How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research

Jose Benitez[a,b,*], Jörg Henseler[c,d], Ana Castillo[b], Florian Schuberth[c]

[a] *Rennes School of Business, Rennes, France*
[b] *Department of Management, School of Business, University of Granada, Granada, Spain*
[c] *Department of Design, Production and Management, Faculty of Engineering Technology, University of Twente, Enschede, the Netherlands*
[d] *Nova Information Management School, Universidade Nova de Lisboa, Lisbon, Portugal*

ARTICLE INFO

ABSTRACT

Partial least squares path modeling (PLS-PM) is an estimator that has found widespread application for causal information systems (IS) research. Recently, the method has been subject to many improvements, such as consistent PLS (PLSc) for latent variable models, a bootstrap-based test for overall model fit, and the heterotrait-to-monotrait ratio of correlations for assessing discriminant validity. Scholars who would like to rigorously apply PLS-PM need updated guidelines for its use. This paper explains how to perform and report empirical analyses using PLS-PM including the latest enhancements, and illustrates its application with a fictive example on business value of social media.

1. Introduction

Structural equation modeling (SEM) has become an important statistical tool in social and behavioral sciences. It is capable of modeling nomological networks by expressing theoretical concepts through constructs and connecting these constructs via a structural model to study their relationships [1]. In doing so, random measurement errors can be taken into account and empirical evidence for postulated theories can be obtained by means of statistical testing.

Two kinds of estimators for SEM can be distinguished: covariance-based and variance-based estimators. While covariance-based estimators minimize the discrepancy between the empirical and model-implied variance–covariance matrix of the observable indicators to obtain the model parameter estimates, variance-based estimators create linear combinations of the indicators as stand-ins for the theoretical concepts and subsequently estimate the model parameters. A widely used variance-based estimator is partial least squares path modeling (PLS-PM). Originally developed by Herman O.A. Wold [2] to analyze high-dimensional data in a low-structure environment, PLS-PM has become a full-fledged estimator for SEM over the past decade [3]. Consequently, PLS-PM has been applied in various fields of business administration research such as strategy [4], marketing [5], operations management [6], human resource management [7], finance [8], tourism [9], and family business [10].

For decades, PLS-PM has been the predominant estimator for structural equation models in the field of information systems (IS) (e.g., [11–15]). IS research usually incorporates complex research problems and questions that require conceptualization and operationalization of theoretical concepts, and investigation of their relationships. Current literature suggests two types of theoretical concepts: concepts from behavioral sciences and concepts from design science [16]. Theoretical concepts from behavioral research are assumed to cause observable indicators and their relationships, i.e., the theoretical concept is the common cause of observable indicators [17]. Typically, these concepts are operationalized by a measurement model. Extant literature suggests two types of measurement models: the reflective [1] and the causal–formative measurement model [18]. Both types of measurement models assume a causal relationship between the indicators and their construct, i.e., the latent variable. In contrast, theoretical concepts from design science, so-called artifacts, are human-made creations that are shaped and built by their ingredients to serve a certain goal [19]. Due to the constructivist nature of this type of theoretical concept, recent literature suggests to operationalize artifacts by the composite model [16]. In contrast to the measurement models, in the composite model, the indicators do not cause the construct, but combine to compose the construct. To highlight this aspect and to pronounce the difference to the latent variable, we refer to constructs that are composed of indicators *emergent variables* [20,21]. In summary, IS scholars can

Table 1
Substantial changes in understanding of PLS-PM.

