

An assessment of the use of partial least squares structural equation modeling in marketing research

Joe F. Hair · Marko Sarstedt · Christian M. Ringle ·
Jeannette A. Mena

Received: 23 December 2010 / Accepted: 17 May 2011 / Published online: 7 June 2011
© Academy of Marketing Science 2011

Abstract Most methodological fields undertake regular critical reflections to ensure rigorous research and publication practices, and, consequently, acceptance in their domain. Interestingly, relatively little attention has been paid to assessing the use of partial least squares structural equation modeling (PLS-SEM) in marketing research—

despite its increasing popularity in recent years. To fill this gap, we conducted an extensive search in the 30 top ranked marketing journals that allowed us to identify 204 PLS-SEM applications published in a 30-year period (1981 to 2010). A critical analysis of these articles addresses, amongst others, the following key methodological issues: reasons for using PLS-SEM, data and model characteristics, outer and inner model evaluations, and reporting. We also give an overview of the interdependencies between researchers' choices, identify potential problem areas, and discuss their implications. On the basis of our findings, we provide comprehensive guidelines to aid researchers in avoiding common pitfalls in PLS-SEM use. This study is important for researchers and practitioners, as PLS-SEM requires several critical choices that, if not made correctly, can lead to improper findings, interpretations, and conclusions.

J. F. Hair (✉)
Kennesaw State University,
KSU Center, DBA Program,
Kennesaw, GA 30144, USA
e-mail: jhair3@kennesaw.edu

M. Sarstedt
Ludwig-Maximilians-University Munich (LMU),
Institute for Market-based Management (IMM),
Kaulbachstr. 45,
80539 Munich, Germany
e-mail: sarstedt@bwl.lmu.de

M. Sarstedt · C. M. Ringle
University of Newcastle,
Faculty of Business and Law,
Newcastle, Australia

C. M. Ringle
Hamburg University of Technology (TUHH),
Institute for Human Resource Management
and Organizations (HRMO),
Schwarzenbergstraße 95 (D),
21073 Hamburg, Germany
e-mail: hrmo@tuhh.de

J. A. Mena
University of South Florida, College of Business,
4202 E. Fowler Avenue,
Tampa, FL 33620-5500, USA
e-mail: jeannette_mena@yahoo.com

Keywords Empirical research methods · Partial least squares · Path modeling · Structural equation modeling

Introduction

Structural equation modeling (SEM) has become a quasi-standard in marketing research (e.g., Babin et al. 2008; Bagozzi 1994; Hulland 1999), as it allows authors to test complete theories and concepts (Rigdon 1998). Researchers especially appreciate SEM's ability to assess latent variables at the observation level (outer or measurement model) and test relationships between latent variables on the theoretical level (inner or structural model)

(Bollen 1989).¹ When applying SEM, researchers must consider two types of methods: covariance-based techniques (CB-SEM; Jöreskog 1978, 1993) and variance-based partial least squares (PLS-SEM; Lohmöller 1989; Wold 1982, 1985). Although both methods share the same roots (Jöreskog and Wold 1982), previous marketing research has focused primarily on CB-SEM (e.g., Bagozzi 1994; Baumgartner and Homburg 1996; Steenkamp and Baumgartner 2000).

Recently, PLS-SEM application has expanded in marketing research and practice with the recognition that PLS-SEM's distinctive methodological features make it a possible alternative to the more popular CB-SEM approaches (Henseler et al. 2009). A variety of PLS-SEM enhancements have been developed in recent years, including (1) confirmatory tetrad analysis for PLS-SEM to empirically test a construct's measurement mode (Gudergan et al. 2008); (2) impact-performance matrix analysis (Slack 1994; Völckner et al. 2010); (3) response-based segmentation techniques, such as finite mixture partial least squares (FIMIX-PLS; Hahn et al. 2002; Sarstedt et al. 2011a); (4) guidelines for analyzing moderating effects (Henseler and Chin 2010; Henseler and Fassott 2010); (5) non-linear effects (Rigdon et al. 2010); and (6) hierarchical component models (Lohmöller 1989; Wetzels et al. 2009). These enhancements expand PLS-SEM's general usefulness as a research tool in marketing and the social sciences.

Most methodological fields have established regular critical reflections to ensure rigorous research and publication practices, and consequently acceptance, in their domain. While reviews of CB-SEM usage have a long tradition across virtually all business disciplines (e.g., Babin et al. 2008; Baumgartner and Homburg 1996; Brannick 1995; Garver and Mentzer 1999; Medsker et al. 1994; Shah and Goldstein 2006; Shook et al. 2004; Steenkamp and van Trijp 1991), relatively little attention has been paid to assessing PLS-SEM use. Hulland (1999) was the first to review PLS-SEM use with an in-depth analysis of four strategic management studies. His review revealed flaws in the PLS-SEM method's application, raising questions about its appropriate use in general, and indicating the need for further examination in a more comprehensive assessment. More recently, Henseler et al. (2009) and Reinartz et al. (2009) assessed PLS-SEM use in (international) marketing research but focused only on the reasons for choosing this method.

¹ It is important to note that the SEM-related literature does not always employ the same terminology when referring to elements of the model. Publications addressing CB-SEM (e.g., Hair et al. 2010) often refer to "structural model" and "measurement models," whereas those focusing on PLS-SEM (e.g., Lohmöller 1989) use the terms "inner model" and "outer models." As this paper deals with PLS-SEM, related terminology is used.

PLS-SEM requires several choices that, if not made correctly, can lead to improper findings, interpretations, and conclusions. In light of its current more widespread application in marketing research and practice, a critical review of PLS-SEM's use seems timely and warranted. Our objectives with this review are threefold: (1) to investigate published PLS-SEM articles in terms of relevant criteria, such as sample size, number of indicators used, and measures reported; (2) to provide an overview of the interdependencies in researchers' choices, identify potential problem areas, and discuss their implications; and (3) to provide guidance on preventing common pitfalls in using PLS-SEM.

Our review shows that PLS-SEM's methodological properties are widely misunderstood, at times leading to the technique's misapplication, even in top tier journals. Furthermore, researchers often do not apply criteria available for model assessment, sometimes even misapplying the measures. Our guidelines on applying PLS-SEM appropriately have important implications, therefore, for correct application and knowledgeable assessment of PLS-SEM-related research studies.

Not "CB-SEM versus PLS-SEM" but "CB-SEM and PLS-SEM"

Wold (1975) originally developed PLS-SEM under the name NIPALS (nonlinear iterative partial least squares), and Lohmöller (1989) extended it. PLS-SEM was developed as an alternative to CB-SEM that would emphasize prediction while simultaneously relaxing the demands on data and specification of relationships (e.g., Dijkstra 2010; Jöreskog and Wold 1982). The methodological concepts underlying both approaches have been compared in several publications, including those by Barclay et al. (1995), Chin and Newsted (1999), Fornell and Bookstein (1982), Gefen et al. (2011), Hair et al. (2011), Jöreskog and Wold (1982), and Lohmöller (1989).

CB-SEM estimates model parameters so that the discrepancy between the estimated and sample covariance matrices is minimized. In contrast, PLS-SEM maximizes the explained variance of the endogenous latent variables by estimating partial model relationships in an iterative sequence of ordinary least squares (OLS) regressions. An important characteristic of PLS-SEM is that it estimates latent variable scores as exact linear combinations of their associated manifest variables (Fornell and Bookstein 1982) and treats them as perfect substitutes for the manifest variables. The scores thus capture the variance that is useful for explaining the endogenous latent variable(s). Estimating models via a series of OLS regressions implies that PLS-SEM relaxes the assumption of multivariate normality needed for maximum likelihood-based SEM estimations

(Fornell and Bookstein 1982; Hwang et al. 2010; Lohmöller 1989; Wold 1982; for a discussion, see Dijkstra 2010). In this context, Lohmöller (1989, p. 64) notes that “it is not the concepts nor the models nor the estimation techniques which are ‘soft,’ only the distributional assumptions.” Furthermore, since PLS-SEM is based on a series of OLS regressions, it has minimum demands regarding sample size and generally achieves high levels of statistical power (Reinartz et al. 2009). Conversely, CB-SEM involves constraints regarding the number of observations and small sample sizes, often leading to biased test statistics (e.g., Hu and Bentler 1995), inadmissible solutions (e.g., Heywood cases), and identification problems—especially in complex model set-ups (e.g., Chin and Newsted 1999). Thus, PLS-SEM is suitable for applications where strong assumptions cannot be fully met and is often referred to as a distribution-free “soft modeling approach.”

Consideration of formative and reflective outer model modes is an important issue for SEM (e.g., Diamantopoulos and Winklhofer 2001; Jarvis et al. 2003). While CB-SEM is applicable for formative outer model specifications only under certain conditions (e.g., Bollen and Davies 2009; Diamantopoulos and Riefler 2011), PLS-SEM can almost unrestrictedly handle both reflective and formative measures (e.g., Chin 1998). Furthermore, PLS-SEM is not constrained by identification concerns, even if models become complex, a situation that typically restricts CB-SEM usage (Hair et al. 2011).

These advantages must be considered, however, in light of several disadvantages. For example, the absence of a global optimization criterion implies a lack of measures for overall model fit. This issue limits PLS-SEM’s usefulness for theory testing and for comparing alternative model structures. As PLS-SEM also does not impose any distributional assumptions, researchers cannot rely on the classic inferential framework and thus have to revert to prediction-oriented, non-parametric evaluation criteria as well as resampling procedures to evaluate the partial model structures’ adequacy (e.g., Chin 2010). A further concern is that PLS-SEM parameter estimates are not optimal regarding bias and consistency (Reinartz et al. 2009)—a characteristic frequently referred to as PLS-SEM bias. This bias is more severe in very complex models, since least squares estimators do not control the contingent and chained effects of one part of the model’s errors to another. Only when the number of observations *and* the number of indicators per latent variable increase to infinity do the latent variable scores (and therefore the parameter estimates) approach the true values. Thus, estimates are asymptotically correct in the qualified sense of consistency at large (Jöreskog and Wold 1982; Lohmöller 1989; Wold 1982).

