Journal of Business Research 69 (2016) 3998–4010

Contents lists available at ScienceDirect

Journal of Business Research

Estimation issues with PLS and CBSEM: Where the bias lies!☆

Marko Sarstedt [a,d,*], Joseph F. Hair [b], Christian M. Ringle [c,d], Kai O. Thiele [c], Siegfried P. Gudergan [d]

[a] Otto-von-Guericke-University Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany
[b] University of South Alabama, Mitchell College of Business, Mobile, AL 36688, USA
[c] Hamburg University of Technology (TUHH), Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany
[d] University of Newcastle (Australia), University Drive, Callaghan, NSW 2308, Australia

A R T I C L E   I N F O

Article history:
Received 31 March 2016
Received in revised form 30 May 2016
Accepted 11 June 2016
Available online xxxx

Keywords:
Common factor models
Composite models
Reflective measurement
Formative measurement
Structural equation modeling
Partial least squares

A B S T R A C T

Discussions concerning different structural equation modeling methods draw on an increasing array of concepts and related terminology. As a consequence, misconceptions about the meaning of terms such as reflective measurement and common factor models as well as formative measurement and composite models have emerged. By distinguishing conceptual variables and their measurement model operationalization from the estimation perspective, we disentangle the confusion between the terminologies and develop a unifying framework. Results from a simulation study substantiate our conceptual considerations, highlighting the biases that occur when using (1) composite-based partial least squares path modeling to estimate common factor models, and (2) common factor-based covariance-based structural equation modeling to estimate composite models. The results show that the use of PLS is preferable, particularly when it is unknown whether the data's nature is common factor- or composite-based.

1. Introduction

The extent to which researchers raise an issue is a subtle indicator of its importance. The benefits and limitations of partial least squares path modeling (PLS) is one such issue that scholars have heatedly debated across a variety of disciplines including marketing (e.g., Fornell & Bookstein, 1982; Hair, Sarstedt, Ringle, & Mena, 2012), strategic management (e.g., Bentler & Huang, 2014; Rigdon, 2012, 2014; Sarstedt, Ringle, Henseler, & Hair, 2014), and management information systems (e.g., Goodhue, Lewis, & Thompson, 2012; Marcoulides & Saunders, 2006; Ringle, Sarstedt, & Straub, 2012). Such scientific debates are important since they serve as a catalyst that sparks further careful examination of a method's properties. Oftentimes, the result is improved understanding of the advantages and disadvantages of the focal method, but also additional research and methodological advances that stem from such objective and constructive discussions among scholars.

Recently, however, the scholarly community has witnessed a surprising level of acrimony towards PLS. Antonakis, Bendahan, Jacquart, and Lalive (2010, p. 1103) allude that "there is no use for PLS whatsoever [...] thus strongly encourage researchers to abandon it." Other authors similarly suggest that the use of PLS "is very difficult to justify" (Rönkkö & Evermann, 2013, p. 443) or that "PLS should not be adopted as a tool for psychological research." (Rönkkö, McIntosh, & Antonakis, 2015, p. 82). This new harshness climaxed in an editorial from the editors in chief of the Journal of Operations Management (Guide & Ketokivi, 2015, p. vii) who declared that they were "desk rejecting practically all PLS-based manuscripts."[1] In a follow-up paper in the very same journal, Rönkkö, McIntosh, Antonakis, and Edwards (2016, p. 16) echo this call by suggesting that "the only logical and reasonable action stemming from objective consideration of these issues is to discontinue the use of PLS."

Leaving aside the tone of these and similar statements, which aim at shutting down any scholarly debate, the question arises why these authors cannot find even a single positive attribute of PLS despite its acceptance in scholarly research. In an effort to disentangle these opposing views, Rigdon (2016) offers an in-depth discussion of PLS and its origins, concluding that critics just as proponents of the method frequently offer

☆ The authors thank George R. Franke (University of Alabama), Edward E. Rigdon (Georgia State University), and the participants of the 2nd International Symposium on Partial Least Squares Path Modeling at the University of Seville for their constructive feedback. The authors also thank Jörg Henseler, University of Twente, The Netherlands, for his support with the replication of Reinartz, Haenlein, and Henseler's study (2009), and with the development of the composite model-based data generation approach. Even though this research does not explicitly refer to the use of the statistical software SmartPLS (http://www.smartpls.com), Ringle acknowledges a financial interest in SmartPLS.
* Corresponding author.
E-mail addresses: marko.sarstedt@ovgu.de (M. Sarstedt), joefhair@gmail.com (J.F. Hair), ringle@tuhh.de (C.M. Ringle), k.thiele@tuhh.de (K.O. Thiele), siggi.gudergan@newcastle.edu.au (S.P. Gudergan).

[1] The positions advocated by these individuals can easily be interpreted as being against the use of structural equation modeling in general, and also extend to any multivariate analysis method, including ANOVA or multiple regression. Their assessment, if taken seriously, basically implies the end of empirical business research at least in the Journal of Operations Management, as long as these individuals are editors of that journal.

http://dx.doi.org/10.1016/j.jbusres.2016.06.007
0148-2963/

incorrect or incomplete rationale for avoiding as well as using PLS. In addition, Rigdon (2016) concludes that many misconceptions about PLS have their roots in the method's conceptual underpinnings and particularly the estimation philosophy it relies on (e.g., Rigdon, 2012).

In fact, when deciding to use PLS, researchers—consciously or unconsciously—opt for a composite-based approach to structural equation modeling (SEM) that linearly combines indicators to form composite variables (Lohmöller, 1989), which serve as proxies for the concepts under investigation (Rigdon, 2016). This approach is different from common factor-based SEM (i.e., covariance-based SEM; CBSEM), which considers the constructs as common factors that explain the covariation between their associated indicators. While this distinction has long been noted (e.g., Jöreskog & Wold, 1982; Schneeweiß, 1991), researchers have traditionally emphasized how PLS "is 'like' factor-based SEM but with advantages and disadvantages across different conditions" (Rigdon, 2012, p. 353)—see, for example, Hair et al. (2012); Kaufmann and Gaeckler (2015); Peng and Lai (2012). Only recently have scholars started calling for the emancipation of PLS from CBSEM by acknowledging its status as a purely composite-based method (e.g., Rigdon, 2012; Sarstedt, Ringle, Henseler et al., 2014). Addressing this call, Henseler, Hubona, and Ray (2016, p. 3) attempt to provide "an updated view on what PLS actually is" and suggest a set of guidelines for the interpretation and reporting of results that explicitly consider the distinction between composite-based SEM and common factor-based SEM. In their guidelines, the authors note that "PLS path models can contain two different forms of construct measurement: factor models or composite models" and continue by explaining that the depicted direction of arrows in the measurement model (i.e., reflective or formative) does not necessarily indicate whether PLS estimates a factor or composite model (Henseler, Hubona et al., 2016, p. 3).

Such statements leave many applied researchers confused as some misperceive the distinction between reflective and formative measurement specification on the one hand, and common factor and composite models on the other. The introduction of consistent PLS (PLSc), which Dijkstra and Henseler (2015) developed in an effort to align common factor and composite-based SEM methods, further contributed to the confusion. For example, some researchers have started using both PLS and PLSc—which assume fundamentally different measurement philosophies—on the same data without explicitly considering the nature of the data, model, and the implications of their choice of methods (e.g., Gelhard & von Delft, 2016). These issues are nicely reflected in a recent query by a thoughtful PhD student from the UK who asked one of this paper's authors, "what is the real difference between reflective constructs and factor models? What would be a reflective composite and what would be a common factor? And how is such difference transferred to the PLS context in terms of model specification?"

These queries constitute the research questions this paper sheds light on. By distinguishing measurement model conceptualization and operationalization from the model estimation perspective, this paper disentangles the confusion between reflective measurement and common factor models as well as formative measurement and composite models. More precisely, this paper's aim is to clarify the interplay between measurement model specification and model estimation via PLS using different estimation modes (i.e., Mode A vs. Mode B) and CBSEM. Understanding this interplay is of fundamental importance when deriving measures that suit a specific SEM method, or when choosing a specific SEM method that aligns with existing measures or a research objective. Results from a simulation study substantiate our conceptual considerations, highlighting the biases that occur when using composite-based PLS to estimate common factor models, and common factor-based CBSEM to estimate composite models. Specifically, our results show that PLS entails practically no bias when estimating data from a composite model population, regardless of the measurement model specification. In contrast, CBSEM and PLSc estimation of reflectively measured constructs when the data stem from a composite population show severe biases in parameter estimates, rendering their use inappropriate in these instances. Further comparisons with common factor model data show that the parameter bias resulting from using an SEM method on discrepant populations is much more severe for CBSEM than for PLS. The real bias results when researchers don't know the underlying data population (i.e., common factor or composite)—as is widespread in social sciences research—making PLS the preferred SEM method for most situations.

Based on our findings, we propose a framework that aligns different measurement and model estimation perspectives. This paper is written with the confidence that it will (1) offer researchers a clear roadmap for the conceptionalization and operationalization of their constructs, (2) provide guidance in their choice of the appropriate SEM method, and (3) ensure a more balanced perspective concerning recent criticism, which largely ignored the common factor vs. composite model distinction.

2. Measurement

2.1. Conceptual variables, constructs, and proxies

Irrespective of whether a deductive or an inductive research approach is undertaken by social science researchers, at some point in their search to better understand and explain theory, they deal with conceptual variables and theoretical models. A theoretical model reflects a set of structural relationships; usually based on a set of equations connecting conceptual variables that formalize a theory and visually represent the relationships (Bollen, 2002). As elements of theoretical models, conceptual variables represent broad ideas or thoughts about abstract concepts that researchers establish and propose to measure in their research (e.g., customer satisfaction).

Constructs represent conceptual variables in statistical models such as in a structural equation model.[2] They are intended to enable empirical testing of hypotheses that concern relationships between conceptual variables (Rigdon, 2012) and are conceptually defined in terms of the attribute and the object (e.g., MacKenzie, Podsakoff, & Podsakoff, 2011). The attribute defines the general type of property to which the focal concept refers, such as an attitude (e.g., attitude towards an advertisement), a perception (e.g., perceived ease of use of technology), or behavioral intention (e.g., purchase intention). The focal object is the entity to which the property is applied. For example, the focus of interest could be a customer's satisfaction with the products, satisfaction with the services, and satisfaction with the prices. In these examples, satisfaction constitutes the attribute, whereas products, services, and prices represent the focal objects.