Traditional view of PLS-PM	Up-to-date view of PLS-PM
1. PLS-PM should be used primarily for exploratory and early-stage research	PLS-PM can be used for various types of research, e.g., confirmatory and explanatory or predictive ([22], [23] [24],)
2. PLS-PM has advantages over covariance-based estimators when the sample size is small	PLS-PM can produce estimates even for very small sample sizes. However, as for other estimators, these estimates are generally less accurate than those obtained by a larger sample ([25] [26],). Hence, the justification of using PLS-PM with small sample sizes should be considered cautiously
3. PLS-PM can only estimate recursive structural models	PLS-PM can also consistently estimate non-recursive structural models by using, e.g., 2SLS or 3SLS instead of OLS (Dijkstra and Henseler 2015b, [27])
4. Model identification plays no role when employing PLS-PM	First, PLS-PM always estimates an underlying composite model, regardless of whether the model consists of latent variables; identification rules of composite models must thus be taken into account ([28] [134],). Consequently, model identification is also important in the case of PLS-PM
5. PLS-PM has greater statistical power than the maximum-likelihood (ML) estimator	This statement is based on inconsistent parameter estimates and has been shown to be invalid [30]. Furthermore, an estimator has no statistical power (one refers to its efficiency, or accuracy, in estimating parameters, usually expressed by the standard error); only a statistical test can be assessed in terms of its statistical power
6. Mode A can be used to consistently estimate reflective measurement models	Regardless of the mode used, PLS-PM creates linear combinations of observed indicators (composites) as proxies for the theoretical concepts ([31] [32],). Therefore, to consistently estimate models containing latent variables, one must correct for attenuation of the construct scores correlations. In the context of PLS-PM, this procedure is known as PLSc (Dijkstra and Henseler 2015a, 2015b)
7. Mode B can be used to estimate causal–formative measurement models consistently	Mode B is just another way to obtain weights to build composites; hence, it does not consistently estimate causal–formative measurement models. However, causal–formative measurement models can be estimated by means of a MIMIC model ([18] [3],)
8. The overall fit of models estimated by PLS-PM cannot be assessed	The overall model can be assessed in two non-exclusive ways ([22], Dijkstra and Henseler 2015b): (1) bootstrap-based tests for overall model fit, and (2) measures of overall model fit. Both assess the discrepancy between the empirical and the model-implied indicator variance–covariance matrix. While the latter is based on heuristic rules, the former is based on statistical inferences
9. Reliability of construct scores obtained by PLS-PM should be assessed using Cronbach's α and Dillon–Goldstein's ρ (also called Jöreskog's ρ or composite reliability)	Currently, Dijkstra–Henseler's ρ_A is the only consistent reliability coefficient for PLS-PM construct scores [16]. Dillon–Goldstein's ρ and Cronbach's α indicate the reliability of sum scores. While Cronbach's α is based on the indicator variance–covariance matrix, Dillon–Goldstein's ρ is based on the factor loadings. Therefore, for the calculation of Dillon–Goldstein's ρ, consistent factor loading estimates should be used. Moreover, Cronbach's α assumes equal population covariances among the indicators of one block; an assumption that is likely not met in empirical research. However, it can be used as a lower bound to reliability ([33] [34],)
10. Discriminant validity should be examined using the Fornell–Larcker criterion	The HTMT [35] should be considered to assess discriminant validity ([36] [37],)

operationalize a theoretical concept in three different ways in their model: reflective and causal–formative measurement model (usually employed for behavioral concepts), and composite model (usually employed for artifacts). See Table 1 in Henseler (2017) for a detailed explanation of the different types of operationalization.

In recent years, PLS-PM has become the subject of scholarly debate. Proponents called PLS-PM a "silver bullet" [38], while opponents criticized PLS-PM's inconsistency for latent variable models and the absence of a test for overall model fit (e.g. [39],). This debate has stimulated the development of several enhancements to PLS-PM. These include consistent PLS (PLSc) to consistently estimate linear and non-linear latent variable models ([40] [41],); a bootstrap-based test to statistically assess overall model fit [42]; measures of overall model fit, such as the standardized root mean squared residual (SRMR), based on heuristic rules to evaluate overall model fit [22]; and the heterotrait-to-monotrait (HTMT) ratio of correlations as a criterion to assess discriminant validity [35]. As a result, PLS-PM has become a full-fledged estimator to SEM that can deal with reflective and causal–formative measurement models as well as composite models. Moreover, it can be applied to confirmatory, explanatory, exploratory, descriptive, and predictive research [24].

For the field of IS to benefit from these methodological and conceptual achievements in PLS-PM, IS scholars need guidelines for their empirical studies that incorporate all these new developments and recently obtained insights. Some of the guidelines papers on PLS-PM in the IS literature were published before 2013 – i.e., before the debate and resulting enhancements (e.g., [12,43–45]). Although several recently published textbooks and articles (e.g. [46–48],) have provided guidelines for causal research that cover some of the latest

enhancements in PLS-PM, neither of these prior PLS-PM guidelines for causal research covered the full range of recent developments.