In light of their distinct statistical concepts, most researchers consider the two approaches to SEM as

complementary, a fact that has already been stressed by the two originators of PLS-SEM and CB-SEM, Jöreskog and Wold (1982). In general, PLS-SEM’s weaknesses are CB-SEM’s strengths, and vice versa (e.g., Jöreskog and Wold 1982; Sosik et al. 2009). Thus, neither of the SEM methods is generally superior to the other. Instead, researchers need to apply the SEM technique that best suits their research objective, data characteristics, and model set-up (e.g., Fornell and Bookstein 1982; Gefen et al. 2011; Reinartz et al. 2009).

Review of PLS-SEM research

Our review of PLS-SEM applications in marketing consists of studies published in the top 30 marketing journals identified in Hult et al.’s (2009) journal ranking. This ranking is similar to those by Baumgartner and Pieters (2003), Hult et al. (1997), and Theoharakis and Hirst (2002). All studies published in the 30-year period from 1981 to 2010 were searched for empirical PLS-SEM applications in the field of marketing. We conducted a full text search in the *Thomson Reuters Web of Knowledge*, *ProQuest ABI/INFORM Global*, and *EBSCO Business Source Premier* databases, using the keywords “partial least squares” and “PLS.” We also looked into the online versions of the journals.² The search across multiple databases using the same keywords allowed us to verify that we had captured all PLS-SEM articles in the targeted marketing journals.

Since Hult et al.’s (2009) ranking includes interdisciplinary journals that cover several functional business areas (e.g., *Management Science*, *Journal of Business Research*, *Journal of International Business Studies*), articles were screened to identify those in marketing. Papers that drew on PLS-SEM simulation studies and empirical PLS-SEM applications to illustrate methodological enhancements were not considered. Similarly, studies applying path analysis, conventional score-based latent variable regression, and PLS regression were excluded from the sample. Ultimately, a total of 24 journals with relevant articles remained (Table 1). The paper selection process included an initial coding by a senior PhD student. Thereafter, two professors proficient in the technique coded each article independently. The coding agreement on the relevant articles was 92%, which compares well with Shook et al.’s (2004) study. To resolve coding inconsistencies, opinions of other experts were obtained.

This search resulted in 204 articles (Table 1) with 311 PLS-SEM estimations, since some articles analyze alternative models and/or use different datasets (e.g., collected in

² The search process was completed on January 31, 2011.

Table 1 PLS-SEM studies in the top 30 marketing journals*Advances in Consumer Research*¹

Guiot 2000
 Mayo and Qualls 1987
 Qualls 1988
 Rego 1998
 Schramm-Klein, Morschett and Swoboda 2007
 Vanhamme and Snelders 2003
 Zinkhan and Fornell 1989
 Zinkhan and Muderrisoglu 1985

European Journal of Marketing
 Antioco and Kleijnen 2010
 Ball, Coelho and Machás 2004
 Baumgarth 2010
 Bruhn 2003
 Chan, Hui, Lo, Tse, Tso and Wu 2003
 Chung 2009
 Chung 2010
 Ekinci, Dawes and Massey 2008
 Guenzi and Georges 2010
 Ha, Janda and Muthaly 2010
 Jayawardhena, Kuckertz, Karjaluoto and Kautonen 2009
 Jensen 2008
 King and Grace 2010
 Klemz, Boshoff and Mazibuko 2006
 Klemz and Boshoff 2001
 Lee 2000
 Massey and Dawes 2007
 Massey and Kyriazis 2007
 O’Cass 2004
 O’Cass and Julian 2003
 O’Cass and Ngo 2007
 O’Cass and Weerawardena 2009
 Ormrod and Henneberg 2010
 Petruzzellis 2010
 Plouffe 2008
 Tortosa, Moliner and Sánchez 2009
 Ulaga and Eggert 2006
 Vandenbosch 1996
 Vlachos, Theotokis, Pramataris and Vrechopoulos 2010
 Voola and O’Cass 2010

Industrial Marketing Management

Baumgarth and Schmidt 2010
 Davis, Golicic and Marquardt 2008
 Dawes, Lee and Midgley 2007
 Eggert and Ulaga 2010
 Eggert, Ulaga and Schultz 2006
 Guenzi, Georges and Pardo 2009
 Guenzi, Pardo and Georges 2007

Katrichis and Ryan 1998
 Lee, Park, Baek and Lee 2008
 Liu, Li, and Xue 2010
 Massey and Dawes 2007
 Menguc and Auh 2008
 Ngo and O’Cass 2009
 O’Cass and Weerawardena 2010
 Plouffe, Sridharan and Barclay 2010
 Rangarajan, Jones and Chin 2005
 Real, Leal and Roldán 2006
 Rodríguez-Pinto, Rodríguez-Escudero and Gutiérrez-Cillán 2008
 van Riel, de Mortanges and Streukens 2005
 Vlachos, Theotokis and Panagopoulos 2010
 Wagner 2010
 Wittmann, Hunt and Arnett 2009
 Yi and Gong 2008

International Journal of Research in Marketing

Becker, Greve and Albers 2009
 Calantone, Graham and Mintu-Wimsatt 1998
 Frenzen, Hansen, Krafft, Mantrala and Schmidt 2010
 Johnson, Lehmann, Fornell and Horne 1992
 Rapp, Ahearne, Mathieu and Rapp 2010
 Reimann, Schilke and Thomas 2010
 Sattler, Völckner, Riediger and Ringle 2010
 Wilken, Cornelißen, Backhaus and Schmitz 2010

International Marketing Review

Alpert, Kamins, Sakano, Onzo and Graham 2001
 Duque and Lado 2010
 Singh, Fassott, Chao and Hoffmann 2006

Journal of Advertising

O’Cass 2002
 Okazaki, Li and Hirose 2009
 Okazaki, Mueller and Taylor 2010

Journal of Advertising Research

Drengner, Gaus and Jahn 2008
 Jagpal 1981
 Martensen, Gronholdt, Bendtsen and Jensen 2007
 Schwaiger, Sarstedt and Taylor 2010

Journal of Business Research

Bruhn, Georgi and Hadwich 2008
 Cepeda and Vera 2007
 Heslop, Papadopoulos, Dowdles, Wall and Compeau 2004
 Huber, Vollhardt, Matthes and Vogel 2010
 Ledden, Kalafatis and Samouel 2007
 Lee 1994
 Lee 2001
 MacMillan, Money, Money and Downing 2005

- Menguc, Auh and Shih 2007
 O’Cass and Ngo 2007
 O’Cass and Pecotich 2005
 Okazaki and Taylor 2008
 Rapp, Trainor and Agnihotri 2010
 Wagner, Eggert and Lindemann 2010
 Weerawardena, O’Cass and Julian 2006
- Journal of Consumer Psychology*
 Nelson 2004
- Journal of Consumer Research*
 Cotte and Wood 2004
 Fornell and Robinson 1983
 Johnson and Fornell 1987
 Mathwick, Wiertz and de Ruyter 2008
 Qualls 1987
- Journal of Interactive Marketing*
 Sääksjärvi and Samiee 2007
 Sanchez-Franco 2009
- Journal of International Business Studies*
 Money and Graham 1999
 Shi, White, Zou and Cavusgil 2010
- Journal of International Marketing*
 Brettel, Engelen, Heinemann and Vadhanasindhu 2008
 Holzmüller and Stöttinger 1996
 Lages, Silva and Styles 2009
 Lee and Dawes 2005
 Navarro, Acedo, Robson, Ruzo and Losada 2010
 Nijssen and Douglas 2008
 Nijssen and van Herk 2009
 Rodríguez and Wilson 2002
 Sichtmann and von Selasinsky 2010
 Tsang, Nguyen and Erramilli 2004
- Journal of Marketing*
 Alpert, Kamins and Graham 1992
 Dawes, Lee and Dowling 1998
 Dellande, Gilly and Graham 2004
 Ernst, Hoyer and Rübсаamen 2010
 Fornell 1992
 Fornell, Johnson, Anderson, Jaesung and Bryant 1996
 Green, Barclay and Ryans 1995
 Hennig-Thurau, Groth, Paul and Gremler 2006
 Hennig-Thurau, Henning and Sattler 2007
 Johnson, Herrmann and Huber 2006
 Lam, Ahearne, Hu and Schillewaert 2010
 McFarland, Bloodgood and Payan 2008
 Smith and Barclay 1997
 Ulaga and Eggert 2006
- Wagner, Hennig-Thurau and Rudolph 2009
 White, Varadarajan and Dacin 2003
 Wuyts and Geyskens 2005
- Journal of Marketing Management*
 Auh, Salisbury and Johnson 2003
 Lings and Owen 2007
 Nakata and Zhu 2006
 Rose and Samouel 2009
 Stewart 1997
 Wolk and Theysohn 2007
- Journal of Marketing Research*
 Ahearne, Mackenzie, Podsakoff, Mathieu and Lam 2010
 Ahearne, Rapp, Hughes and Jindal 2010
 Barclay 1991
 Fornell and Bookstein 1982
 Reinartz, Krafft and Hoyer 2004
 Zinkhan, Joachimsthaler and Kinnear 1987
- Journal of Product Innovation Management*
 Akgün, Keskin and Byrne 2010
 Antioco, Moenaert and Lindgreen 2008
 de Brentani, Kleinschmidt and Salomo 2010
 De Luca, Verona and Vicari 2010
 Kaplan, Schoder and Haenlein 2007
 Moenaert, Robben, Antioco
 De Schamphelaere and Roks 2010
 Persaud 2005
 Plouffe, Vandenbosch and Hulland 2001
 Salomo, Keinschmidt and de Brentani 2010
 Salomo, Talke and Strecker 2008
 Wynstra, von Corswant and Wetzels 2010
- Journal of Public Policy and Marketing*
 Martin and Johnson 2010
- Journal of Retailing*
 Arnett, Laverie and Meiers 2003
 Kleijnen, de Ruyter and Wetzels 2007
 Lei, de Ruyter and Wetzels 2008
 Mathwick, Wagner and Unni 2010
 Richardson and Jain 1996
 Sirohi, McLaughlin and Wittink 1998
- Journal of Service Research*
 Chen, Tsou and Huang 2009
 Cooil, Aksoy, Keiningham and Maryott 2009
 Daryanto, de Ruyter and Wetzels 2010
 Eisingerich, Rubera and Seifert 2009
 Olsen and Johnson 2003
 Storey and Kahn 2010
 Völckner, Sattler, Hennig-Thurau and Ringle 2010