Establishing a construct definition also includes determination of the dimensionality that describes the conceptual variable, with each dimension representing a different aspect (e.g., Law, Wong, & Mobley, 1998). A conceptual variable is not per se characterized as unidimensional or multidimensional, let alone two-, three- or four-dimensional (Bollen, 2011). Rather it depends on the context-specific definition of the conceptual variable and the denotation that comes with it. The denotation can, in principle, be infinite, since the same conceptual variable can represent different levels of theoretical abstraction across contexts (Diamantopoulos, 2005; Law & Wong, 1999). Thus, a construct definition is subject to the context within which a conceptual variable is examined such that the definition can change from one study to another and, accordingly, can differ in terms of dimensionality and the object of interest. For example, a customer's satisfaction with the service can be broken down into more concrete subdimensions, such as satisfaction with the speed of service, the servicescape, and the staff. The latter dimension can be differentiated into more concrete subdimensions such as satisfaction with the friendliness, competence, and outer appearance

[2] Note that researchers frequently distinguish between latent variables/constructs and composites, depending on the type of relationship assumed between the latent variable (composite) and its indicators (e.g., MacCallum & Browne, 1993). We use the term latent variable/construct to refer to the entities that represent conceptual variables in a structural equation model.

of the service staff. Each of these aspects can, in principle, be further broken down into yet more concrete subdimensions (e.g., Rossiter, 2011). Finally, the construct definition also clarifies how the abstract, conceptual variable relates to measurable, observable quantities. That is, the construct definition guides the conceptualization of the measurement models, which entails deciding whether to measure a construct reflectively or formatively.

Constructs are not just theoretical concepts under a different name as implied by commonly used definitions of this term (e.g., Bollen, 2002; Pedhazur & Pedhazur Schmelkin, 1991), but representations of conceptual variables in a statistical model. Importantly, constructs do not represent conceptual variables perfectly since any concept and any construct definition has some degree of ambiguity associated with it (e.g., Gilliam & Voss, 2013). In addition, constructs stem from data and therefore share the data's idiosyncrasies (Cliff, 1983; MacCallum, Browne, & Cai, 2007), which further detach them from the concepts they intend to represent. In this context, Michell (2013, p. 20) notes that constructs "are contrived in a way that is detached from the actual structure of testing phenomena and held in place by an array of quantitative methods, such as factor analysis, which gratuitously presume quantitative structure rather than infer it from the relevant phenomena (…)." Similarly MacCallum et al. (2007, p. 153) state that factor analytical procedures such as CBSEM cannot fully represent "the undoubtedly large number of minor common factors that influence measured variables and account in part for their intercorrelations. There are many other sources of error in such models. At best, a factor analysis model is an approximation of real-world phenomena." Against this background, Rigdon (2012, pp. 343–344) concludes that constructs should rather be viewed as "something created from the empirical data which is intended to enable empirical testing of propositions regarding the concept." That is, all measures of conceptual variables are approximations of or proxies for conceptual variables, independent from how they were derived (e.g., Wickens, 1972). Thus, irrespective of the quality with which a conceptual variable is theoretically substantiated and operationally defined and the rigor that encompasses measurement model development, any measurement in structural equation models produces only proxies for latent variables (Rigdon, 2012). This assessment is in line with the proliferation of all sorts of instruments that claim to measure essentially the same construct, albeit often with little chance to convert one instrument's measures into any other instrument's measures (Salzberger, Sarstedt, & Diamantopoulos, 2016). For example, business research and practice has brought forward a multitude of measurement instruments for corporate reputation, which rest on the same definition of the concept but differ fundamentally in terms of their underlying conceptualizations and measurement items (e.g., Sarstedt, Wilczynski, & Melewar, 2013).

2.2. Measurement model conceptualization and operationalization

Based on the construct definition, the next step is to specify a measurement model, which expresses how to measure the construct by means of a set of indicators (e.g., Jarvis, MacKenzie, & Podsakoff, 2003; MacKenzie, 2003). Generally, there are two broad ways to conceptualize measurement models (Coltman, Devinney, Midgley, & Venaik, 2008; Diamantopoulos & Winklhofer, 2001), which entail fundamentally different approaches to generating items (e.g., Churchill, 1979; Diamantopoulos & Winklhofer, 2001; MacKenzie et al., 2011). The first approach is referred to as reflective measurement. In a reflective measurement model the indicators are considered to be error-prone manifestations of an underlying construct with relationships going from the construct to its indicators (Bollen, 1989). The relationship between an observed and an unobserved variable is usually modeled as expressed in the following equation:

$$x = l \cdot Y + e, \qquad (1)$$

where x is the observed indicator variable, Y is the latent variable, the loading l is a regression coefficient quantifying the strength of the relationship between x and Y, and e represents the random measurement error.

Fig. 1 shows a reflective measurement model for a latent variable Y_1, measured with four indicators x_1, x_2, x_3, and x_4 as well as the conceptual variable the construct seeks to represent, illustrated by a triangle in the upper part of the figure (Rigdon, 2012). Reflective indicators, also referred to as effect indicators, can be viewed as a representative sample of all the possible items available within the conceptual domain of the construct (Nunnally & Bernstein, 1994). Since a reflective measurement model dictates that all items reflect the same construct, indicators associated with a particular construct should be highly correlated with each other (Edwards & Bagozzi, 2000). In addition, individual items should be interchangeable, and any single item can generally be left out without changing the meaning of the construct, as long as the construct has sufficient reliability (Jarvis et al., 2003). The fact that the relationship goes from the construct to its indicators implies that if the evaluation of the latent trait changes (e.g., because of a change in the standard of comparison), all indicators will change simultaneously (e.g., Diamantopoulos & Winklhofer, 2001).

The second approach is formative measurement. In a formative measurement model the indicators form the construct by means of linear combinations (Fig. 1). A change in an indicator's value due to, for example, a change in a respondent's assessment of the trait being captured by the indicator, changes the value of the construct. That is, "variation in the indicators precedes variation in the latent variable" (Borsboom, Mellenbergh, & van Heerden, 2003, p. 208), which means that, by definition, constructs with a formative measurement model are inextricably tied to their measures (Diamantopoulos, 2006). Besides the difference in the relationship between indicator(s) and construct, formative measurement models do not require correlated indicators. In practical applications, however, indicators in formative measurement models may be highly correlated, yielding satisfactory levels in reliability and validity statistics whose use, from a conceptual perspective, should be restricted to reflective measurement models (Hair et al., 2012).

Despite these clear conceptual differences, deciding whether to specify measurement models reflectively or formatively is not clear-cut in practice, as constructs do not inherently follow a reflective or formative measurement logic (e.g., Baxter, 2009). Rather, the researcher has the flexibility to conceptualize a measurement model based on the construct definition the researcher specifies. As Baxter (2009, p. 1377) notes, "there are often quite different possibilities for conceptualization of what might at first sight appear to be the same construct and, most importantly, there may be quite distinct lines of enquiry underlying the multiple possible conceptualizations." Consider, for example, the concept of perceived switching costs. Jones, Mothersbaugh, and Beatty (2000, p. 262) define perceived switching costs as "consumer perceptions of the time, money, and effort associated with changing service providers." Their measurement approach in the context of banking services draws on three items, which constitute reflections or consequences of perceived switching costs ("In general it would be a hassle changing banks," "It would take a lot of time and effort changing banks," and "For me, the costs in time, money, and effort to switch banks are high"). Hence, the authors implicitly assume that there is a concept of perceived switching costs, which can be manifested by querying a set of (e.g., three) items. Barroso and Picón (2012, p. 532), on the other hand, consider perceived switching costs as "a latent aggregate construct that is expressed as an algebraic composition of its different dimensions." These authors identify a set of six dimensions (benefit loss costs, personal relationship loss costs, economic risks costs, evaluation costs, set-up costs, and monetary loss costs), which represent certain specific characteristics, each covering an independent part of the perceived switching costs concept. As such, Barroso and Picón's (2012) construct definition of perceived switching costs follows a formative measurement model logic. Of course, the underlying items can

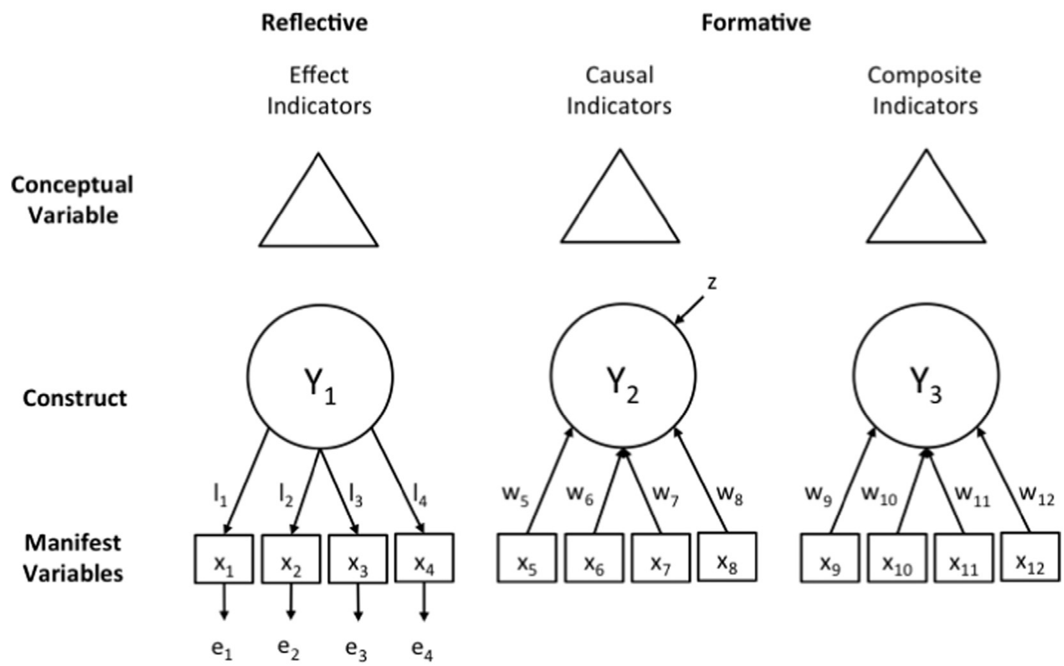


Fig. 1. Measurement model conceptualization and operationalization.

be empirically correlated, and perhaps causally related, but they are not actually exchangeable in the way the reflective measurement model conceptualization assumes they are (Rigdon et al., 2011). That is, their correlation is not because the construct of perceived switching costs is assumed to be their common cause. There are many more examples of constructs that carry the same label but which rely on different (i.e., reflective vs. formative) measurement model conceptualizations—see, for example, Albers (2010), Baxter (2009), and Chang, Franke, and Lee (2016).