To fill this gap on guidelines in the current IS literature, this study provides updated guidelines for using PLS-PM in causal research (confirmatory and explanatory research), employing all the most recently proposed standards. In so doing, the paper addresses *why* and *how* to perform and report PLS-PM estimation in confirmatory and explanatory IS research. In confirmatory IS research, the scholar aims to understand the causal relationships between theoretical concepts of interest for the IS community. In doing so, the scholar aims to confirm a postulated theory, i.e., obtain empirical evidence for his/her description of the working mechanism of the world. This is tried to be achieved by imposing testable restrictions on the indicator variance–covariance matrix, e.g., by fixing path coefficients to a certain value, assuming that the correlation between two indicators is the result of an underlying latent variable like in the classical reflective measurement model, or in the composite model that the correlations of the indicators forming an emergent variable with a variable not forming the emergent variable are proportional. The dominant statistical tool in the context of confirmatory research is the test for overall model fit. Testing model fit only makes sense if the number of correlations among observable variables exceeds the number of model parameters to be estimated, i.e., it is indispensable to have a certain amount of parsimony (a positive number of degrees of freedom in the sense of SEM).

As in confirmatory research, in explanatory IS research, the analyst aims to understand the causal relationship among the theoretical concepts. However, this type of research primarily focuses on the explanation of a specific phenomenon which is treated as a dependent variable in the model. In doing so, the primary focus is on the

coefficient of determination (R^2) and the significance of path coefficient estimates. Such models can be saturated (i.e., have zero degrees of freedom in the sense of SEM). Although the two types of research can be theoretically distinguished, in empirical IS research, scholars very often combine confirmatory and explanatory IS research, e.g., testing the measurement model (confirmatory research) and focusing on the explanation of a specific construct in structural model (explanatory research). This paper explains why and how to perform and report empirical analyses using PLS-PM in causal IS research following the latest enhancements, and illustrates this analysis with a fictive example on business value of social media.

2. Foundations of PLS-PM

In its current form, PLS-PM is a full-fledged variance-based estimator for SEM that can estimate linear, non-linear, recursive, and non-recursive structural models ([40] [42],). Moreover, it is capable of dealing with models that contain emergent and latent variables [41], second-order emergent variables built by latent variables [49], and ordinal categorical indicators [29]. It can incorporate sampling weights known as weighted partial least squares (WPLS, see [50]), deal with correlated measurement errors within a block of indicators [51], and address multicollinearity among the constructs in the structural model [52]. It can also be used for multiple group comparison ([53], [54],), and potential sources of endogeneity can be addressed [55]. Finally, important-performance map analysis can be used to illustrate the results of the structural model [56]. For a recent overview of the methodological research on PLS-PM, we refer to [57].

2.1. Model specification

To employ PLS-PM, scholars must transfer their proposed theory into a statistical model [58]. In the context of SEM, this means that the theoretical concepts and their hypothesized relationships must be transferred into a structural model. "Theoretical concepts refer to ideas that have some unity or something in common. The meaning of a theoretical concept is spelled out in a theoretical definition" [59]. We distinguish between two types of theoretical concepts: behavioral concepts, and design concepts, so-called artifacts. Typically, theoretical concepts are represented by constructs in the structural model [32]. Although constructs and latent variables are often equated [60], we deliberately distinguish between a latent variable, i.e., a construct that represents a behavioral concept, and an emergent variable, i.e., a construct that represents an artifact. The operationalization of theoretical concepts, i.e., the specification of the theoretical concepts in the structural model, requires special attention because estimates are likely to be inconsistent if a concept's operationalization is not in accordance with the concept's nature [61].

PLS-PM can deal with two kinds of constructs: emergent variables and latent variables. Latent variables refer to variables that are not directly observed but instead inferred through a measurement model from other observed variables (directly measured; [46,62]). They usually represent theoretical concepts of behavioral research such as personality traits, individual behavior, and individual attitude [63]. This theoretical reasoning rests on the assumption that behavioral concepts of interest exist in nature, irrespective of scholarly investigation [64]. The existing literature proposes two ways to measure behavioral concepts [59]: reflective and causal–formative measurement model.

The reflective measurement model – also known as the common factor model – is grounded in the true score theory [65]. It assumes that a set of indicators is a measurement error–prone manifestation of an underlying latent variable [66]. Some indicators can thus be interchanged without altering the meaning of the latent variable. As the measurement errors of a block of indicators are usually assumed to be uncorrelated and independent of the latent variable, the reflective

measurement model imposes restrictions on the variance–covariance matrix of indicators belonging to one latent variable. In its classical form, the correlations among the indicators of one block are zero when controlled for the latent variable, also known as the axiom of local independence [67]. This fact is typically exploited to draw conclusions about the existence of the latent variable.