Journal of the Academy of Marketing Science
 Antioco, Moenaert, Feinberg and Wetzels 2008
 Antioco, Moenaert, Lindgreen and Wetzels 2008
 Davis and Golicic 2010
 Gelbrich 2010
 Grégoire, Laufer and Tripp 2010
 Gilly, Graham, Wolfenbarger and Yale 1998
 Harmancioglu, Droge and Calantone 2009
 Hennig-Thurau, Houston and Walsh 2006
 Mintu-Wimsatt and Graham 2004
 Reimann, Schilke and Thomas 2010
 Sarkar, Echambadi, Cavusgil and Aulakh 2001
 Slotegraaf and Dickson 2004
 Sundaram, Schwarz, Jones and Chin 2007

Management Science

Ahuja, Galleta and Carley 2003
 Fornell, Lorange and Roos 1990
 Fornell, Robinson and Wernerfelt 1985
 Graham, Mintu and Rodgers 1994
 Gray and Meister 2004
 Im and Rai 2008
 Mitchell and Nault 2007
 Venkatesh and Agarwal 2006
 Xu, Venkatesh, Tam and Hong 2010

Marketing Letters

Burke 1996
 Das, Echambadi, McCardle and Luckett 2003
 Ettlé and Johnson 1994
 Grégoire and Fisher 2006
 Mooradian, Matzler and Szykman 2008

Psychology and Marketing

Bodey and Grace 2007
 Fisher and Grégoire 2006
 Guiot 2001
 Johnson and Horne 1992
 O’Cass and Grace 2008
 Smith 1998
 Smith and Bristor 1994
 Yi and Gong 2008a
 Yi and Gong 2008b

Harvard Business Review, California Management Review, Journal of Business, Quantitative Marketing and Economics, Marketing Science, and Sloan Management Review did not produce any relevant articles.

¹ We excluded studies by Lee et al. (2009), as well as Lawson (2010) published in *Advances in Consumer Research* as these were only published as extended abstracts.

different years and/or countries). The *European Journal of Marketing* (30 articles, 14.71%), *Industrial Marketing Management* (23 articles, 11.27%), and *Journal of Marketing* (17 articles, 8.33%) published the highest numbers of PLS-SEM studies. To assess the growth trend in PLS-SEM use, we conducted a time-series analysis with the number of studies applying this method as the dependent variable. Model estimation using a linear term results in a significant model ($F=25.01$, $p \leq 0.01$) in which the time effect is significant ($t=5.00$, $p \leq 0.01$). Next, we used both linear and quadratic time effects. The regression model is significant ($F=35.13$, $p \leq 0.01$) and indicates that 72.22% of PLS-SEM applications are explained by linear and quadratic time effects. The additionally considered quadratic effect is significant ($t=4.94$, $p \leq 0.01$), indicating that the use of PLS-SEM in marketing has accelerated over time. Clearly, PLS-SEM has gained popularity over the past decades—most notably, 51 studies appeared in 2010 alone. In contrast, Baumgartner and Homburg’s (1996) analysis of early CB-SEM diffusion provided no indication of its accelerated use prior to their analysis.

Critical issues in PLS-SEM applications

Each article was evaluated according to a wide range of criteria which allow PLS-SEM’s critical issues and common misapplications to be identified. The following review focuses on six key issues in PLS-SEM: (1) reasons for using PLS-SEM; (2) data characteristics; (3) model characteristics; (4) outer model evaluation; (5) inner model evaluation; and (6) reporting. Where possible, we also indicate best practices as guidelines for future applications and suggest avenues for further research. In addition to the PLS-SEM application analysis, we also contrast two time periods to assess whether the usage has changed. In light of Chin’s (1998), Chin and Newsted’s (1999) as well as Hulland’s (1999) influential PLS-SEM articles, published in the late 1990s, we differentiate between studies published before 2000 (39 studies with 69 models) and those published in 2000 and beyond (165 studies with 242 models).³

Reasons for using PLS-SEM

Since predominantly covariance-based SEM techniques have been used to estimate models in marketing, PLS-SEM use often requires a more detailed explanation of the rationale for selecting this method (Chin 2010). The most often used reasons relate to data characteristics, such as the

³ In the following, we consistently use the term “studies” when referring to the 204 journal articles and the term “models” when referring to the 311 PLS-SEM applications in these articles.

analysis of non-normal data (102 studies, 50.00%), small sample sizes (94 studies, 46.08%), and the formative measurement of latent variables (67 studies, 32.84%). Furthermore, researchers stated that applying PLS-SEM is more consistent with their study objective. For example, in 57 studies (27.94%), authors indicated their primary research objective is to explain the variance of the endogenous constructs. This is closely related to the rationale of using the method for exploratory research and theory development, which 35 studies (17.16%) mentioned. The latter two reasons comply with PLS-SEM's original purpose of prediction in research contexts with rich data and weak theory (Wold 1985). Nevertheless, 25 studies (12.25%) incorrectly justified the use of PLS-SEM by citing its appropriateness for testing well-established complex theories. Further justifications for its use relate to the ability to cope with highly complex models (27 studies, 13.24%) and categorical variables (26 studies, 12.75%). In studies published prior to 2000, the authors mentioned non-normal data, prediction orientation (both $p \leq 0.01$), and the use of categorical variables ($p \leq 0.05$) significantly more often than did authors of recent studies, suggesting changes over time. Conversely, the formative indicator argument was significantly ($p \leq 0.01$) more prevalent in studies published in 2000 and beyond.

Several of these characteristics have been extensively discussed in the methodological literature on PLS-SEM. The sample size argument in particular has been the subject of much debate (e.g., Goodhue et al. 2006; Marcoulides and Saunders 2006). While many discussions on this topic are anecdotal in nature, few studies have systematically evaluated PLS-SEM's performance when the sample size is small (e.g., Chin and Newsted 1999; Hui and Wold 1982). More recently, Reinartz et al. (2009) showed that PLS-SEM achieves high levels of statistical power—in comparison to its covariance-based counterpart—even if the sample size is relatively small (i.e., 100 observations). Similarly, Boomsma and Hoogland's (2001) study underscores CB-SEM's need for relatively large sample sizes to achieve robust parameter estimates. PLS-SEM is therefore generally more favorable with smaller sample sizes and more complex models. However, as noted by Marcoulides and Saunders (2006), as well as Sosik et al. (2009), PLS-SEM is not a silver bullet for use with samples of any size, or a panacea for dealing with empirical research challenges. All statistical techniques require consideration of the sample size in the context of the model and data characteristics, and PLS-SEM is no exception.

Theoretical discussions (Beebe et al. 1998) as well as simulation studies (Cassel et al. 1999) indicate that the PLS-SEM algorithm transforms non-normal data in accordance with the central limit theorem (see also Dijkstra 2010). These studies show that PLS-SEM results are robust if data are highly skewed, also when formative measures

are used (Ringle et al. 2009). In contrast, maximum likelihood-based CB-SEM requires normally distributed indicator variables. Since most empirical data do not meet this requirement, several studies have investigated CB-SEM with non-normal data and reported contradictory results (e.g., Babakus et al. 1987; Reinartz et al. 2009). Given the multitude of alternative estimation procedures for CB-SEM, such as weighted least squares and unweighted least squares, it is questionable whether the choice of PLS-SEM over CB-SEM can be justified solely by distribution considerations.

CB-SEM can accommodate formative indicators, but to ensure model identification, researchers must follow rules that require specific constraints on the model (Bollen and Davies 2009; Diamantopoulos and Riefler 2011). These constraints often contradict theoretical considerations, and the question arises whether model design should guide theory or vice versa. In contrast, similar problems do not arise in PLS-SEM, which only requires the constructs to be structurally linked. As a result, PLS-SEM provides more flexibility when formative measures are involved.

Data characteristics

Sample size is a basic PLS-SEM application issue. In line with the frequently noted argument that PLS-SEM works particularly well with small sample sizes, the average sample size in our review (5% trimmed mean=211.29) is clearly lower than that reported by Shah and Goldstein (2006) in their review of CB-SEM studies (mean=246.4). The same holds for Baumgartner and Homburg's (1996) review; these authors reported a median sample size of 180 (median in this study=159.00). It is interesting to note that several models exhibited very large sample sizes considered atypical in PLS-SEM (Johnson et al. 2006, $n=2,990$; Sirohi et al. 1998, $n=16,096$; Xu et al. 2010, $n=2,333$). Conversely, 76 of 311 models (24.44%) have less than 100 observations, with Lee (1994) having the smallest sample size ($n=18$). An assessment over time shows that the sample size was higher (albeit not significantly) in recent models (5% trimmed mean=229.37) than in earlier models (5% trimmed mean=175.22).

As a popular rule of thumb for robust PLS-SEM estimations, Barclay et al. (1995) suggest using a minimum sample size of ten times the maximum number of paths aiming at any construct in the outer model (i.e., the number of formative indicators per construct) and inner model (i.e., the number of path relationships directed at a particular construct). Although this rule of thumb does not take into account effect size, reliability, the number of indicators, and other factors known to affect power and can thus be misleading, it nevertheless provides a rough estimate of minimum sample size requirements. While most models meet this rule of thumb, 28 of the estimated models

(9.00%) do not; on average they are 45.18% below the recommended sample size. Moreover, 82.61% of the models published before 2000 meet this rule of thumb, but the percentage increased significantly ($p \leq 0.01$) in more recent models (93.39%). Thus, researchers seem to be more aware of sample size issues in PLS-SEM.

The use of holdout samples to evaluate the results' robustness is another area of concern (Hair et al. 2010). Only 13 of 204 studies (6.37%) included a holdout sample analysis. While this may be due to data availability issues, PLS-SEM's distribution-free character, which relies on resampling techniques such as bootstrapping for significance testing (Henseler et al. 2009), might also be a reason for not using a holdout sample. In light of the PLS-SEM bias, however, substantiating parameter estimates' robustness through holdout samples is of even greater importance in PLS-SEM than in CB-SEM.

