdam: North-Holland.

Sharma, S., Durvasula, S., & Dillon, W. R. (1989). Some results on the behavior of alternate covariance structure estimation procedures in the presence of non-normal data. *Journal of Marketing Research*, *26*(2), 214–221.

Sirohi, N., McLaughlin, E. W., & Wittink, D. R. (1998). A model of consumer perceptions and store loyalty intentions for a supermarket retailer. *Journal of Retailing*, *74*(2), 223–245.

Smith, J. B., & Barclay, D. W. (1997). The effects of organizational differences and trust on the effectiveness of selling partner relationships. *Journal of Marketing*, *61*(1), 3–21.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y. -M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis*, *48*(1), 159–205.

Ulaga, W., & Eggert, A. (2006). Value-based differentiation in business relationships: Gaining and sustaining key supplier status. *Journal of Marketing*, *70*(1), 119–136.

Vale, D. C., & Maurelli, V. A. (1983). Simulating multivariate nonnormal distributions. *Psychometrika*, *48*(3), 465–471.

Vanhamme, J., & Snelders, D. (2003). What if you surprise your customers… Will they be more satisfied? Findings from a pilot experiment. *Advances in Consumer Research*, *30*(1), 48–55.

Velicer, W. F., & Fava, J. L. (1987). An evaluation of the effects of variable sampling on component, image and factor analysis. *Multivariate Behavioral Research*, *22*(2), 193–209.

Venkatesh, V., & Agarwal, R. (2006). Turning visitors into customers: A usability-centric perspective on purchase behavior in electronic channels. *Management Science*, *52*(3), 367–382.

White, J. C., Varadarajan, P. R., & Dacin, P. A. (2003). Market situation interpretation and response: The role of cognitive style, organizational culture and information use. *Journal of Marketing*, *67*(3), 63–79.

Williams, L. J., Edwards, J. R., & Vandenberg, R. J. (2003). Recent advances in causal modeling methods for organizational and management research. *Journal of Management*, *29*(6), 903–936.

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, & V. Capecchi (Eds.), *Quantitative sociology: International perspectives on mathematical and statistical modeling.* New York: Academic Press, Inc.

Zinkhan, G. M., & Fornell, C. (1989). A test of the learning hierarchy in high- and low-involvement situations. *Advances in Consumer Research*, *16*(1), 152–159.