Besides the reflective measurement model, the literature proposes the causal–formative measurement of behavioral concepts [68,69]. In contrast to the reflective measurement model, the causal–formative measurement model reverses the direction of causality between the indicators and the construct and assumes that the observed indicators cause the latent variable. This model thus does not restrict the covariances of the indicators belonging to one block. The remaining causes not represented by the indicators are captured in an error term, which is by assumption uncorrelated with the causal indicators. Although a violation of this assumption, i.e., omission of causal indicators, leads to biased parameter estimates of the causal indicators, recent literature shows that the meaning of the latent variable is not affected by omitting causal indicators and the remaining model parameters can be consistently estimated [70]. However, the causal–formative measurement model is not identified on its own, i.e., the model parameters cannot be uniquely retrieved from the population indicator variance–covariance matrix [71,72]. To obtain an identified causal–formative measurement model, the latent variable must be connected to at least two other variables not affecting the latent variable [18], for example, using a multiple-indicators, multiple-causes (MIMIC) model.

Typical examples of behavioral concepts in IS that have been operationalized by a measurement model are behavioral intention to use information technology (IT), and IT interaction behavior. Behavioral intention to use IT indicates the degree to which a person has formulated conscious plans to perform or not to perform a specified future behavior involving IT use. This concept has been operationalized by a reflective measurement model in past IS research using the following items: user's intention, prediction, and plan to use IT in future months (e.g., [73,74]). IT interaction behavior refers to the user's interaction with IT to accomplish an individual or organizational task. This concept has been operationalized by a causal–formative measurement model in past IS research. For example, Barki et al. [75] employed a MIMIC model to operationalize IT interaction behavior using six tasks (causes) that motivated users to interact with IT (problem solving, justifying decisions, exchanging information with people, planning or following up, coordinating activities, and serving customers); and two measurements of this behavior using two reflective indicators (importance of IT and time invested using IT).

Emergent variables are an alternative representation of theoretical concepts [20,21]. They have been recently referred in empirical IS research as "composite constructs" (e.g. [76],). Although these labels could be used interchangeably, we recommend using the term "emergent variable" to highlight that the construct emerges from the indicators. Emergent variables can help model artifacts [3,16]. An artifact is a human- or firm-made object composed of its ingredients. Thus, in contrast to behavioral concepts, they are not assumed to exist in nature, but are products of theoretical thinking and/or theoretically justified constructions usually made to fulfill a certain purpose. To operationalize these human- or firm-made concepts, the composite model can be employed [16]. Examples from the IS research are IT capability and IT ambidexterity [77,78].

The composite model can be understood as a recipe for how ingredients (the components) should be mixed and matched to build the artifact. The composite model assumes a definitorial rather than a causal relationship between indicators and the emergent variable ([63], 2017). In the classical composite model, the indicators forming an emergent variable are assumed to be free of measurement errors. In contrast to the reflective measurement model, the composite model imposes no restrictions on the covariance structure of indicators belonging to the same construct. The reflective measurement model is

thus nested within the composite model, as the composite model relaxes the assumption that all covariation among a block of indicators is explained by one latent variable [22]. Yet, the composite model constraints the correlations between the indicators forming an emergent variable and variables not forming the emergent variable, i.e., they are proportional [29]. Similar to the causal–formative measurement model, the composite model is not identified when isolated in the structural model. To ensure identification, a necessary condition is that each emergent variable must be linked to at least one variable not forming the emergent variable [28,134].

Because the artifact as a type of theoretical concept was introduced only recently, it is helpful to illustrate this type of theoretical concept with an example. Based on theory, bread is made from wheat, water, salt, and yeast. Although the correlations between the amounts of wheat, water, salt, and yeast in a sample of loaves of bread are likely to be high, one would not conclude that bread is something that should be measured, i.e., that bread causes (or is caused by) wheat, water, salt, and yeast. Rather, wheat, water, salt, and yeast are the simple entities (ingredients) combined to form the emergent variable representing the artifact we call bread. Clearly, the temporal precedence of the ingredients also suggests that bread cannot be the common cause of its ingredients.

Because IS Science analyzes and aims at explaining how IT affects organizations, individuals, and society, artifacts play a pivotal role in IS research. For example, the theoretical concept IT infrastructure capability refers to a firm's ability to use and leverage its IT resource infrastructure for business activities [79–82]. IT infrastructure capability is a "human-made/firm-made" concept that can be operationalized by the composite model [76,81,83]. Of course, no single "true" recipe exists for creating this artifact. Just as different bakeries can produce different types of bread or different breweries produce different types of beer, different scholars can produce different recipes for the same concept. The beer analogy can be extraordinarily instructive. Different recipes exist worldwide to design and manufacture beer. For example, Spanish breweries use one recipe, German breweries another. Recipes can even vary by region within a country. Such diversity makes each recipe an idiosyncratic way to understand and design beer, but all of these recipes ultimately produce beer.