Although prior research has provided evidence of PLS-SEM's robustness in situations in which data are extremely non-normal (e.g., Cassel et al. 1999; Reinartz et al. 2009), researchers should nevertheless consider the data distribution. Highly skewed data inflate bootstrap standard errors (Chernick 2008) and thus reduce statistical power, which is especially problematic given PLS-SEM's tendency to underestimate inner model relationships (Wold 1982). Despite this concern, only 19 studies (9.31%) report the extent to which the data are non-normal, and no significant differences were evident over time.

Several researchers stress that PLS-SEM generally works with nominal, ordinal, interval, and ratio scaled variables (e.g., Fornell and Bookstein 1982; Haenlein and Kaplan 2004; Reinartz et al. 2009). It is therefore not surprising that researchers routinely use categorical (14 studies, 6.86%) or even binary variables (43 studies, 21.08%). However, this practice should be considered with caution. For example, researchers may decide to use a binary single indicator to measure an endogenous construct to indicate a choice situation. In this set-up, however, the latent construct becomes its measure (Fuchs and Diamantopoulos 2009), which proves problematic for approximations in the PLS-SEM algorithm since path coefficients are estimated by OLS regressions. Specifically, OLS requires the endogenous latent variable scores to be continuous, a property that cannot be met in such a set-up. Likewise, using binary indicators in reflective models violates this OLS assumption, because reflective indicators are regressed on the latent variable scores when estimating outer weights. Correspondingly, Jakobowicz and Derquenne (2007, p. 3668) point out that "when working with continuous data [...], PLS does not face any problems, but when working with nominal or binary data it is not possible to suppose there is any underlying continuous distribution." In a similar vein, Lohmöller (1989) argues that standard procedures for applying linear

equations cannot be used for categorical variables. Based on Lohmöller's (1989) early efforts to include categorical variables, Jakobowicz and Derquenne (2007) developed a modified version of the PLS-SEM algorithm based on generalized linear models that is, however, restricted to reflective measures. The standard PLS-SEM algorithm's application does not account for these extensions and thus often violates fundamental OLS principles when used on categorical variables.⁴ As a consequence, researchers should not use categorical variables in endogenous constructs and should carefully interpret the meaning of categorical variables in exogenous constructs. Alternatively, categorical variables can be used to split the data set for PLS multigroup comparisons (Sarstedt et al. 2011b).

Model characteristics

Table 2 provides an overview of model characteristics in PLS-SEM studies. The average number of latent variables in PLS path models is 7.94, which is considerably higher than the 4.70 reported in Shah and Goldstein's (2006) review of CB-SEM studies. In addition, the average number of latent variables is significantly ($p \leq 0.01$) higher in models published in 2000 and beyond (8.43) than in earlier models (6.29). The average number of inner model relationships (10.56) has also increased (but not significantly) over time (9.10 before 2000 and 10.99 thereafter). Overall, model complexity in PLS-SEM studies has clearly increased.

Our review examines three types of models (i.e., focused, unfocused, and balanced). Focused models have a small number of endogenous latent variables that are explained by a rather large number of exogenous latent variables (i.e., the number of exogenous latent variables is at least twice as high as the number of endogenous latent variables). A total of 109 focused models (35.05%) were used. An unfocused model was defined as having many endogenous latent variables and mediating effects, and a comparatively smaller number of exogenous latent variables (i.e., the number of endogenous latent variables is at least twice as high as the number of exogenous latent variables). A total of 85 unfocused models (27.33%) were employed. The remaining 117 models (37.62%) were balanced models, identified as between the focused and unfocused model types. Focused models were significantly ($p \leq 0.01$) more prevalent recently, while unfocused ($p \leq 0.05$) and balanced ($p \leq 0.01$) models appeared significantly more often in studies published before 2000.

Focused and balanced models meet PLS-SEM's prediction goal, while CB-SEM may be more suitable for

⁴ The same holds for comparative scales such as rank order, paired comparison, or constant sum scales.

Table 2 Descriptive statistics for model characteristics

Criterion	Results (<i>n</i> =311)	Proportion (%)	Before 2000 (<i>n</i> =69)	2000 onward (<i>n</i> =242)	Other leading journals (<i>n</i> =250)	Top tier journals (<i>n</i> =61)
Number of latent variables						
Mean	7.94	–	6.29	8.43***	7.76	8.69*
Median	7.00		6.00	8.00	7.00	8.00
Range	(2; 29)		(2; 16)	(2; 29)	(2; 29)	(2; 20)
Number of inner model path relations ¹						
Mean	10.56	–	9.10	10.99	10.10	12.41*
Median	8.00		6.00	9.00	8.00	10.00
Range	(1; 38)		(1; 37)	(1; 38)	(1; 38)	(1; 35)
Model type ¹						
Focused	109	35.05	8	101***	88	21
Unfocused	85	27.33	26**	59	59	26***
Balanced	117	37.62	35***	82	103***	14
Mode of outer models						
Only reflective	131	42.12	23	108*	115***	16
Only formative	20	6.43	12***	8	17	3
Reflective and formative	123	39.55	26	97	88	35***
Not specified	37	11.90	8	29	30	7
Number of indicators per reflective construct						
Mean	3.99	–	2.90	4.25***	4.15**	3.35
Median	3.50		2.50	4.00	3.50	3.00
Range	(1; 27)		(1; 16)	(1; 27)	(1; 27)	(1; 16)
Number of indicators per formative construct						
Mean	4.62	–	3.63	4.98***	4.83**	3.99
Median	4.00		3.00	4.00	4.00	3.50
Range	(1; 20)		(1; 15)	(1; 20)	(1; 18)	(1; 20)
Total number of indicators in models						
Mean	29.55	–	19.58	32.51***	29.47	29.85
Median	24.00		17.00	27.00	24.00	25.00
Range	(4; 131)		(4; 101)	(5; 131)	(4; 131)	(7; 101)
Number of models with single-item constructs	144	46.30	42	102	111	33

¹ Focused model: number of exogenous latent variables at least twice as high as the number of endogenous latent variables; unfocused model: number of endogenous latent variables at least twice as high as the number of exogenous latent variables; balanced model: all remaining types of models

*** (**, *) indicates a significant difference between “before 2000”/“2000 onward” and “other leading journals”/“top tier journals,” respectively, at a 1% (5%, 10%) significance level; results based on independent samples t-tests and (one-tailed) Fisher’s exact tests (no tests for median differences)

explaining unfocused models. Only 11 of the 57 applications that explicitly stated they used PLS-SEM for prediction purposes also examined a focused model. In contrast, 23 of 57 models supposedly designed for prediction actually examined an unfocused model. Thus, there appears to be a lack of awareness of the relationship between PLS-SEM’s prediction goal and the type of model examined.

Regarding the outer models, PLS path models typically have been assumed to be composed either of solely reflectively measured latent variables (131 models,

42.12%) or a combination of reflectively and formatively measured latent variables (123 models, 39.55%). Much fewer PLS path models were assumed to be based exclusively on formative measures (20 models, 6.43%). Surprisingly, 37 models (11.90%) do not offer a description of the constructs’ measurement modes, despite the extensive debate on measurement specification (e.g., Bollen 2011; Diamantopoulos and Sigauw 2006; Diamantopoulos and Winklhofer 2001; Jarvis et al. 2003). Interestingly, neither the proportion of models incorporating reflective and formative measures nor those lacking a description of

the measurement mode changed significantly over time (Table 2).

The average number of indicators is 3.99 for reflective constructs, which is significantly ($p \leq 0.01$) higher in recent models (2.90 before 2000; 4.25 thereafter). The higher number of indicators for formative constructs (4.62) is logical since the construct should be represented by the entire population of indicators (Diamantopoulos et al. 2008). As with reflective constructs, the number has increased significantly ($p \leq 0.01$) recently (3.63 before 2000; 4.98 thereafter). The total number of indicators used is large, and much higher than in CB-SEM models. This is primarily a result of the larger number of constructs, however, and not a larger average number of indicators per construct. Our review identified an average of 29.55 indicators per PLS path model, with a significant ($p \leq 0.01$) increase in recent models (32.51) compared to earlier models (19.58). Furthermore, the number of indicators is much higher than the 16.30 indicators per CB-SEM model reported by Shah and Goldstein (2006). Similarly, Baumgartner and Homburg (1996) reported a median value of 12 indicators per CB-SEM analysis, which is also considerably lower than in our review (median=24).

Researchers have argued that if a construct's scope is narrow, unidimensional, and unambiguous for the respondents, using single-item measures is the best approach (e.g., Nunnally 1967; Sackett and Larson 1990), an argument which Bergkvist and Rossiter (2007, 2009) have recently empirically supported.⁵ Unlike with CB-SEM, where the inclusion of single items generally leads to model under-identification (e.g., Fuchs and Diamantopoulos 2009), PLS-SEM is not restricted in this respect. Thus, it is not surprising that 144 models (46.30%) included single-item measures. Although using single-item measures can also prove beneficial from an analytic perspective as they generally increase response rates (e.g., Fuchs and Diamantopoulos 2009; Sarstedt and Wilczynski 2009), one has to keep in mind that the utilization of single-item measures is contrary to PLS-SEM's concept of consistency at large. For example, Reinartz et al. (2009) showed that only with reasonable outer model quality (in terms of indicators per construct and loadings), does PLS-SEM yield acceptable parameter estimates when the sample size is restricted. While the conflict between psychometric properties and consistency at large has not yet been addressed in research, the use of

single-item measures should be considered with caution when using PLS-SEM.

Outer model evaluation

Outer model assessment involves examining individual indicator reliabilities, the reliabilities for each construct's composite of measures (i.e., internal consistency reliability), as well as the measures' convergent and discriminant validities. When evaluating how well constructs are measured by their indicator variables, individually or jointly, researchers need to distinguish between reflective and formative measurement perspectives (e.g., Diamantopoulos et al. 2008). While criteria such as Cronbach's alpha and composite reliability are commonly applied to evaluate reflective measures, an internal consistency perspective is inappropriate for assessing formative ones (e.g., Diamantopoulos and Winklhofer 2001). As Diamantopoulos (2006, p. 11) points out, when "formative measurement is involved, reliability becomes an irrelevant criterion for assessing measurement quality." Similarly, formative measures' convergent and discriminant validities cannot be assessed by empirical means (e.g., Hair et al. 2011). Our review examines if and how authors evaluate the suitability of constructs, and whether the concepts typically related to reflective outer models' assessment are also applied in formative settings.⁶

While the decision on outer model set-up should be based primarily on theoretical grounds (e.g., Diamantopoulos and Winklhofer 2001; Jarvis et al. 2003), Gudergan et al. (2008) propose a confirmatory tetrad analysis technique for PLS-SEM (CTA-PLS) that allows researchers to empirically test constructs' measurement modes (Bollen and Ting 2000). Since CTA-PLS has been introduced only recently, it may not have been applied in any of the reviewed studies. Future research should, however, routinely employ this technique as a standard means for model assessment (Coltman et al. 2008).