Further contributing to the difficulties of deciding on the measurement perspective is the fact that there is not one type of formative measurement model—as had been implied in the early works on formative measurement (e.g., Diamantopoulos & Winklhofer, 2001) and the use of formative measurement models in statistical analysis (e.g., Hair, Ringle, & Sarstedt, 2011). Rather, two types of indicators exist in formative measurement models: causal indicators and composite indicators (Bollen, 2011; Bollen & Bauldry, 2011). Models with causal indicators follow a realist approach to measurement, which acknowledges that under any definition of a conceptual variable, there is a true value but this can never be measured with complete accuracy (e.g., Grace & Bollen, 2008). Therefore, the indicators should have conceptual unity in that all the indicators correspond to the researcher's definition of the concept (Bollen & Diamantopoulos, 2016). Breadth of coverage of the domain is extremely important to ensure that the domain of content is adequately captured: Omitting important indicators implies omitting a part of the conceptual variable that the construct represents (e.g., Bollen & Lennox, 1991).

Since causal indicators are expected to cover all aspects of the content domain (Bollen & Bauldry, 2011), constructs measured with causal indicators (Y_2 in Fig. 1) have an error term (z in Fig. 1). This error term captures all the other “causes” of the construct not included in the model (Diamantopoulos, 2006). Or as Diamantopoulos, Riefler, and Roth (2008, p. 1211–1212) note, “formative latent variables have a number of proximal causes, which researchers try to identify when conceptually specifying the construct. In many cases, however, researchers will be unable to detect all possible causes, as there may be some causes that have neither been discussed in prior literature nor revealed by exploratory research. The construct-level error term represents these missing causes.” Causal indicators themselves are, by definition, error free—that is, they are not subject to any systematic or random error. While this characteristic is fully comprehensible from a model estimation perspective (see Diamantopoulos, 2006), from a measurement perspective, there is no reason to assume that the sources of error that have traditionally been associated with reflective indicators do not apply to causal indicators. For example, why would the use of double-barreled items or of suggestive item wordings trigger error in a reflective indicator but not in a causal indicator? The following equation represents a measurement model comprised of causal indicators, where w_i indicates the contribution of x_i ($i = 1, \ldots, I$) to Y, and z is an error term associated with Y:

$$Y = \sum_{i=1}^{I} w_i \cdot x_i + z \quad (2)$$

The other type of indicators, referred to as composite indicators, closely resembles that of causal indicators except for one aspect. In contrast to constructs measured with causal indicators, constructs measured with composite indicators do not have an error term (see construct Y_3 in Fig. 1). This distinction has an important implication for the conceptualization of formative measurement models (Henseler et al., 2014) because composite indicators operate as contributors to a construct rather than truly “causing” it (Bollen, 2011; Bollen & Bauldry, 2011). They form the composite representing the construct in full by means of linear combinations. Therefore, a formative construct measured with composite indicators does not have an error term (i.e., the error term is set to zero). As with causal indicators, composite indicators are assumed to be error free. The following equation illustrates a measurement model with composite indicators, where Y is a linear combination of indicators x_i, each weighted by an indicator weight w_i (Bollen, 2011; McDonald, 1996):

$$Y = \sum_{i=1}^{I} w_i \cdot x_i \quad (3)$$

Although researchers have often used composite models and causal indicator models synonymously (e.g., Bollen & Lennox, 1991), more recently they have started distinguishing composite from causal indicators (e.g., Bollen, 2011; Bollen & Diamantopoulos, 2016; Howell, Breivik, & Wilcox, 2013). Thus, there is still some ambiguity regarding their nature and areas of application. For example, Bollen (2011, p. 366) notes that "it seems unlikely that there are many situations where an error term would be absent (…). This would mean that the latent variable that represents the unidimensional concept is an exact linear function of its indicators, which would seem to be a rarity." Bollen (2011) therefore treats the latent variables as if they were indeed the conceptual variables from a theoretical model (also see Bollen & Bauldry, 2011; Bollen & Diamantopoulos, 2016). However, viewing latent variables as proxies for a conceptual variable seems more reasonable and realistic (Rigdon, 2012), blurring the conceptual distinction between composite and causal indicators.

Furthermore, Bollen (2011, p. 366) asserts that "composite indicators need not share unidimensional conceptual unity. That is, composite indicators might be combined into a composite as a way to conveniently summarize the effect of several variables that do not tap the same concept although they may share a similar 'theme.'" Following this logic, measurement models with composite indicators only offer a means to model conceptual variables, for which elements are combined to form a new entity (Henseler, Hubona et al., 2016). This is particularly the case when analyzing secondary data, which typically lack a comprehensive theoretical substantiation and are collected for a purpose other than SEM (Rigdon, 2013). For example, a measurement model conceptualization of information search activities could be based on capturing the sum of the activities that customers engage in when seeking information from dealers, promotional materials, the Internet and other sources. Another researcher might choose a different set of variables to form a measure of information search activities. Thus, the items ultimately determine the meaning of the construct, which implies that adding or omitting an indicator potentially alters the nature of the construct. While this interpretation of composite indicators may be convenient for communication, it remains largely unclear where to draw a line between items having "conceptual unity" and sharing "a similar theme" (Bollen, 2011, p. 366).

In practice, researchers naturally choose items in operationalizing measurement models that match their construct definition, regardless of whether the actual measurement conceptualization draws on reflective, causal or composite indicators. That is, they treat the constructs in their studies as unitary entities just like Barroso and Picón (2012) do when offering an in-depth literature review of the nature and dimensionality of the perceived switching costs concept prior to deriving indicators in their operationalization of the construct's measurement model. As such, they fully comply with Rönkkö et al. (2016) who note that only a guiding conceptual framework and careful development of the indicator content imbues theoretical meaning upon factors. In fact, assuming that researchers use measures of composite indicators merely as convenient summaries of the data (Bollen, 2011) implies that the common practice of aggregating items as composites to represent constructs, even though commonly done in practically all non-SEM studies in every field of research, is without any theoretical justification and undermines the fundamentals of appropriate measurement. However, the very same measures in most instances have been carefully developed and tested following conventional measurement model evaluation guidelines—as extensively documented in standard measurement scale handbooks (e.g., Bearden, Netemeyer, & Haws, 2011; Bruner, James, & Hensel, 2001). Thus, the very activity of forming composites from validated measurement scales interweaves composite and causal indicators, casting doubt on the notion that the use of composites to represent conceptual variables is an outright abandonment of measurement theory as Rönkkö et al. (2016) imply.

Thus, composite indicators not only offer a way to conveniently summarize the data but can be used to measure any type of property to which the focal concept refers, including attitudes, perceptions, and behavioral intentions (e.g., Rigdon, 2012). As with any type of measurement conceptualization, however, researchers need to offer a clear construct definition and specify items that closely match this definition—that is, they must share conceptual unity.

Alternatively, measurement models with composite indicators can be interpreted as a prescription for dimension reduction, where the aim is to condense the measures so they adequately cover a conceptual variable's salient features (Dijkstra & Henseler, 2011). For example, a researcher may be interested in measuring the salient aspects of perceived switching costs by means of three (composite) indicators, which cover aspects particularly relevant to the study at hand (e.g., evaluation costs, set-up costs, and monetary loss costs).

3. Model estimation

3.1. PLS and CBSEM

The previous sections described different routes to operationalize constructs as proxies for conceptual variables. This measurement perspective needs to be complemented with the model estimation perspective, which explains how the different SEM techniques arrive at a solution and which assumptions underlie them. Researchers typically use two approaches to estimate structural equation models. One is the more widely applied CBSEM approach (Bollen, 1989; Diamantopoulos, 1994; Jöreskog, 1978); the other is PLS (Hair, Hult, Ringle, & Sarstedt, 2017; Lohmöller, 1989; Wold, 1982). While both complementary methods share the same basic aim, which is to estimate the relationships among constructs and indicators, they differ fundamentally in their statistical conceptions and particularly in the way they treat measurement models of constructs (Jöreskog & Wold, 1982).

CBSEM initially divides the variance of each indicator into two parts: (1) the common variance, which is estimated from the variance shared with other indicators in the measurement model of a construct, and (2) the unique variance, which consists of both specific and error variance (Bollen, 1989; Rigdon, 1998). The specific variance is assumed to be systematic and reliable while the error variance is assumed to be random and unreliable (i.e., measurement, sampling, and specification error). CBSEM initially calculates the covariances of a set of variables (common variance), and only that variance is included in any solutions derived. CBSEM, therefore, follows a common factor model approach in the estimation of the construct measures, which assumes that the variance of a set of indicators can be perfectly explained by the existence of one unobserved variable (the common factor) and individual random error (Spearman, 1927; Thurstone, 1947). The common factor model estimation approach conforms to the measurement philosophy underlying reflective measurement models.

In principle, CBSEM can also accommodate formative measurement models even though the method follows a common factor model estimation approach (e.g., Temme, Diamantopoulos, & Pfegfeidel, 2014). Analogous to the scientific realist perspective assumed in the method's treatment of reflective measurement models, formative measurement models in CBSEM typically assume causal indicators (Diamantopoulos, 2011). To estimate models with causal indicators, researchers must follow rules that require specific constraints on the model to ensure model identification (Bollen & Davies, 2009; Diamantopoulos & Riefler, 2011). As Hair et al. (2012, p. 420) note, "these constraints often contradict theoretical considerations, and the question arises whether model design should guide theory or vice versa."