Reflective outer models Assessment of reflective outer models involves determining indicator reliability (squared standardized outer loadings), internal consistency reliability (composite reliability), convergent validity (average variance extracted, AVE), and discriminant validity (Fornell-Larcker criterion, cross-loadings) as described by, for example, Henseler et al. (2009) and Hair et al. (2011). The findings

⁵ It has to be noted that, contrary to Bergkvist and Rossiter's (2007, 2009) findings, several studies have shown that single-item measures do not necessarily match multi-item measures in terms of psychometric properties (e.g., Gardner et al. 1989; Kwon and Trail 2005; Sarstedt and Wilczynski 2009).

⁶ We also computed the means, medians, and ranges of statistics related to outer (and inner) model assessment (e.g., composite reliability, indicator weights). These are available from the authors upon request.

of the marketing studies reviewed are shown in Table 3 (Panel A).

Overall, 254 of 311 models (81.67%) reported that reflectively measured constructs were included.⁷ Many, but surprisingly not all, models commented on reliability. Specifically, 157 of 254 models (61.81%) reported outer loadings, and thus indirectly specify the indicator reliability, with only 19 models explicitly addressing this criterion. Support for indicator reliability was significantly ($p \leq 0.01$) more prevalent in early models than in more recent ones.

Internal consistency reliability was reported for 177 models (69.69%). Prior assessments of reporting practices in the CB-SEM context (e.g., Shah and Goldstein 2006; Shook et al. 2004) revealed that Cronbach's alpha is the most common measure of internal consistency reliability. However, Cronbach's alpha is limited by the assumption that all indicators are equally reliable (tau-equivalence), and efforts to maximize it can seriously compromise reliability (Raykov 2007). In contrast, composite reliability does not assume tau-equivalence, making it more suitable for PLS-SEM, which prioritizes indicators according to their individual reliability. The majority of models report composite reliability, either exclusively (73 models; 28.74%) or in conjunction with Cronbach's alpha (69 models; 27.17%). A total of 35 models (13.78%) reported only Cronbach's alpha. Application of composite reliability, individually or jointly with Cronbach's alpha, was significantly ($p \leq 0.01$) more prevalent recently (Table 3).

Convergent validity was examined in 153 of 254 models (60.24%). Authors primarily relied on the AVE (146 models), while in the remaining seven models, they incorrectly interpreted composite reliability or the significance of the loadings as indicative of convergent validity. Moreover, a total of 154 models (60.63%) provided evidence of discriminant validity, with most (111 models) solely comparing the constructs' AVEs with the inter-construct correlations (Fornell and Larcker 1981), a practice that was significantly ($p \leq 0.01$) more prevalent recently (Table 3, Panel A). Alternatively, authors examined only cross-loadings (12 models), a criterion which can generally be considered more liberal in terms of discriminant validity. In the 31 remaining models both criteria were reported.

Formative outer models Overall, 143 of 311 models (45.98%) contained at least one formatively measured construct. Table 3 (Panel B) shows the results of this analysis. A total of 33 models (23.08%) with formatively measured

constructs inappropriately evaluated the corresponding measures using reflective outer model assessments. This mistake was significantly ($p \leq 0.01$) more prevalent in earlier studies, in which 21 of 38 models (55.26%) used reflective criteria to evaluate formative measures, compared to 12 of 105 models (11.43%) in more recent studies.

Several statistical criteria have been suggested to evaluate the quality of formative measures. The primary statistic for assessing formative indicators is their weight, which was reported in 33 of 143 models (23.08%). Evaluation of indicator weights should also include examining their significance by means of resampling procedures. While most studies applied blindfolding or bootstrapping, these procedures were primarily used to evaluate inner model parameter estimates rather than the significance of formative indicators' weights. Only 25 models (17.48%) reported t-values or corresponding p-values. In fact, most researchers did not comment on this important issue.

Multicollinearity between indicators is an important issue in assessing formative measures because of the potential for unstable indicator weights (Cenfetelli and Basselier 2009). Since formative indicator weights are frequently smaller than reflective indicators' loadings, this can lead to misinterpretations of the indicator relevance for the construct domain (Diamantopoulos and Winklhofer 2001). Only 22 of 143 models (15.38%) using formative measures—all of which appeared in 2000 and beyond—assessed multicollinearity, relying primarily on the tolerance and variance inflation factor (VIF).

Overall, our findings regarding outer model assessments give rise to concern. First, given that reliability and validity assessments play a vital role in outer model assessment, the proportion of studies that do not report reliability and validity measures is disconcerting. If measures lack reliability and validity, inner model estimates may be substantially biased, leading researchers to overlook relationships that could be significant. Indeed, since PLS-SEM applications often serve as a basis for theory development, promising research avenues might have been overlooked. Further, despite the broad discussion on the inappropriateness of internal consistency-based measures for evaluating blocks of formative indicators (e.g., Diamantopoulos 2006; Diamantopoulos and Winklhofer 2001; Diamantopoulos et al. 2008), it is distressing to see that authors still apply a cookbook-like recipe used with reflective measures to assess formative ones. The limited multicollinearity assessment in earlier studies, can potentially be traced back to Diamantopoulos and Winklhofer's (2001) article, which raised considerable awareness of this issue among marketing scholars.

⁷ This analysis includes only those models specified as including only reflective and a mixture of reflective and formative latent variables. Models in which the measurement mode was not specified are not included.

Table 3 Evaluation of outer models

Panel A: Reflective outer models

	Empirical test criterion in PLS-SEM	Number of models reporting (n=254)	Proportion reporting (%)	Before 2000 (n=49)	2000 onward (n=205)	Other leading journals (n=203)	Top tier journals (n=51)
Indicator reliability	Indicator loadings ¹	157	61.81	38***	119	126	31
Internal consistency reliability	Only composite reliability	73	28.74	5	68***	56	17
	Only Cronbach's alpha	35	13.78	9	26	31*	4
	Both	69	27.17	6	63***	60*	9
Convergent validity	AVE	146	57.48	18	128	119	27
	Other	7	2.76	1	6	6	1
Discriminant validity	Only Fornell-Larcker criterion	111	43.70	15	96***	93	18
	Only cross-loadings	12	4.72	3	9	9	3
	Both	31	12.20	4	27	20	11**

Panel B: Formative outer models

	Empirical test criterion in PLS-SEM	Number of models reporting (n=143)	Proportion reporting (%)	Before 2000 (n=38)	2000 onward (n=105)	Other leading journals (n=105)	Top tier journals (n=38)
–	Reflective criteria used to evaluate formative constructs	33	23.08	21***	12	29**	4
Indicator's absolute contribution to the construct	Indicator weights	33	23.08	7	26	14	19***
Significance of weights	Standard errors, significance levels, t-values/p-values for indicator weights	25	17.48	3	22*	13	12***
Multicollinearity	Only VIF/tolerance	17	11.89	0	17***	12	5
	Only condition index	1	0.70	0	1	1	0
	Both	4	2.80	0	4	4	0

¹ Single item constructs were excluded from this analysis
 *** (,) indicates a significant difference between “before 2000”/“2000 onward” and “other leading journals”/“top tier journals,” respectively, at a 1% (5%, 10%) significance level; results based on (one-tailed) Fisher's exact tests

With the emerging interest in the use of formative measurement in the marketing discipline, PLS-SEM is likely to be more widely applied. But PLS-SEM researchers need to pay closer attention to the validation of formative measures by taking into account standard (e.g., significance, multicollinearity) as well as recently proposed evaluation steps (e.g., absolute and relative indicator contributions, suppressor effects) (Bollen 2011; Cenfetelli and Bassellier 2009; MacKenzie et al. 2011).

Inner model evaluation

If the outer model evaluation provides evidence of reliability and validity, it is appropriate to examine inner model estimates. The classic CB-SEM–related distinction between variance and covariance fit is not applicable in PLS-SEM, which is primarily due to the PLS-SEM assumption of distribution-free variance. Thus, when using PLS-SEM, researchers must focus their evaluation on variance-based, non-parametric evaluation criteria to assess the inner model’s quality (e.g., Chin 1998, 2010; Henseler et al. 2009). Table 4 provides an overview of the results regarding inner model evaluation.

The primary criterion for inner model assessment is the coefficient of determination (R^2), which represents the amount of explained variance of each endogenous latent variable. In our review, 275 models (88.42%) reported R^2 values to assess the quality of their findings. Only 16 models (5.14%) considered a particular exogenous latent variable’s relative impact on an endogenous latent variable by means of changes in the R^2 values, based on the effect size f^2 (Cohen 1988). Sample re-use techniques proposed by Stone (1974) and Geisser (1974) can be used to assess the model’s predictive validity by means of the cross-validated redundancy measure Q^2 . This technique is a synthesis of cross-validation and function fitting, and Wold (1982, p. 30) argues that it fits PLS-SEM “like hand in glove.” Nevertheless, only 51 models (16.40%) reported this criterion. Like f^2 , the Q^2 can assess an individual construct’s predictive relevance for the model by omitting selected inner model relationships and computing changes in the criterion’s estimates (q^2). None of the models reported this statistic, although research has stressed its importance for inner model evaluation (e.g., Chin 1998; Henseler et al. 2009).