As an alternative, CBSEM scholars have proposed the multiple indicators and multiple causes (MIMIC) model (e.g., Bollen, 1989; Jöreskog & Goldberger, 1975)—that includes both formative and reflective indicators (e.g., Diamantopoulos & Riefler, 2011; Diamantopoulos et al., 2008). While MIMIC models enable researchers to deal with the identification problem, they do not overcome the problem that formative measurement models with causal indicators invariably underrepresent the variance in the construct, since correlated indicators are required by the CBSEM common factor model to produce a valid proxy and thereby adequately represent a conceptual variable. As Lee and Cadogan (2013, p. 243) note, "researchers should not be misled into thinking that achieving statistical identification allows one to obtain information about the variance of a formative latent variable." Clearly, CBSEM at best only allows for approximating formative measurement models with causal indicators.

Similarly, CBSEM can accommodate formative measurement models with composite indicators (e.g., Diamantopoulos, 2011). Since constructs measured with composite indicators are defined by having zero variances, the identification of the construct's error variance is not an issue. Problems arise, however, with regard to the identification of all paths leading to as well as flowing out from the construct. Grace and Bollen (2008) suggest solving this problem by specifying a single incoming or outgoing path relationship to 1.0. While such specifications overcome parameter identification issues, they severely limit the interpretability of the estimates of the magnitude and significance of the fixed paths in the structural model (Grace & Bollen, 2008). Because of these limitations, several researchers conclude that CBSEM is not well suited for estimating formative measurement models (Hair et al., 2012; Peng & Lai, 2012; Reinartz, Haenlein, & Henseler, 2009).

Different from CBSEM, PLS does not divide the variance into common and unique variance. More precisely, the objective of PLS is to account for the total variance in the observed indicators rather than to explain only the correlations between the indicators (e.g., Tenenhaus, Esposito Vinzi, Chatelin, & Lauro, 2005). The logic of the PLS approach is, therefore, that in estimating the model relationships, all of the variance (common, unique and error) that the exogenous variables have in common with the endogenous variables should be included (e.g., McDonald, 1996). The underlying notion is that the indicators can be (linearly) combined to form composite variables that are comprehensive representations of the latent variables, and that these linear combinations are valid proxies of the conceptual variables under investigation (e.g., Henseler, Hubona et al., 2016). As such, PLS follows a composite model approach in the estimation of the construct measures, which generally conforms to the measurement philosophy underlying formative measurement models.

PLS's designation as composite-based refers only to the method's way to represent constructs that approximate the conceptual variables in a model. Although PLS draws on composites whose use has traditionally been considered to be consistent with formative measurement models but not reflective measurement models (e.g., Grace & Bollen, 2008), the method readily accommodates both measurement model types without identification issues (Hair et al., 2011). In estimating the model parameters, however, PLS always follows a composite model approach. That is, regardless of whether measurement models are reflective or formative, PLS always computes composite variables from sets of indicator variables as representations of the conceptual variables in the model. Three aspects are important in this regard.

First, in formative measurement models, PLS treats all indicators as composite indicators. That is, the method does not allow for the explicit modeling of a construct's error term measured with causal indicators (i.e., the error term z in Fig. 1 is constrained to zero). As a consequence and analogous to CBSEM, PLS only allows for approximating formative measurement models with causal indicators. Note, however, that actually no method can estimate formative measurement models unless reflective measures are simultaneously available.

Second, researchers have long noted that since PLS is based on the composite model logic, the method only approximates common factor-based reflective measurement models (Hui & Wold, 1982; also see Rigdon, et al., 2014). That is, from a model estimation perspective, PLS will produce "biased" estimates if the common factor model holds—just like CBSEM will produce "biased" estimates when using the method to estimate data generated from a composite model, as this study will show. However, the deviations in parameter estimates should not be considered a "bias" as both methods estimate different things and therefore may yield different values.

Third, to estimate the model parameters, PLS uses two modes, which relate to the way the method estimates the indicators weights that represent each indicator's contribution to the composite. Mode A corresponds to correlation weights derived from bivariate correlations between each indicator and the construct; Mode B corresponds to regression weights, the standard in ordinary least squares regression analysis. Regression weights not only take the correlation between each indicator and the construct into account but also the correlations between the indicators. No matter which mode for estimating the indicator weights is used, the resulting latent variable is always modeled as a composite (Henseler, Ringle, & Sarstedt, 2016). That is, since all multi-item measures are converted into weighted components—even in Mode A—PLS computes components by means of linear combinations of indicators.

PLS by default uses Mode A for reflectively specified constructs and Mode B for formatively specified constructs. Recent research, however, suggests that selecting the appropriate weighting mode requires a more thoughtful approach. Specifically, Becker, Rai, and Rigdon (2013) show that for formatively specified constructs, Mode A estimation yields better out-of-sample prediction for sample sizes larger than 100 and when the R^2 is moderate to large (i.e., $R^2 \geq 0.30$). For large sample sizes and large R^2 values, Mode A and Mode B perform equally well in terms of out-of-sample prediction. In terms of parameter accuracy in the structural model, Mode A performs best when sample size or R^2 values are small to medium. For larger sample sizes or R^2 values, Mode A and Mode B estimations do not differ in terms of parameter accuracy.

From a measurement perspective, PLS and CBSEM both share an approximation character as constructs do not necessarily fully correspond to the conceptual variables they represent. As noted by Rigdon (2016, p. 19), "common factor proxies cannot be generally assumed to carry greater significance than composite proxies in regard to the existence or nature of conceptual variables." A similar view is echoed in the intense debates on the relative advantages of component versus common factor analysis in the 90s, which witnessed a series of articles and commentaries on the conceptual and philosophical underpinnings of the methods. Summarizing these debates, Bandalos and Boehm-Kaufman (2009, p. 70) note that "although methodologists still disagree about which model is most appropriate, component analysis and common factor analysis have different goals and are based on different philosophies." Rejecting the reflex-like adherence to the common factor model, researchers have long warned that the common factor model rarely holds in applied research (Schönemann & Wang, 1972). For example, among 72 articles published during 2012 in what Atinc, Simmering, and Kroll (2012) consider the four leading management journals (*Academy of Management Journal*, *Journal of Applied Psychology*, *Journal of Management*, and *Strategic Management Journal*) that tested one or more common factor model(s), fewer than 10% contained a common factor model that did not have to be rejected. In light of these results, Henseler et al. (2014, p. 184) conclude "from a philosophical standpoint, there is no need for modeling constructs as common factors (…), and reducing SEM to common factor models is a very restrictive (unnecessarily restrictive, we would argue) view about SEM."

4. Using PLS to estimate common factor models vs. using CBSEM to estimate composite models

4.1. The parameter estimation bias

The previous discussions showed that PLS and CBSEM assume different ways of how the data represent measurement models that the researcher—in line with a set of construct definitions—specifies in a reflective or formative way. CBSEM assumes the data follow a common factor model in which the indicator covariances define the nature of the data, whereas PLS adheres to a composite model approach in which data are defined by means of linear combinations of indicators. So while the measurement models may follow a reflective (or formative) specification, the underlying data model may be composite-based (or common factor-based).

Numerous studies have explored PLS's performance in terms of parameter accuracy when data are assumed to follow a common factor model approach (e.g., Barroso, Cepeda Carrión, & Roldán, 2010; Hwang, Malhotra, Kim, Tomiuk, & Hong, 2010; Marcoulides, Chin, & Saunders, 2012; Reinartz et al., 2009). Overall, these studies suggest that the bias that PLS produces when estimating common factor models is comparably small provided that the measurement models meet minimum recommended standards in terms of the number of indicators and indicator loadings. Recent efforts to dramatize the differences between CBSEM and PLS estimates (Rönkkö et al., 2016) in, for example, Reinartz et al.'s (2009) study focused on descriptive differences between population values and parameter estimates only, disregarding the concept of statistical inference. As Reinartz et al. (2009, p. 338; emphasis added by the authors) note in their results description of all simulation conditions, "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)." Only when the model estimation draws on a very large sample size (N = 10,000) and includes measurement models with many indicators with high loadings, did statistically significant differences occur. Correspondingly, empirical studies using both methods suggest that the divergence between PLS and CBSEM results when estimating common factor models is of little practical relevance for the results' implications (e.g., Astrachan, Patel, & Wanzenried, 2014).

The question, however, is whether the bias identified in prior studies results from using composite-based PLS on common factor model data or if the method is inherently biased, including when estimating composite models. Similarly, while the (supposed) PLS bias has been extensively debated in the literature, the bias that CBSEM produces when mistakenly estimating composite models has not yet been explored. For this reason, the following simulation studies focus on revealing the biases that occur when using (1) composite-based PLS to estimate common factor models, and (2) common factor-based CBSEM to estimate composite models. Furthermore, both studies consider PLSc.

4.2. Simulation studies

Our studies replicate Reinartz et al.'s (2009) simulation study on the comparative performance of PLS and CBSEM, which in its original form assumed a common factor model. We extended the original study, however, by additionally generating composite model-based data. Furthermore, our studies also consider PLSc, which follows a composite modeling logic but mimics a common factor model (Sarstedt, Ringle, & Hair, 2014). To do so, the method first estimates the model parameters using the standard PLS algorithm and corrects these estimates for attenuation using the consistent reliability coefficient ρ_A. This correction only applies to reflective measurement models, while formative measurement models remain unchanged.

The path model and path coefficient specifications used in the simulations (Fig. 2) are identical to Reinartz et al. (2009) with low (i.e., 0.15; p_1, p_2, p_{12}), medium (i.e., 0.30; p_5), and high (i.e., 0.50; p_3, p_4, p_6, p_9, p_{10}, p_{11}) pre-specified path coefficients. Accounting for corresponding calls in the literature (Marcoulides et al., 2012), we extended the original model by adding a construct (Y_5) with two null paths (p_7 and p_8). Also analogous to Reinartz et al. (2009), all measurement models are reflective. Table 1 illustrates the design factors and their levels manipulated in the simulation study. The simulation study uses a factorial design. We conducted 300 replications of each factor-level combination to obtain stable average outcomes for our analysis. In summary, the analysis includes $4 \cdot 4 \cdot 3 \cdot 5 \cdot 300 = 72{,}000$ datasets for Study I (i.e., the common factor-based simulation) and $8 \cdot 3 \cdot 5 \cdot 300 = 36{,}000$ datasets for Study II (i.e., the composite-based simulation), which results in a total number of 324,000 computations for the three methods under research.

In line with related research in the field (e.g., Becker, Rai, Ringle, & Völckner, 2013; Reinartz et al., 2009), common factor model-based data generation was performed by means of Mattson's (1997) method (also see Reinartz, Echambadi, & Chin, 2002), where univariate random variables initially serve the generation of the latent variables in the structural model, followed by the computation of the observed variables. The composite model-based data generation used in this study draws on a procedure similar to the one that Schlittgen (2015) presents in his SEGIRLS package for the statistical R software (R Core Team, 2014). We first generate the model-implied covariance matrix of the indicators, followed by a Cholesky decomposition, and finally extract the indicator data for a pre-specified number of observations and the sought data distribution. For model estimation based on PLS, PLSc, and CBSEM, we use the semPLS (Monecke & Leisch, 2012), matrixpls R (Rönkkö, 2016), and sem (Fox et al., 2015) packages of the R software. As in Reinartz et al. (2009), CBSEM estimation draws on the standard maximum likelihood approach; PLS uses Mode A estimation while PLSc uses Mode A estimation followed by the correction for attenuation in both studies.

5. Results

The assessment of the methods' parameter accuracy occurs on the grounds of the mean absolute error MAE, which is defined as

$$\text{MAE} = \frac{1}{t}\sum_{j=1}^{t}\left|\hat{\theta}_j - \theta_j\right|, \quad (4)$$

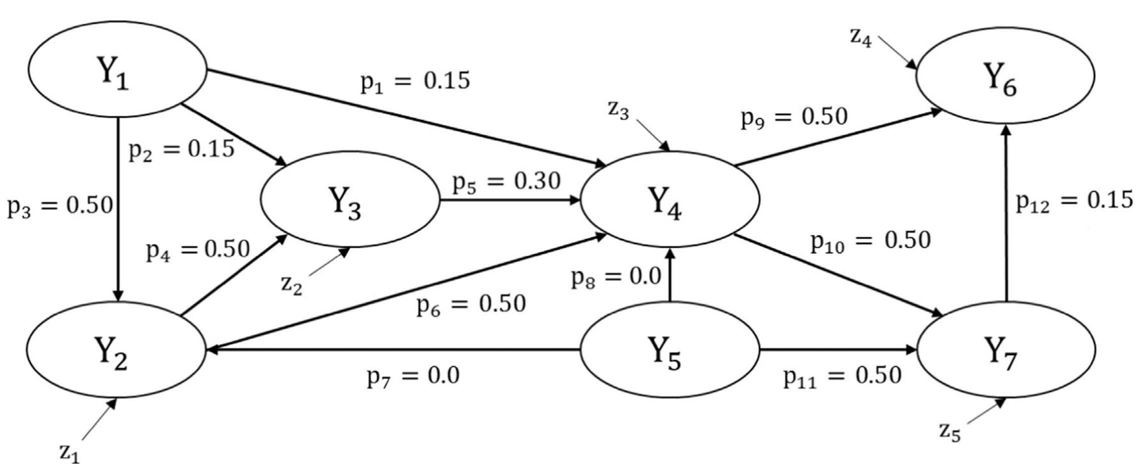


Fig. 2. Simulation model.

Table 1
Simulation design.

Design factors	Study	
	Study I: Common factor-based simulation	Study II: Composite-based simulation
Representation of the constructs and data generation	Common factor model	Composite model
Structural model	Reinartz et al. (2009), extended by null paths	Reinartz et al. (2009), extended by null paths
Loadings/weights and number of indicators	All possible combinations of number of indicators (2, 4, 6, 8) and indicator loadings (equal: 0.50, 0.70 or 0.90; and unequal: half of the indicators 0.50, the other half 0.90)	Eight combinations of number of indicators and indicator weights: Unequal weights • 0.30/0.90; • 0.10/0.30/0.50/0.70; • 0.10/0.175/0.25/0.325/0.40/0.475; and • 0.075/0.125/0.175/0.225/0.275/0.325/0.375/0.425 Equal weights • 0.625/0.625; • 0.40/0.40/0.40/0.40; • 0.30/0.30/0.30/0.30/0.30/0.30; and • 0.25/0.25/0.25/0.25/0.25/0.25/0.25/0.25
Data distribution	Three variations of skewness/kurtosis: none, moderate, and high	Three variations of the normal distribution: symmetric normal, log-normal, and diff-normal
Sample size	100, 250, 500, 1000, and 10,000	100, 250, 500, 1000, and 10,000

Notes: In composite models, the indicators fully explain the latent variable, which imposes some restrictions on the possibilities of cases that can be drawn; log-normal = random variable which has a standard normally distributed logarithm; diff-normal = difference of two log-normal distributions.

where t equals the number of parameters, θ_j is the prespecified parameter and $\hat{\theta}_j$ is the parameter estimate in any replication. Tables 2 and 3 illustrate the results of the simulation studies. Our illustration focuses on the case of normally distributed data as the analysis of non-normal data yields highly similar results.

Our results confirm the well-known PLS bias when using the method to estimate the path model with common factor model-based data. PLS produces biased estimates with average MAE values of around 0.10 when the measurement models only have two indicators or when the loadings are low (i.e., 0.50). Confirming PLS's consistency at large (Hui & Wold, 1982), PLS's MAE values decrease for greater numbers of indicators per measurement model or higher sample sizes. Compared to PLS, CBSEM achieves lower MAE values across all conditions except for small sample sizes of 100. In this condition, PLSc shows pronounced MAE values of up to 0.34 and also performs weak at 250 observations when measurement models have only two indicators or low loadings. However, PLSc's performance increases considerably with more indicators and higher sample sizes. On average across all simulation conditions, PLS and PLSc have a higher MAE (0.07) compared to CBSEM (0.05). Clearly, the differences between the three methods when used on common factor model-based data are overall only marginal, however.

A different picture emerges when estimating data from a composite model population. Whereas PLS has an overall MAE value of 0.04, the parameter biases of CBSEM (0.76) and particularly PLSc (3.70) are much more pronounced. PLSc shows a bewildering performance across the simulation conditions with MAE values ranging from 0.64 to 17.89. Specifically, in conditions with four indicators, equal weights and 500, and 10,000 observations, respectively, MAE values bounce up to values higher than 10. To rule out potential problems resulting from the PLSc implementation of the matrixpls package, we re-ran the simulation study using our PLSc extension of the semPLS package (Monecke & Leisch, 2012). Results from this additional analysis provided support for the extent and variation of PLSc's bias with MAE values well above 10 for several simulation conditions. Similar to PLSc, CBSEM shows pronounced parameter estimation biases across all simulation conditions but at a much lower level with MAE values ranging from 0.51 to 1.72. Nevertheless, CBSEM's bias when estimating composite models is on average 11 times higher than PLS's bias when estimating common factor models. Clearly, the use of PLS to estimate common factor models is much less of an issue than using CBSEM on data consistent with the composite model. Finally, while PLS's MAE values decrease when sample sizes increase, this is not the case with CBSEM and PLSc. For these two methods, the MAE values show no clear pattern. For example, for measurement models with 2 indicators and unequal weights, PLSc's MAE values decrease from 100 to 500 observations, increase at 1000 observations, and finally decrease at 10,000 observations. Overall our simulation study results suggest that when the underlying model type is unknown, researchers are well advised to draw on PLS in order to avoid substantial parameter biases that result from using PLSc or CBSEM in case the composite model holds.

In summary, we find that the methods' parameter bias depends on the underlying model and data. If one assumes a common factor model and draws on data of such a nature, CBSEM generally performs—as expected—very well. The same generally holds for PLSc, except when the sample size is small. The PLS method offers a very good approximation in this case. At the same time, PLS performs—as expected—very well for composite models, if one draws on data of such a nature. In this case, however, CBSEM and PLSc perform very poorly.

Table 4 summarizes the results of prior research on the methods' performance in terms of parameter bias when estimating common factor models with effect, causal, and composite indicators. Furthermore, the table summarizes the results of this paper's simulation studies regarding the methods' performances when estimating composite models with effect indicators. In line with the nature of each data generation approach and the methods' way of treating construct measures (e.g., Diamantopoulos, 2011; Henseler, Hubona et al., 2016), we differentiate between (1) effect indicator models and causal indicator models when the underlying population is common factor-based, and (2) between effect indicator models and composite indicator models when the underlying population is composite-based.

6. Conclusion

"Professional statisticians tend to know little about factor analysis and seldom practice it. Indeed, statisticians mostly have a cool negative attitude towards the subject. They hardly ever write about it. […] I can see nothing advantageous in factor analytic methods. Factor analysis is technically under-developed and at times appears almost cretinous. Its practitioners seem to be largely unaware of the technical and methodological problems, which they let themselves in for." This text, which is more than fifty years old and taken from Ehrenberg's (1962, p. 191 and p. 206) article "Some Questions About Factor Analysis", appears

Table 2
Coefficients' mean absolute error (MAE) in the common factor model situation.

Design factor			Mean absolute error (MAE)		
Observations	Group	Loadings	PLS	PLSc	CBSEM
100	2 indicators	Mixed	0.11	0.30	0.13
	4 indicators		0.09	0.13	0.10
	6 indicators		0.08	0.09	0.08
	8 indicators		0.08	0.12	0.08
	Loadings: 0.5	Equal[a]	0.13	0.34	0.16
	Loadings: 0.7		0.09	0.10	0.10
	Loadings: 0.9		0.06	0.07	0.07
	Loadings: 0.5/0.9	Unequal[a]	0.08	0.10	0.08
250	2 indicators	Mixed	0.10	0.13	0.09
	4 indicators		0.08	0.07	0.06
	6 indicators		0.06	0.05	0.05
	8 indicators		0.06	0.06	0.05
	Loadings: 0.5	Equal[a]	0.12	0.14	0.10
	Loadings: 0.7		0.07	0.06	0.06
	Loadings: 0.9		0.04	0.04	0.04
	Loadings: 0.5/0.9	Unequal[a]	0.06	0.05	0.05
500	2 indicators	Mixed	0.10	0.07	0.06
	4 indicators		0.07	0.05	0.04
	6 indicators		0.06	0.04	0.04
	8 indicators		0.05	0.04	0.04
	Loadings: 0.5	Equal[a]	0.12	0.08	0.07
	Loadings: 0.7		0.07	0.04	0.04
	Loadings: 0.9		0.03	0.03	0.03
	Loadings: 0.5/0.9	Unequal[a]	0.05	0.04	0.03
1000	2 indicators	Mixed	0.09	0.05	0.04
	4 indicators		0.06	0.03	0.03
	6 indicators		0.05	0.03	0.03
	8 indicators		0.05	0.03	0.03
	Loadings: 0.5	Equal[a]	0.12	0.05	0.05
	Loadings: 0.7		0.06	0.03	0.03
	Loadings: 0.9		0.03	0.02	0.02
	Loadings: 0.5/0.9	Unequal[a]	0.05	0.03	0.02
10,000	2 indicators	Mixed	0.09	0.01	0.01
	4 indicators		0.06	0.01	0.01
	6 indicators		0.04	0.01	0.01
	8 indicators		0.04	0.01	0.01
	Loadings: 0.5	Equal[a]	0.11	0.02	0.01
	Loadings: 0.7		0.06	0.01	0.01
	Loadings: 0.9		0.02	0.01	0.01
	Loadings: 0.5/0.9	Unequal[a]	0.04	0.01	0.01
Total			**0.07**	**0.07**	**0.05**

[a] Across all numbers of indicators.