Tenenhaus et al. (2004) proposed a global criterion for goodness-of-fit (i.e., the GoF index). This criterion is defined by the geometric mean of the average communality and the model’s average R^2 value. Thus, the GoF does not represent a true global fit measure (even though its name suggests this), and threshold values for an acceptable “goodness-of-fit” can hardly be derived because acceptable R^2 values depend on the research context (Hair et al. 2011) and the construct’s role in the model (e.g., key target

Table 4 Evaluation of inner models

Criterion	Empirical test criterion in PLS-SEM	Number of models reporting (n=311)	Proportion reporting (%)	Before 2000 (n=69)	2000 onward (n=242)	Other leading journals (n=250)	Top tier journals (n=61)
Endogenous constructs’ explained variance	R^2	275	88.42	56	219**	221	54
Effect size	f^2	16	5.14	0	16**	14	2
Predictive relevance	Cross-validated redundancy Q^2	51	16.40	5	46**	46**	5
Relative predicted relevance	q^2	0	0.00	–	–	–	–
Overall goodness-of-fit	GoF	16	5.14	0	16**	12	4
Path coefficients	Absolute values	298	95.82	68	230	239	59
Significance of path coefficients	Standard errors, significance levels, t-values, p-values	287	92.28	60	227*	230	57
Confidence intervals	–	0	0.00	–	–	–	–
Criterion	Empirical test criterion in PLS-SEM	Number of studies reporting (n=204)	Proportion reporting (%)	Before 2000 (n=39)	2000 and beyond (n=165)	Other leading journals (n=163)	Top tier journals (n=41)
Observed heterogeneity	Categorical moderator	47	23.04	8	39	41	6
Unobserved heterogeneity	Continuous moderator	15	7.35	1	14	12	3
	Response-based segmentation techniques (e.g., FIMIX-PLS)	0	0.00	–	–	–	–

*** (,) indicates a significant difference between “before 2000”/“2000 onward” and “other leading journals”/“top tier journals,” respectively, at a 1% (5%, 10%) significance level; results based on (one-tailed) Fisher’s exact tests

construct versus mediating construct). Moreover, the GoF is not universally applicable for PLS-SEM as it is based on reflective outer models' communalities. Thus, the proposed GoF is conceptually inappropriate whenever outer models are formative, or when single indicator constructs are involved. Despite these concerns, 16 studies (5.14%) reported this relatively new measure, five of which included single items or formatively measured constructs.

Standardized path coefficients provide evidence of the inner model's quality, and their significance should be assessed using resampling procedures. A total of 298 models (95.82%) reported path coefficients, and 287 (92.28%) commented on their significance by, for example, providing t-value statistics and/or corresponding p-values. But none of the articles reported resampling-based confidence intervals (Henseler et al. 2009).

PLS-SEM applications are usually based on the assumption that the data stem from a single population. In many real-world applications, however, this assumption of homogeneity is unrealistic, as different population parameters are likely to occur for different subpopulations, such as segments of consumers, firms, industries, or countries. A total of 47 studies (23.04%) accounted for observed heterogeneity by considering categorical moderating variables and comparing corresponding group-specific path coefficient estimates, for example, by using multigroup comparison techniques (Sarstedt et al. 2011b). Other studies (15; 7.35%) evaluated interaction effects by modeling (continuous) moderator variables that potentially affect the strengths or direction of specific path relationships (e.g., Henseler and Chin 2010). However, while the consideration of observed heterogeneity generally proves valuable from a theory perspective, heterogeneity is often unobservable and cannot be attributed to any predetermined variable(s). The impact of unobserved heterogeneity on SEM results can be considerable, and if not carefully taken into account, may entail misleading interpretations (e.g., Jedidi et al. 1997). As a consequence, PLS-SEM analyses require the use of complementary techniques for response-based segmentation that allow testing for and dealing with unobserved heterogeneity. Finite mixture partial least squares (FIMIX-PLS; Hahn et al. 2002; Sarstedt et al. 2011a) is currently regarded the primary approach in the field (e.g., Rigdon et al. 2010). Based on a mixture regression concept, FIMIX-PLS simultaneously estimates inner model parameter and ascertains the data structure's heterogeneity by calculating the probability of the observations' segment membership so that they fit into a predetermined number of segments. However, in contrast to conventional mixture regressions, models in FIMIX-PLS can comprise a multitude of interrelated endogenous latent variables. In light of the approach's performance in prior

studies (e.g., Ringle et al. 2010a, b; Sarstedt and Ringle 2010) and its availability through the software application SmartPLS (Ringle et al. 2005), Hair et al. (2011) suggest that researchers should routinely use the technique to evaluate whether the results are distorted by unobserved heterogeneity. While the issue of unobserved heterogeneity is important in many marketing studies using PLS-SEM, none of the reviewed studies carried out this type of analysis.

Overall, our review shows that even though researchers have made a significantly ($p \leq 0.05$) greater use of model evaluation criteria (i.e., R^2 , f^2 , Q^2 , and GoF) in recent years (Table 4), they apply few of the criteria available for inner model assessment. We urge researchers to use a greater number of measures to assess the inner model's quality. Gathering more evidence for or against the model's quality is particularly required for PLS-SEM since it does not allow assessment of model fit as CB-SEM does. Moreover, such an analysis should always be supplemented by checking for heterogeneity.

Reporting

Reporting is a crucial issue in empirical marketing studies, and articles should provide readers with sufficient information to enable them to replicate the results and to fully assess the study's quality, regardless of which method is applied (Stewart 2009). In general, PLS-SEM studies should provide information on (1) the population and sample structure, (2) the distribution of the data, (3) the conceptual model, including a description of the inner and outer models, as well as the measurement modes, and (4) the statistical results to corroborate the subsequent interpretation and conclusions (Chin 2010). In addition, researchers should report specific technicalities related to the software and computational options used as well as the parameter settings of ancillary analysis procedures.

It is crucial to know which software application was used as different programs exhibit different default settings. Despite this fact and license agreement requirements, only 100 studies (49.02%) indicated which software was used for model estimation. Of those providing this information, 64 studies used PLS Graph (Chin 2003), 24 used SmartPLS (Ringle et al. 2005), and 12 used LVPLS (Lohmöller 1987).

In addition, the initial values for the outer model relationships, parameter settings, computational options, and the resulting (maximum) number of iterations need to be reported, which none of the studies did. The selection of initial values for outer weights may impose changes in the outer models and/or the inner model estimates (e.g., Henseler et al. 2009). Specific parameter settings (e.g., stop criterion) and computational options (e.g., weighting scheme for determining inner model proxies) can also entail

different model estimation outcomes and are sometimes inadequate in certain model configurations. For example, the centroid weighting scheme ensures the PLS-SEM algorithm's convergence (Henseler 2010), but it should not be used for estimating higher order component models (Henseler et al. 2009). Finally, although the PLS-SEM algorithm usually converges for reasonably small stop criterion parameter settings (e.g., 10^{-5} ; Wold 1982), it may not converge in some extreme data constellations (Henseler 2010). Moreover, convergence is usually not reached if the stop criterion is extremely low (e.g., 10^{-20}). To assess if the PLS-SEM algorithm converged before reaching the pre-specified stop criterion (i.e., the actual number of iterations is smaller than the pre-specified maximum number of iterations), and thus provides the optimum model estimates, one needs to know the (maximum) number of iterations.

While all PLS-SEM evaluations should rely on resampling procedures, only 135 studies (66.18%) explicitly mention the use of bootstrapping or jackknifing in model evaluation. The number reporting resampling is significantly ($p \leq 0.01$) more prevalent in recent studies (115 before 2000; 20 thereafter), but it still needs to be higher. Moreover, only 66 of 135 studies, all of which appeared more recently, reported the concrete parameter settings. Reporting of this detailed information is important since misspecifications in, for example, the bootstrapping parameter settings (e.g., with regard to the number of bootstrap samples) can lead to biased standard error estimates (e.g., Chernick 2008).

Overall, none of the reviewed studies provided sufficient information to replicate and validate the analytical finding with only 10 studies (4.90%) reporting the empirical covariance/correlation matrix for the indicator variables. While readers usually gain a basic understanding of the analysis, overall transparency leaves much to be desired.

Impact of journal quality on PLS-SEM use

A final analysis addresses the question of whether a comparison of top tier journals with other leading journals reveals significant differences in PLS-SEM use. We therefore compare PLS-SEM use in the top five journals according to Hult et al.'s (2009) ranking (i.e., *Journal of Marketing*, *Journal of Marketing Research*, *Journal of Consumer Research*, *Marketing Science*, and *Journal of the Academy of Marketing Science*; 41 studies with 61 models) with those in other leading journals (i.e., those journals on positions 6 to 30 in Hult et al.'s (2009) ranking; 163 studies with 250 models) to determine the impact of journal quality on application and reporting practices. Surprisingly, our

assessment reveals that there are few significant differences between the journal tiers.⁸

Meeting the ten times rule of thumb for the minimum sample size is significantly ($p \leq 0.10$) more prevalent in models published in other leading journals (231 models; 92.40%) than in the top tier journals (52 models; 85.25%). This finding is likely due to differences in model characteristics (Table 2). Specifically, PLS path models in the top tier journals include a significantly ($p \leq 0.10$) larger number of latent variables and inner model relationships. Likewise, models in top tier journals are to a significant extent ($p \leq 0.01$) more frequently unfocused (i.e., contain a relatively high number of endogenous latent variables relative to exogenous latent variables) and more often incorporate both reflective and formative measures.

In terms of outer model evaluation (Table 3), articles in other leading journals use Cronbach's alpha significantly ($p \leq 0.10$) more frequently for reliability assessment and more often apply reflective criteria to evaluate formative measures ($p \leq 0.05$), both of which are inappropriate. Models in top tier journals significantly ($p \leq 0.01$) more often report formative indicator weights and the corresponding significance levels. Most other aspects of outer and inner model assessment (Tables 3 and 4) yield no significant differences, indicating that, in general, reporting practices are equally prevalent among the top tier and other leading journals.

Conclusion

Our review substantiates that PLS-SEM has become a more widely used method in marketing research. But PLS-SEM's methodological properties are widely misunderstood, which at times leads to misapplications of the technique, even in top-tier marketing journals. For instance, our findings suggest that researchers rely on a standard set of reasons from previous publications to substantiate the use of PLS-SEM rather than carefully pinpointing specific reasons for its use. Likewise, researchers do not fully capitalize on the criteria available for model assessment and sometimes even misapply measures. A potential reason for many researchers', reviewers', and editors' unfamiliarity with the principles of PLS-SEM might be that textbooks on multivariate data analysis do not discuss PLS-SEM at all (e.g., Churchill and Iacobucci 2010; Malhotra 2010), or only superficially (e.g.,

⁸ In addition, we contrasted PLS-SEM use between studies that appeared in the five highest ranked and the five lowest ranked journals (20 studies with 23 model estimations) considered in this specific analysis. These analyses produced highly similar results compared to those reported here and are available from the authors upon request.