Table 3
Coefficients' mean absolute error (MAE) in the composite model situation.

Design factor			Mean absolute error (MAE)		
Observations	Indicators	Weights	PLS	PLSc	CBSEM
100	2	Equal	0.07	2.90	0.84
	4[a]		0.07	4.42	0.74
	6[a]		0.07	3.57	0.63
	8[a]		0.07	2.05	0.52
	2	Unequal	0.07	5.05	0.61
	4[a]		0.08	3.03	0.92
	6[a]		0.07	5.83	0.54
	8[a]		0.07	2.89	0.48
250	2	Equal	0.05	5.33	0.81
	4[a]		0.04	5.56	0.81
	6[a]		0.05	3.22	0.73
	8[a]		0.04	3.71	0.57
	2	Unequal	0.05	2.94	0.54
	4[a]		0.05	4.06	0.82
	6[a]		0.05	2.76	0.58
	8[a]		0.05	4.20	0.59
500	2	Equal	0.03	5.21	0.90
	4[a]		0.03	11.55	0.81
	6[a]		0.03	2.87	0.82
	8[a]		0.03	2.89	0.75
	2[a]	Unequal	0.03	2.29	0.55
	4[a]		0.03	1.17	0.62
	6[a]		0.03	2.52	0.60
	8[a]		0.03	4.96	0.56
1000	2	Equal	0.02	2.28	0.98
	4[a]		0.02	5.52	0.80
	6[a]		0.02	1.38	0.95
	8[a]		0.02	2.74	0.88
	2[a]	Unequal	0.03	4.39	0.55
	4[a]		0.02	0.81	0.65
	6[a]		0.02	2.29	0.65
	8[a]		0.02	6.88	0.65
10,000	2	Equal	0.01	1.00	1.34
	4		0.01	17.89	0.74
	6[a]		0.01	0.85	1.72
	8[a]		0.01	1.06	1.22
	2[a]	Unequal	0.01	2.44	0.66
	4[a]		0.01	0.64	0.51
	6[a]		0.01	1.04	1.02
	8[a]		0.01	1.59	0.58
Total			**0.04**	**3.70**	**0.76**

[a] Instances in which CBSEM converged in <50% of the simulation runs.

surprising considering that today factor analysis is one of the success stories of statistical analysis (Cudeck & MacCallum, 2007). This assessment sounds familiar to everyone who has been exposed to recent papers critically referring to the PLS method. Authors have repeatedly suggested that PLS has "largely been ignored in research methods journals" (Rönkkö & Evermann, 2013, p. 426), that its use is restricted to few domains (Rönkkö et al., 2016; Rönkkö et al., 2015) and that "PLS is not useful for statistical estimation and testing" (Rönkkö et al., 2015, p. 76). While we do not suggest that PLS will undergo a similar development as factor analysis, the statements about the limitations of factor analysis and PLS nicely show how unsubstantiated some methodological discussions can become. As noted elsewhere, "any extreme position that (oftentimes systematically) neglects the beneficial features of the other technique, and may result in prejudiced boycott calls, is not good research practice and does not help to truly advance our understanding of methods and any other research subject" (Sarstedt, Ringle, Henseler et al., 2014, p. 158).

Our discussions show that researchers need to clearly distinguish between (conceptual) measurement approaches and the (statistical) estimation perspectives when judging the appropriateness of or choosing a specific SEM method. Model estimation does not occur in a methodological vacuum detached from measurement considerations but rests on specific assumptions, which need to be considered when conceptualizing and operationalizing models and vice versa. Despite frequent warnings (Chin, 2010; Henseler et al., 2014; Marcoulides et al., 2012), research on the performance of PLS has repeatedly ignored the implications of using a composite-based method for estimating common factor models (Becker, Rai, & Rigdon, 2013). Recent efforts to align reflective measurement and composite-based modeling (Dijkstra & Henseler, 2015; Henseler, Hubona et al., 2016)—while commendable from a methodological viewpoint—have instead contributed to the confusion, leaving researchers with little guidance regarding when to apply each method and how to align their use with measurement considerations.

The framework in Fig. 3 merges our theoretical discussions and simulation results. Whereas the theoretical layer serves to define the conceptual variable, the conceptual layer delivers the operational definition of the conceptual variables, which then serves as the basis for the measurement operationalization using effect, causal, or composite indicators on the operational layer. This conceptualization and operationalization of construct measures represents the measurement perspective. This perspective needs to be complemented with the model estimation perspective. The estimation layer intertwines with the measurement model layer that expresses how the data represent reflectively or formatively specified measurement models.

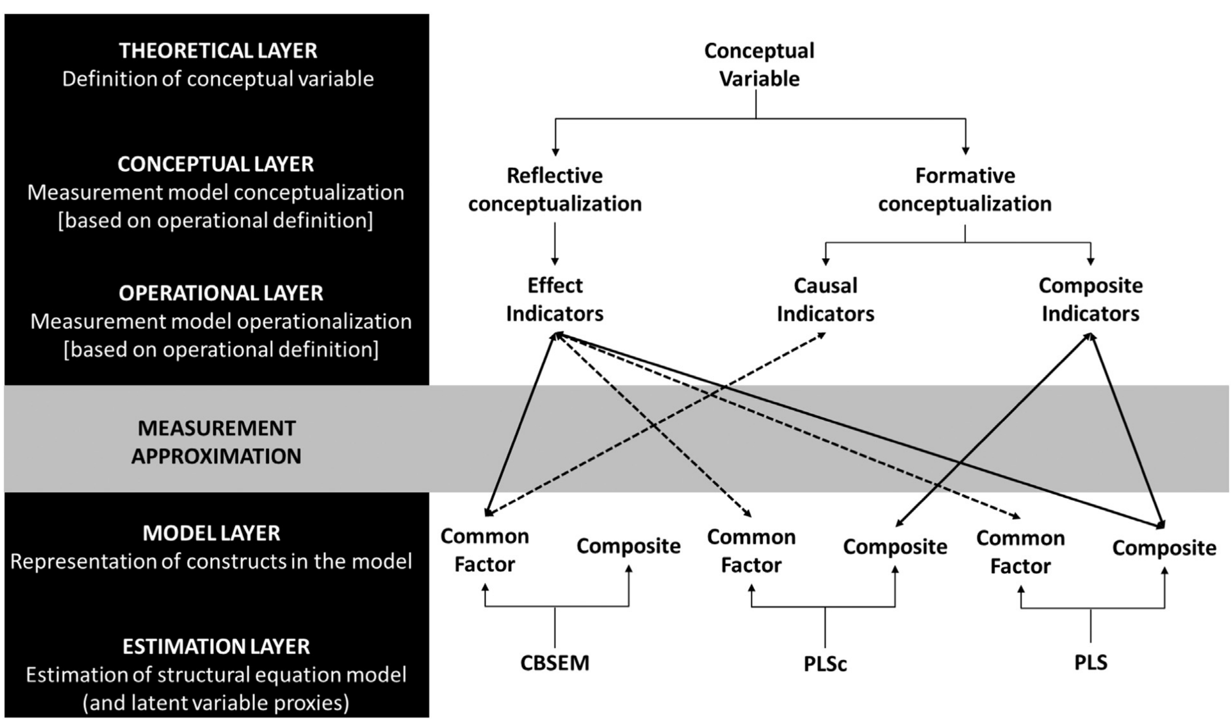
By exploring the performance of CBSEM, PLS, and PLSc when estimating composite models, the simulation studies overcome a crucial limitation of prior studies, which univocally relied on data from

Table 4
Efficacy of PLS, PLSc, and CBSEM for estimating common factor and composite models.

Method	Data type	Measurement model operationalization	Performance	Description	Sample references
PLS	Common factor	Effect indicators	(+)	Small bias	Reinartz et al. (2009), this study
		Causal indicators	−	PLS cannot model the construct-level error term	Diamantopoulos (2011)
	Composite	Effect indicators	+	Very small bias	This study
		Composite indicators	+	Small bias for small sample sizes, which approaches zero for increasing sample sizes	Becker, Rai, Rigdon (2013)
PLSc	Common factor	Effect indicators	(+)	Small bias, which is however higher than that of CBSEM for small sample sizes	Dijkstra and Henseler (2015), this study
		Causal indicators	−	PLSc cannot model the construct-level error term	Diamantopoulos (2011)
	Composite	Effect indicators	−	Pronounced bias	This study
		Composite indicators	+	Results parallel those from PLS as no correction for attenuation occurs	Becker, Rai, Rigdon (2013)
CBSEM	Common factor	Effect indicators	+	Very small bias except for small sample sizes of 100	Reinartz et al. (2009), this study
		Causal indicators	(+)	Model identification via MIMIC models or by specification of select parameters	Diamantopoulos and Riefler (2011)
	Composite	Effect indicators	−	Increased bias	This study
		Composite indicators	−	Model identification by specification of select parameters but strong limitations in terms of inference	Grace and Bollen (2008)

Notes: − not recommended/possible, (+) acceptable, + recommended.

common factor model populations to judge their universal efficacy (Chin, 2010; Marcoulides & Chin, 2013). Therefore, our studies address corresponding calls for future research, such as expressed by Hwang et al. (2010, p. 710) in their comparative study on parameter recovery of common factor-based SEM, PLS, and generalized structured component analysis: "We generated simulated data on the basis of covariance structure analysis. This data generation procedure may have had an unfavorable effect on the performance of partial least squares and generalized structured component analysis. We adopted the procedure because it was rather difficult to arrive at an impartial way of generating synthetic data for all three different approaches. Nevertheless, the same procedure has been used in other studies that compared the performance of covariance structure analysis with that of partial least squares (…). In any case, it appears necessary in future studies to



Notes: Dashed lines indicate acceptable types of measurement approximation; solid lines represent recommended types of measurement approximation. The PLSc results when estimating composite model data and composite indicators parallel those from PLS as no correction for attenuation occurs.