Table 5 Guidelines for applying PLS-SEM

Criterion	Recommendations / rules of thumb	Suggested references
<i>Data characteristics</i>		
General description of the sample	Use “ten times rule” as rough guidance for minimum sample size	Barclay et al. 1995
Distribution of the sample	Robust when applied to highly skewed data; report skewness and kurtosis	Cassel et al. 1999, Reinartz et al. 2009
Use of holdout sample	30% of original sample	Hair et al. 2010
Provide correlation / covariance matrix (or raw data in online appendix)	–	–
Measurement scales used	Do not use categorical variables in endogenous constructs; carefully interpret categorical variables in exogenous constructs	–
<i>Model characteristics</i>		
Description of the inner model	Provide graphical representation illustrating all inner model relations	–
Description of the outer models	Include a complete list of indicators in the appendix	–
Measurement mode of latent variables	Substantiate measurement mode by using CTA-PLS	Diamantopoulos et al. 2008; Gudergan et al. 2008; Jarvis et al. 2003
<i>PLS-SEM algorithm settings and software used</i>		
Starting values for weights for initial approximation of the latent variable scores	Use an uniform value of 1 as an initial value for each of the outer weights	Henseler 2010
Weighting scheme	Use path weighting scheme	Henseler 2010; Henseler et al. 2009
Stop criterion	Sum of the outer weights’ changes between two iterations $<10^{-5}$	Wold 1982
Maximum number of iterations	300	Ringle et al. 2005
Software used	Report software, including version to indicate default settings	–
<i>Parameter settings for procedures used to evaluate results</i>		
Bootstrapping		Efron 1981
Sign change option	Use individual sign changes	Henseler et al. 2009
Number of bootstrap samples	5,000; must be greater than the number of valid observations	Hair et al. 2011
Number of bootstrap cases	Equal to the number of valid observations	Hair et al. 2011
Blindfolding	Use cross-validated redundancy	Chin 1998; Geisser 1974; Stone 1974
Omission distance d	Number of valid observations divided by d must not be an integer; choose $5 \leq d \leq 10$	Chin 1998
CTA-PLS	5,000 bootstrap samples; rejection of reflective measurement approach if a non-redundant vanishing tetrad is significantly (bias-corrected confidence interval) different from zero (Bonferroni correction for multiple tests)	Coltman et al. 2008; Gudergan et al. 2008
Multigroup comparison	Use distribution-free approaches to multigroup comparison	Sarstedt et al. 2011b
FIMIX-PLS		Hahn et al. 2002; Sarstedt et al. 2011a
Stop criterion	$\ln(L)$ change $<10^{-15}$	Ringle et al. 2010a
Maximum number of iterations	15,000	Ringle et al. 2010a
Number of segments	Use AIC ₃ and CAIC jointly; also consider EN	Sarstedt et al. 2011a
Ex post analysis	Use multinomial or binary logistic regression, CHAID, C&RT, crosstabs	Sarstedt and Ringle 2010
<i>Outer model evaluation: reflective</i>		
Indicator reliability	Standardized indicator loadings ≥ 0.70 ; in exploratory studies, loadings of 0.40 are acceptable	Hulland 1999
Internal consistency reliability	Do not use Cronbach’s alpha; composite reliability ≥ 0.70 (in exploratory research 0.60 is considered acceptable)	Bagozzi and Yi 1988
Convergent validity	AVE ≥ 0.50	Bagozzi and Yi 1988

Table 5 (continued)

Criterion	Recommendations / rules of thumb	Suggested references
Discriminant validity		
Fornell-Larcker criterion	Each construct's AVE should be higher than its squared correlation with any other construct	Fornell and Larcker 1981
Cross loadings	Each indicator should load highest on the construct it is intended to measure	Chin 1998; Grégoire and Fisher 2006
<i>Outer model evaluation: formative</i>		
Indicators' relative contribution to the construct	Report indicator weights	–
Significance of weights	Report t-values, p-values or standard errors	–
Multicollinearity	VIF < 5 / tolerance > 0.20; condition index <30	Hair et al. 2011
<i>Inner model evaluation</i>		
R ²	Acceptable level depends on research context	Hair et al. 2010
Effect size f ²	0.02, 0.15, 0.35 for weak, moderate, strong effects	Cohen 1988
Path coefficient estimates	Use bootstrapping to assess significance; provide confidence intervals	Chin 1998; Henseler et al. 2009
Predictive relevance Q ² and q ²	Use blindfolding; Q ² > 0 is indicative of predictive relevance; q ² : 0.02, 0.15, 0.35 for weak, moderate, strong degree of predictive relevance	Chin 1998; Henseler et al. 2009
Observed and unobserved heterogeneity	Consider categorical or continuous moderating variables using a priori information or FIMIX-PLS	Henseler and Chin 2010; Rigdon et al. 2010; Sarstedt et al. 2011a, b

Hair et al. 2010). In fact, there is still no introductory textbook on PLS-SEM, and the topic is seldom found in research methodology class syllabi. Journal editors and reviewers should more strongly emphasize making all information available, including the data used, to allow the replication of statistical analyses (e.g., in online appendices). Progressing toward the highest possible level of transparency will substantially improve the way in which research is conducted and its quality, thus allowing accelerated development paths.

Several points should be considered when applying PLS-SEM, some of which, if not handled properly, can seriously compromise the analysis's interpretation and value. Many of these issues, such as the performance of different weighting schemes for algorithm convergence (Henseler 2010) and the limitations of conventional statistical tests in multigroup comparisons (e.g., Rigdon et al. 2010; Sarstedt et al. 2011b), have been reported in the literature. Researchers should consider the PLS-SEM's methodological foundations and complementary analysis techniques more strongly. Similarly, discussions in related fields, such as the recent debate on the validation of formative measures in management information systems research (e.g., Bollen 2011; Cenfetelli and Bassellier 2009; MacKenzie et al. 2011), can provide important guidance for PLS-SEM applications.

Based on the results of our review, we present guidelines for applying PLS-SEM which provide researchers, editors,

and reviewers with recommendations, rules of thumb, and corresponding references (Table 5).

While offering many beneficial properties, PLS-SEM's "soft assumptions" should not be taken as carte blanche to disregard standard psychometric assessment techniques. The quality of studies employing PLS-SEM hopefully will be enhanced by following our recommendations, so that the method's value in research and practice can be clarified.

Acknowledgments The authors would like to thank three anonymous reviewers, Jörg Henseler (University of Nijmegen), and Edward E. Rigdon (Georgia State University) for their helpful remarks on earlier versions of this article.

References

- Babakus, E., Ferguson, C. E., & Jöreskog, K. G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research*, 24(2), 222–228.
- Babin, B. J., Hair, J. F., & Boles, J. S. (2008). Publishing research in marketing journals using structural equation modeling. *Journal of Marketing Theory & Practice*, 16(4), 279–285.
- Bagozzi, R. P. (1994). Structural equation models in marketing research: Basic principles. In R. P. Bagozzi (Ed.), *Principles of marketing research* (pp. 317–385). Oxford: Blackwell.
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16(1), 74–94.

- Barclay, D. W., Higgins, C. A., & Thompson, R. (1995). The partial least squares approach to causal modeling: personal computer adoption and use as illustration. *Technology Studies*, 2(2), 285–309.
- Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: a review. *International Journal of Research in Marketing*, 13(2), 139–161.
- Baumgartner, H., & Pieters, R. (2003). The structural influence of marketing journals: a citation analysis of the discipline and its subareas over time. *Journal of Marketing*, 67(2), 123–139.
- Beebe, K. R., Pell, R. J., & Seasholtz, M. B. (1998). *Chemometrics: A practical guide*. New York: Wiley.
- Bergkvist, L., & Rossiter, J. R. (2007). The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of Marketing Research*, 44(2), 175–184.
- Bergkvist, L., & Rossiter, J. R. (2009). Tailor-made single-item measures of doubly concrete constructs. *International Journal of Advertising*, 28(4), 607–621.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bollen, K. A. (2011). Evaluating effect, composite, and causal indicators in structural equation models. *MIS Quarterly*, 35(2), 359–372.
- Bollen, K. A., & Davies, W. R. (2009). Causal indicator models: identification, estimation, and testing. *Structural Equation Modeling. An Interdisciplinary Journal*, 16(3), 498–522.
- Bollen, K. A., & Ting, K.-F. (2000). A tetrad test for causal indicators. *Psychological Methods*, 5(1), 3–22.
- 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.
- Brannick, M. T. (1995). Critical comments on applying covariance structure modeling. *Journal of Organizational Behavior*, 16(3), 201–213.
- Cassel, C., Hackl, P., & Westlund, A. H. (1999). Robustness of partial least-squares method for estimating latent variable quality structures. *Journal of Applied Statistics*, 26(4), 435–446.
- Cenfetelli, R. T., & Bassellier, G. (2009). Interpretation of formative measurement in information system research. *MIS Quarterly*, 33(4), 689–707.
- Chernick, M. R. (2008). *Bootstrap methods. A guide for practitioners and researchers* (2nd ed.). Hoboken, New Jersey: Wiley.
- Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Chin, W. W. (2003). PLS Graph 3.0. Houston: Soft Modeling Inc.
- Chin, W. W. (2010). How to write up and report PLS analyses. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields* (pp. 655–690). Berlin: Springer.
- Chin, W. W., & Newsted, P. R. (1999). Structural equation modeling analysis with small samples using partial least squares. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 307–341). Thousand Oaks: Sage.
- Churchill, G. A., & Iacobucci, D. (2010). *Marketing research: Methodological foundations*. Mason: South-Western Cengage Learning.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- 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.
- 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., & Riefler, P. (2011). Using formative measures in international marketing models: A cautionary tale using consumer animosity as an example. *Advances in International Marketing*, forthcoming.
- Diamantopoulos, A., & Siguaw, J. A. (2006). Formative vs. reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management*, 17(4), 263–282.
- 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. (2010). Latent variables and indices: Herman Wold's basic design and partial least squares. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields* (pp. 23–46). Berlin: Springer.
- Efron, B. (1981). Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods. *Biometrika*, 68(3), 589–599.
- 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.
- Fornell, C. G., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50.
- Fuchs, C., & Diamantopoulos, A. (2009). Using single-item measures for construct measurement in management research. Conceptual issues and application guidelines. *Die Betriebswirtschaft*, 69(2), 195–210.
- Gardner, D. G., Dunham, R. B., Cummings, L. L., & Pierce, J. L. (1989). Focus of attention at work: construct definition and empirical validation. *Journal of Occupational Psychology*, 62(1), 61–77.
- Garver, M. S., & Mentzer, J. T. (1999). Logistics research methods: employing structural equation modeling to test for construct validity. *Journal of Business Logistics*, 20(1), 33–57.
- Gefen, D., Rigdon, E. E., & Straub, D. (2011). Editor's comments: an update and extension to SEM guidelines for administrative and social science research. *MIS Quarterly*, 35(2), III–XIV.
- Geisser, S. (1974). A predictive approach to the random effects model. *Biometrika*, 61(1), 101–107.
- Goodhue, D., Lewis, W., & Thompson, R. (2006). PLS, small sample size, and statistical power in MIS research. In *HICSS'06: Proceedings of the 39th annual Hawaii international conference on system sciences*. Washington: IEEE Computer Society.
- Grégoire, Y., & Fisher, R. J. (2006). The effects of relationship quality on customer retaliation. *Marketing Letters*, 17(1), 31–46.
- Gudergan, S. P., Ringle, C. M., Wende, S., & Will, A. (2008). Confirmatory tetrad analysis in PLS path modeling. *Journal of Business Research*, 61(12), 1238–1249.
- Haenlein, M., & Kaplan, A. M. (2004). A beginner's guide to partial least squares analysis. *Understanding Statistics*, 3(4), 283–297.
- Hahn, C., Johnson, M. D., Herrmann, A., & Huber, F. (2002). Capturing customer heterogeneity using a finite mixture PLS approach. *Schmalenbach Business Review*, 54(3), 243–269.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Englewood Cliffs: Prentice Hall.
- 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.