Fig. 3. Measurement and model estimation framework.

investigate whether a particular data generation procedure may influence the relative performance of the different approaches." Our results show that Hwang et al.'s (2010) notion can be answered with a resounding "yes."

The results outlined in this paper show that PLS entails practically no bias when estimating data from a composite model population, regardless of whether the measurement models are reflective or formative (Table 3). Biases are somewhat higher for common factor model populations (Table 2), but low in absolute terms. Clearly, PLS is optimal for estimating composite models while simultaneously allowing the approximation of common factor models involving effect indicators with practically no limitations (see the solid lines between composite indicators/effect indicators and composite model and the dashed line between effect indicators and common factor model in Fig. 3; also see Table 4). In contrast, CBSEM and PLSc estimation of reflectively measured constructs when the data stem from a composite population entails severe biases in parameter estimates, rendering their use inappropriate in these instances (no line between effect indicators and composite model in Fig. 3; also see Table 4). Particularly PLSc shows a bewildering behavior with strong biases across practically all conditions, which do not diminish as sample size increases. When using PLSc to estimate measurement models with composite indicators using data that stem from a composite model population, the PLSc results parallel those from PLS as no correction for attenuation occurs (see the solid line between composite indicators and composite model in Fig. 3; also see Table 4).

When estimating data from common factor populations, CBSEM's parameter bias is small for a sample size of 250 and quickly diminishes for higher sample sizes (see solid line between effect indicators and common factor model in Fig. 3; also see Table 4). PLSc shows a similar pattern when estimating data from common factor populations but performs less well for small sample sizes of 100, where MAE values peak at 0.34 (see dashed line between effect indicators and common factor model in Fig. 3; also see Table 4). In this situation, PLS outperforms the other methods but overall, the differences are marginal (Chin, 1998; Fornell & Bookstein, 1982; also see Goodhue et al., 2012). Note that other CBSEM estimators than ML (e.g., GLS, ULS, and ADF) entail further biases when estimating common factor models; see for example Boomsma and Hoogland (2001) and Dijkstra and Henseler (2015).

The obvious problem with these observations is that researchers can hardly know whether the data's nature is common factor- or composite-based. Fit measures such as the standardized root mean square residual (SRMR) may provide an indication of whether the data follow a common factor model. If the specific measurement model does not meet the required level (e.g., 0.08 and smaller for the SRMR; Hu & Bentler, 1998), this result suggests that the data follow a composite model. Alternatively, an improper CBSEM solution may point to an underlying composite model population. Our results show that CBSEM produces improper solutions in up to 99% of cases when the composite model holds. In consideration that in practical applications improper solutions often occur in CBSEM use (Rigdon, 2012; Sarstedt, Ringle, Henseler et al., 2014), these results offer a potential explanation why, more often than not, the common factor model cannot be supported in practice (Atinc et al., 2012). At the same time, however, reasons for improper solutions are manifold and not restricted to the misspecification of the model type. Therefore, interpreting improper solutions as clear evidence for an underlying composite model is not reasonable. As an alternative, researchers can follow a multi-methods approach, in which they combine CBSEM with PLS to see whether the results align in that specific research situation. Substantial differences between the methods indicate that the underlying population is composite-based, supporting the use of composite-based SEM methods. Nevertheless, in light of the biases that come with a CBSEM and PLSc-based estimations of composite model data, PLS is certainly the safer option when estimating data from an unknown population until research has proposed clear guidelines on how to identify the population type.

Our findings suggest that composite-based methods are going to play a greater role in future SEM applications. To date our understanding of this strand of methods is incomplete, however, as prior assessments universally drew on common factor model-based data and thereby relied on misspecified populations (Rigdon, 2016). Therefore, future research should aim at broadening our knowledge of the relative performance of the different approaches on the grounds of composite model-based data. For example, studies should contrast PLS's performance with other composite-based SEM techniques such as generalized structured components analysis (Hwang et al., 2010) or regularized generalized canonical correlation analysis (Tenenhaus & Tenenhaus, 2011).

In doing so, future research should consider a broader range of model constellations and more complex model structures such as hierarchical component models, moderating effects, or nonlinear effects. Such assessments would help disclose the different methods' efficacy for different situations that researcher encounter in their studies. By examining CBSEM's performance on composite model data, this study complements prior research, which univocally examined PLS's performance on common factor model data. However, future research should compare PLS, PLSc, and CBSEM on data where both models fit in the population. Such a design would provide supplementary insights, because CBSEM may work well on some composite measures but not others. In addition, with regards to PLS, future research should explore the interplay between measurement specifications, population type, and PLS's estimation modes (i.e., Mode A and B). These results would help clarifying the estimation modes' efficacy for out-of-sample prediction, in-sample-prediction, and parameter bias under different model specification and data conditions.

References

Albers, S. (2010). PLS and success factor studies in marketing. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications (Springer handbooks of computational statistics series, vol. II)* (pp. 409–425). Heidelberg, Dordrecht, London, New York: Springer.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*(6), 1086–1120.

Astrachan, C. B., Patel, V. K., & Wanzenried, G. (2014). A comparative study of CB-SEM and PLS-SEM for theory development in family firm research. *Journal of Family Business Strategy*, *5*(1), 116–128.

Atinc, G., Simmering, M. J., & Kroll, M. J. (2012). Control variable use and reporting in macro and micro-management research. *Organizational Research Methods*, *15*(1), 57–74.

Bandalos, D. L., & Boehm-Kaufman, M. R. (2009). Four common misconceptions in exploratory factor analysis. In C. E. Lance, & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 61–87). New York, NY: Routledge.

Barroso, C., & Picón, A. (2012). Multi-dimensional analysis of perceived switching costs. *Industrial Marketing Management*, *41*(3), 531–543.

Barroso, C., Cepeda Carrión, G., & Roldán, J. L. (2010). Applying maximum likelihood and PLS on different sample sizes: Studies on SERVQUAL model and employee behavior model. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares* (pp. 427–447). Springer Berlin Heidelberg.

Baxter, R. A. (2009). Reflective and formative metrics of relationship value: A commentary essay. *Journal of Business Research*, *62*(12), 1370–1377.

Bearden, W. O., Netemeyer, R. G., & Haws, K. L. (2011). *Handbook of marketing scales. Multi-item measures of marketing and consumer behavior research* (3rd ed.). Sage.

Becker, J. -M., Rai, A., & Rigdon, E. E. (2013). *Predictive validity and formative measurement in structural equation modeling: Embracing practical relevance. Proceedings of the 34th International Conference on Information Systems*, Milan, Italy.

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

Bentler, P. M., & Huang, W. (2014). On components, latent variables, PLS and simple methods: Reactions to Rigdon's rethinking of PLS. *Long Range Planning*, *47*(3), 138–145.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, *53*(1), 605–634.

Bollen, K. A. (2011). Evaluating effect, composite, and causal indicators in structural equation models. *MIS Quarterly*, *35*(2), 359–372.

Bollen, K. A., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods*, *16*(3), 265–284.

Bollen, K. A., & Davies, W. R. (2009). Causal indicator models: Identification, estimation, and testing. *Structural Equation Modeling*, *16*(3), 498–522.

Bollen, K. A., & Diamantopoulos, A. (2016). In defense of causal-formative indicators: A minority report. *Psychological Methods* (forthcoming).

Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *110*(2), 305–314.

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future* (pp. 139–168). Chicago: Scientific Software International.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203–219.

Bruner, G. C., James, K. E., & Hensel, P. J. (2001). *Marketing scales handbook: A compilation of multi-item measures: Butterworth-Heinemann.*

Chang, W., Franke, G. R., & Lee, N. (2016). Comparing reflective and formative measures: New insights from relevant simulations. *Journal of Business Research* (in press).

Chin, W. W. (1998). Commentary: Issues and opinion on structural equation modeling. *MIS Quarterly*, *22*(1), xii–xvi.

Chin, W. W. (2010). How to write up and report PLS analyses. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications (Springer handbooks of computational statistics series, vol. II)* (pp. 655–690). Heidelberg, Dordrecht, London, New York: Springer.

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

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, *18*(1), 115–126.

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research*, *61*(12), 1250–1262.

Cudeck, R., & MacCallum, R. C. (Eds.). (2007). *Factor analysis at 100: Historical developments and future directions*. Mahwah, NJ: Lawrence Erlbaum Associates.

Diamantopoulos, A. (1994). Modelling with LISREL: A guide for the uninitiated. *Journal of Marketing Management*, *10*(1–3), 105–136.

Diamantopoulos, A. (2005). The C-OAR-SE procedure for scale development in marketing: A comment. *International Journal of Research in Marketing*, *22*(1), 1–9.

Diamantopoulos, A. (2006). The error term in formative measurement models: Interpretation and modeling implications. *Journal of Modelling in Management*, *1*(1), 7–17.

Diamantopoulos, A. (2011). Incorporating formative measures into covariance-based structural equation models. *MIS Quarterly*, *35*(2), 335–358.

Diamantopoulos, A., & Riefler, P. (2011). Using formative measures in international marketing models: A cautionary tale using consumer animosity as an example. *Advances in International Marketing*, *10*(22), 11–30.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research*, *38*(2), 269–277.

Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of Business Research*, *61*(12), 1203–1218.

Dijkstra, T. K., & Henseler, J. (2011). Linear indices in nonlinear structural equation models: Best fitting proper indices and other composites. *Quality & Quantity*, *45*(6), 1505–1518.

Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly*, *39*(2), 297–316.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*(2), 155–174.

Ehrenberg, A. S. C. (1962). Some questions about factor analysis. *Journal of the Royal Statistical Society Series D (The Statistician)*, *12*(3), 191–208.

Fornell, C. G., & 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.