- Henseler, J. (2010). On the convergence of the partial least squares path modeling algorithm. *Computational Statistics*, 25(1), 107–120.
- Henseler, J., & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(1), 82–109.
- Henseler, J., & Fassott, G. (2010). Testing moderating effects in PLS path models: An illustration of available procedures. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields* (pp. 713–735). Berlin: Springer.
- Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in international marketing*, 20, 277–319.
- Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Thousand Oaks, California: Sage.
- Hui, B. S., & Wold, H. (1982). Consistency and consistency at large of partial least squares estimates. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Part II* (pp. 119–130). Amsterdam: North Holland.
- Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: a review of four recent studies. *Strategic Management Journal*, 20(2), 195–204.
- Hult, G. T. M., Neese, W. T., & Bashaw, R. E. (1997). Faculty perceptions of marketing journals. *Journal of Marketing Education*, 19(1), 37–52.
- Hult, G. T. M., Reimann, M., & Schilke, O. (2009). Worldwide faculty perceptions of marketing journals: Rankings, trends, comparisons, and segmentations. *GlobalEDGE Business Review*, 3(3), 1–23.
- 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.
- Jakobowicz, E., & Derquenne, C. (2007). A modified PLS path modeling algorithm handling reflective categorical variables and a new model building strategy. *Computational Statistics & Data Analysis*, 51(8), 3666–3678.
- 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.
- Jedidi, K., Jagpal, H. S., & DeSarbo, W. S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science*, 16(1), 39–59.
- Johnson, M. D., Herrmann, A., & Huber, F. (2006). The evolution of loyalty intentions. *Journal of Marketing*, 70(2), 122–132.
- Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, 43(4), 443–477.
- Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294–316). Newbury Park: Sage.
- Jöreskog, K. G., & Wold, H. (1982). The ML and PLS techniques for modeling with latent variables: Historical and comparative aspects. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Part I* (pp. 263–270). Amsterdam: North-Holland.
- Kwon, H., & Trail, G. (2005). The feasibility of single-item measures in sport loyalty research. *Sport Management Review*, 8(1), 69–89.
- Lee, D. Y. (1994). The impact of firms' risk-taking attitudes on advertising budgets. *Journal of Business Research*, 31(2–3), 247–256.
- Lohmöller, J.-B. (1987). *LVPLS 1.8*. Cologne.
- Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Heidelberg: Physica.
- 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), 293–334.
- Malhotra, N. K. (2010). *Marketing research: An applied orientation*. Upper Saddle River: Prentice Hall.
- Marcoulides, G. A., & Saunders, C. (2006). PLS: a silver bullet? *MIS Quarterly*, 30(2), III–IX.
- Medsker, G. J., Williams, L. J., & Holahan, P. J. (1994). A review of current practices for evaluating causal models in organizational behavior and human resources management research. *Journal of Management*, 20(2), 439–464.
- Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw Hill.
- Raykov, T. (2007). Reliability if deleted, not “alpha if deleted”: evaluation of scale reliability following component deletion. *British Journal of Mathematical and Statistical Psychology*, 60(2), 201–216.
- Reinartz, W. J., Haenlein, M., & Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of Market Research*, 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., Ringle, C. M., & Sarstedt, M. (2010). Structural modeling of heterogeneous data with partial least squares. In N. K. Malhotra (Ed.), *Review of Marketing Research*. Armonk: Sharpe, 7, 255–296.
- Ringle, C., Wende, S., & Will, A. (2005). *SmartPLS 2.0 (Beta)*. Hamburg. (www.smartpls.de).
- Ringle, C. M., Götz, O., Wetzels, M., & Wilson, B. (2009). On the use of formative measurement specifications in structural equation modeling: A Monte Carlo simulation study to compare covariance-based and partial least squares model estimation methodologies. In *METEOR Research Memoranda (RM/09/014)*: Maastricht University.
- Ringle, C. M., Sarstedt, M., & Mooi, E. A. (2010a). Response-based segmentation using finite mixture partial least squares: theoretical foundations and an application to American Customer Satisfaction Index data. *Annals of Information Systems*, 8, 19–49.
- Ringle, C. M., Wende, S., & Will, A. (2010b). Finite mixture partial least squares analysis: Methodology and numerical examples. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields* (pp. 195–218). Berlin: Springer.
- Sackett, P. R., & Larson, J. R. (1990). Research strategies and tactics in I/O psychology. In M. D. Dunnette, P. L. Ackerman, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd edition, Vol. 1, pp. 419–488). Palo Alto: Consulting Psychology Press.
- Sarstedt, M., Becker, J.-M., Ringle, C. M., & Schwaiger, M. (2011a). Uncovering and treating unobserved heterogeneity with FIMIX-PLS: which model selection criterion provides an appropriate number of segments? *Schmalenbach Business Review*, 63(1), 34–62.
- Sarstedt, M., Henseler, J., & Ringle, C. M. (2011b). Multi-group analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results. *Advances in International Marketing*, forthcoming.
- Sarstedt, M., & Ringle, C. M. (2010). Treating unobserved heterogeneity in PLS path modeling: a comparison of FIMIX-PLS with different data analysis strategies. *Journal of Applied Statistics*, 37(7–8), 1299–1318.

- Sarstedt, M., & Wilczynski, P. (2009). More for less? A comparison of single-item and multi-item measures. *Die Betriebswirtschaft*, 69(2), 211–227.
- Shah, R., & Goldstein, S. M. (2006). Use of structural equation modeling in operations management research: looking back and forward. *Journal of Operations Management*, 24(2), 148–169.
- Shook, C. L., Ketchen, D. J., Hult, G. T. M., & Kacmar, K. M. (2004). An assessment of the use of structural equation modeling in strategic management research. *Strategic Management Journal*, 25(4), 397–404.
- Sirohi, N., McLaughlin, E. W., & Wittink, D. R. (1998). A model of consumer perceptions and store loyalty intentions for a super-market retailer. *Journal of Retailing*, 74(2), 223–245.
- Slack, N. (1994). The importance-performance matrix as a determinant of improvement priority. *International Journal of Operations and Production Management*, 14(5), 59–75.
- Sosik, J. J., Kahai, S. S., & Piovosio, M. J. (2009). Silver bullet or voodoo statistics? A primer for using partial least squares data analytic technique in group and organization research. *Group & Organization Management*, 34(1), 5–36.
- Steenkamp, J.-B. E. M., & Baumgartner, H. (2000). On the use of structural equation models for marketing modeling. *International Journal of Research in Marketing*, 17(2–3), 195–202.
- Steenkamp, J.-B. E. M., & van Trijp, H. C. M. (1991). The use of LISREL in validating marketing constructs. *International Journal of Research in Marketing*, 8(4), 283–299.
- Stewart, D. W. (2009). The role of method: some parting thoughts from a departing editor. *Journal of the Academy of Marketing Science*, 37(4), 381–383.
- Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society*, 36(2), 111–147.
- Tenenhaus, M., Amato, S., & Esposito Vinzi, V. (2004). A global goodness-of-fit index for PLS structural equation modeling. In *Proceedings of the XLII SIS Scientific Meeting* (pp. 739–742). Padova: CLEUP.
- Theoharakis, V., & Hirst, A. (2002). Perceptual differences of marketing journals: a worldwide perspective. *Marketing Letters*, 13(4), 389–402.
- Völckner, F., Sattler, H., Hennig-Thurau, T., & Ringle, C. M. (2010). The role of parent brand quality for service brand extension success. *Journal of Service Research*, 13(4), 379–396.
- Wetzels, M., Oderkerken-Schröder, G., & van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: guidelines and empirical illustration. *MIS Quarterly*, 33(1), 177–195.
- Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, & V. Capecchi (Eds.), *Quantitative sociology: International perspectives on mathematical and statistical modeling* (pp. 307–357). New York: Academic.
- Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observations: Part II* (pp. 1–54). Amsterdam: North-Holland.
- Wold, H. (1985). Partial least squares. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of statistical sciences* (pp. 581–591). New York: Wiley.
- Xu, X., Venkatesh, V., Tam, K. Y., & Hong, S.-J. (2010). Model of migration and use of platforms: role of hierarchy, current generation, and complementarities in consumer settings. *Management Science*, 56(8), 1304–1323.