Fox, J., Nie, Z., Byrnes, J., Culbertson, M., DebRoy, S., Friendly, M., ... Monette, G. (2015). R package sem: Structural equation models (version 3.1-6). http://cran.r-project.org/web/packages/sem/

Gelhard, C., & von Delft, S. (2016). The role of organizational capabilities in achieving superior sustainability performance. *Journal of Business Research* (forthcoming).

Gilliam, D. A., & Voss, K. (2013). A proposed procedure for construct definition in marketing. *European Journal of Marketing*, *47*(1/2), 5–26.

Goodhue, D. L., Lewis, W., & Thompson, R. (2012). Does PLS have advantages for small sample size or non-normal data? *MIS Quarterly*, *36*(3), 981–1001.

Grace, J. B., & Bollen, K. A. (2008). Representing general theoretical concepts in structural equation models: The role of composite variables. *Environmental and Ecological Statistics*, *15*(2), 191–213.

Guide, V. D. R., & Ketokivi, M. (2015). Notes from the editors: Redefining some methodological criteria for the journal. *Journal of Operations Management*, *37*, v–viii (July 2015).

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–151.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, *40*(3), 414–433.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Thousand Oaks, CA: Sage.

Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., ... Calantone, R. J. (2014). Common beliefs and reality about partial least squares: Comments on Rönkkö & Evermann (2013). *Organizational Research Methods*, *17*(2), 182–209.

Henseler, J., Hubona, G. S., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems*, *116*(1), 1–19.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using partial least squares. *International Marketing Review*, *33*(3), 405–431.

Howell, R. D., Breivik, E., & Wilcox, J. B. (2013). Formative measurement: A critical perspective. *The DATA BASE for Advances in Information Systems*, *44*(4), 44–55.

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

Hui, B. S., & Wold, H. O. A. (1982). Consistency and consistency at large of partial least squares estimates. In K. G. Jöreskog, & H. O. A. Wold (Eds.), *Systems under indirect observation, part II* (pp. 119–130). Amsterdam: North Holland.

Hwang, H., Malhotra, N. K., Kim, Y., Tomiuk, M. A., & Hong, S. (2010). A comparative study on parameter recovery of three approaches to structural equation modeling. *Journal of Marketing Research*, *47*(4), 699–712.

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.

Jones, M. A., Mothersbaugh, D. L., & Beatty, S. E. (2000). Switching barriers and repurchase intentions in services. *Journal of Retailing*, *76*(2), 259–274.

Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, *43*(4), 443–477.

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., & Wold, H. O. A. (1982). The ML and PLS techniques for modeling with latent variables: Historical and comparative aspects. In H. O. A. Wold, & K. G. Jöreskog (Eds.), *Systems under indirect observation, part I* (pp. 263–270). Amsterdam: North-Holland.

Kaufmann, L., & Gaeckler, J. (2015). A Structured Review of Partial Least Squares in Supply Chain Management Research. *Journal of Purchasing and Supply Management* (forthcoming).

Law, K. S., & Wong, C. -S. (1999). Multidimensional constructs in structural equation analysis: An illustration using the job perception and job satisfaction constructs. *Journal of Management*, *25*(2), 143–160.

Law, K. S., Wong, C. -S., & Mobley, W. H. (1998). Toward a taxonomy of multidimensional constructs. *The Academy of Management Review*, *23*(4), 741–755.

Lee, N., & Cadogan, J. W. (2013). Problems with formative and higher-order reflective variables. *Journal of Business Research*, *66*(2), 242–247.

Lohmöller, J. -B. (1989). *Latent variable path modeling with partial least squares*. Heidelberg: Physica.

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., & Cai, L. (2007). Factor analysis models as approximations: Some history and some implications. In R. Cudeck, & R. C. MacCallum (Eds.), *Factor analysis at 100: Historical developments and future directions* (pp. 153–175). Mahwah, NJ: Laurence Erlbaum Associates.

MacKenzie, S. B. (2003). The dangers of poor construct conceptualization. *Journal of the Academy of Marketing Science*, *31*(3), 323–326.

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, *35*(2), 293–A295.

Marcoulides, G. A., & Chin, W. W. (2013). You write, but others read: Common methodological misunderstandings in PLS and related methods. In H. Abdi, W. W. Chin, V. Esposito Vinzi, G. Russolillo, & L. Trinchera (Eds.), *New Perspectives in Partial Least Squares and Related Methods. Vol. 56*. (pp. 31–64). New York: Springer.

Marcoulides, G. A., & Saunders, C. (2006). PLS: A silver bullet? *MIS Quarterly*, *30*(2), III–IIX.

Marcoulides, G. A., Chin, W. W., & Saunders, C. (2012). When imprecise statistical statements become problematic: A response to Goodhue, Lewis, and Thompson. *MIS Quarterly*, *36*(3), 717–728.

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

McDonald, R. P. (1996). Path analysis with composite variables. *Multivariate Behavioral Research*, *31*(2), 239–270.

Michell, J. (2013). Constructs, inferences and mental measurement. *New Ideas in Psychology*, *31*(1), 13–21.

Monecke, A., & Leisch, F. (2012). semPLS: Structural equation modeling using partial least squares. *Journal of Statistical Software*, *48*(3), 1–32.

Nunnally, J. C., & Bernstein, I. (1994). *Psychometric theory* (3 ed.). New York: McGraw Hill.

Pedhazur, E. J., & Pedhazur Schmelkin, L. (1991). *Measurement, design, and analysis: An integrated approach*. Mahwah, NJ: Lawrence Erlbaum.

Peng, D. X., & Lai, F. (2012). Using partial least squares in operations management research: A practical guideline and summary of past research. *Journal of Operations Management*, *30*(6), 467–480.

R Core Team (2014). *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. J., Haenlein, M., & Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of Research in Marketing*, *26*(4), 332–344.

Rigdon, E. E. (1998). Structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 251–294). Mahwah: Erlbaum.

Rigdon, E. E. (2012). Rethinking partial least squares path modeling: In praise of simple methods. *Long Range Planning*, *45*(5–6), 341–358.

Rigdon, E. E. (2013). Partial least squares path modeling. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course, seconded. Vol. 1*. (pp. 81–116). Charlotte: Information Age.
Rigdon, E. E. (2014). Rethinking partial least squares path modeling: Breaking chains and forging ahead. *Long Range Planning*, *47*(3), 161–167.
Rigdon, E. E. (2016). Choosing PLS path modeling as analytical method in European management research: A realist perspective. *European Management Journal* (forthcoming).
Rigdon, E. E., Preacher, K. J., Lee, N., Howell, R. D., Franke, G. R., & Borsboom, D. (2011). Overcoming measurement dogma: A response to Rossiter. *European Journal of Marketing*, *45*(11/12), 1589–1600.
Rigdon, E. E., Becker, J. M., Rai, A., Ringle, C. M., Diamantopoulos, A., Karahanna, E., & Dijkstra, T. K. (2014). Conflating antecedents and formative indicators: A comment on Aguirre-Urreta and Marakas, *25*(4), 780–784.
Ringle, C. M., Sarstedt, M., & Straub, D. W. (2012). A critical look at the use of PLS-SEM in MIS quarterly. *MIS Quarterly*, *36*(1), iii–xiv.
Rönkkö, M. (2016). R package matrixpls: Matrix-based partial least squares estimation (version 0.7.0). http://cran.r-project.org/web/packages/matrixpls/
Rönkkö, M., & Evermann, J. (2013). A critical examination of common beliefs about partial least squares path modeling. *Organizational Research Methods*, *16*(3), 425–448.
Rönkkö, M., McIntosh, C. N., & Antonakis, J. (2015). On the adoption of partial least squares in psychological research: Caveat emptor. *Personality and Individual Differences*, *87*, 76–84.
Rönkkö, M., McIntosh, C. N., Antonakis, J., & Edwards, J. R. (2016). Partial least squares path modeling: Time for some serious second thoughts. *Journal of Operations Management* (forthcoming).
Rossiter, J. R. (2011). Marketing measurement revolution: The C-OAR-SE method and why it must replace psychometrics. *European Journal of Marketing*, *45*(11/12), 1561–1588.
Salzberger, T., Sarstedt, M., & Diamantopoulos, A. (2016). Measurement in the social sciences: Where *C*-OAR-SE delivers and where it does not. *European Journal of Marketing* (forthcoming).
Sarstedt, M., Wilczynski, P., & Melewar, T. C. (2013). Measuring reputation in global markets — A comparison of reputation measures' convergent and criterion validities. *Journal of World Business*, *48*(3), 329–339.
Sarstedt, M., Ringle, C. M., & Hair, J. F. (2014). PLS-SEM: Looking back and moving forward. *Long Range Planning*, *47*(3), 132–137.
Sarstedt, M., Ringle, C. M., Henseler, J., & Hair, J. F. (2014). On the emancipation of PLS-SEM: A commentary on Rigdon (2012). *Long Range Planning*, *47*(3), 154–160.
Schlittgen, R. (2015). R package SEGIRLS (version 0.5). http://www.wiso.uni-hamburg.de/fileadmin/bwl/statistikundoekonometrie/Schlittgen/SoftwareUndDaten/SEGIRLS_0.5.tar.gz
Schneeweiß, H. (1991). Models with latent variables: LISREL versus PLS. *Statistica Neerlandica*, *45*(2), 145–157.
Schönemann, P. H., & Wang, M. -M. (1972). Some new results on factor indeterminacy. *Psychometrika*, *37*(1), 61–91.
Spearman, C. (1927). *The abilities of man*. London: MacMillan.
Temme, D., Diamantopoulos, A., & Pfegfeidel, V. (2014). Specifying formatively-measured constructs in endogenous positions in structural equation models: Caveats and guidelines for researchers. *International Journal of Research in Marketing*, *31*(3), 309–316.
Tenenhaus, A., & Tenenhaus, M. (2011). Regularized generalized canonical correlation analysis. *Psychometrika*, *76*(2), 257–284.
Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y. -M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics & Data Analysis*, *48*(1), 159–205.
Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago, IL: The University of Chicago Press.
Wickens, M. R. (1972). A note on the use of proxy variables. *Econometrica*, *40*(4), 759–761.
Wold, H. O. A. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog, & H. O. A. Wold (Eds.), *Systems under indirect observations: Part II* (pp. 1–54). Amsterdam: North-Holland.