For example, based on Melville et al. [84] study, Ajamieh et al. [81] define the artifact IT infrastructure capability as composed of IT technological infrastructure capability, IT managerial infrastructure capability, and IT technical infrastructure capability. Further, some prior IS research [85,86] considers IT capability – a concept similar to IT infrastructure capability – as composed of IT technical infrastructure, human IT resources, and IT-enabled intangibles. IT infrastructure flexibility and post-merger and acquisition (M&A) IT integration capability are two examples of artifacts recently considered in IS research [83]. IT infrastructure flexibility refers to the capability of the infrastructure to adapt to environmental changes. A flexible firm IT infrastructure has the following characteristics: IT compatibility, IT connectivity, modularity, and IT personnel skills flexibility [83]. Similarly, post-M&A IT integration capability is the firm's ability to integrate the IT technical infrastructure, IT personnel, and IT and business processes of the target/acquired firm with the IT technical infrastructure, IT personnel, and IT and business processes of the acquirer after an M&A [83]. Thus, post-M&A IT integration capability can be understood as an artifact built by integrating IT technical infrastructure, IT personnel, and IT and business processes. These are two examples of artifacts that have recently been examined in the field of IS.

While this study argues for the use of the composite model to operationalize artifacts, recently it has been suggested to employ the composite model to operationalize behavioral concepts [58]. This notion assumes that both latent and emergent variables serve as a proxy for behavioral concepts [61]. Following this reasoning, the validity gap occurs between the concept and its construct and not between the construct and the observable variables [32].

Once the theoretical concepts are operationalized, the constructs representing the theoretical concepts can be related via the structural model. The structural model typically represents the core of the theory proposed. The structural model generally consists of a set of regression equations, illustrating the relationship hypothesized between the theoretical concepts. In each equation, a dependent construct is explained by one or more independent constructs. Because a dependent construct is typically not fully explained by its independent constructs, an error term accounts for the remaining variance in the dependent construct. By assumption, the error term is independent of the explanatory constructs of its equation. To avoid violating this assumption, in causal IS research, the scholar should make every effort to include all relevant constructs (those that affect the dependent construct and correlate with at least one explanatory construct in the corresponding equation). Otherwise, the path coefficient estimates obtained by ordinary least squares (OLS) suffer from omitted variable bias [87]. One potential way to address this problem of endogeneity is to use the two-stage least squares (2SLS) estimator for the structural model ([42] [27,55],). In the following, we consider only recursive structural models, structural models without feedback loops, and/or correlated error terms.

2.2. Parameter estimation

In its current form, PLS-PM estimates model parameters in three steps. In the first, the iterative PLS-PM algorithm determines the weights to create scores for each construct (latent variables and emergent variables; [88]). As construct scores of latent variables contain measurement errors, the second step corrects for attenuation in correlations between latent variables. In doing so, PLSc divides the construct scores correlations by the geometric mean of the constructs' reliabilities [41], making the main outcome of the second step a consistent construct correlation matrix. Finally, the third step estimates the model parameters (weights, loadings, and path coefficients). Based on the consistent construct correlation matrix, OLS can be used to estimate the path coefficients of recursive structural models. In case of non-recursive structural models, the 2SLS or three-stage least squares (3SLS) estimator can be used, instead of the OLS estimator, to obtain consistent path coefficient estimates ([42] [27,83],).

2.3. Substantial changes in the understanding of PLS-PM

In recent years, PLS-PM practices have been examined, debated, and improved. The recent literature on PLS-PM has been thus substantially changed and improved, requiring that we identify the changes in the understanding and practice of PLS-PM. Table 1 summarizes these changes in the understanding of PLS-PM in the context of confirmatory and explanatory research.

Traditional view 1: PLS-PM should be used primarily for exploratory and early-stage research. Although PLS-PM was originally developed for exploratory research [2], enhancements such as PLSc and the bootstrap-based test for overall model fit make PLS-PM suitable for causal research, i.e., confirmatory and explanatory research. However, as originally developed, PLS-PM can also be applied in descriptive and predictive research [23,24].

Traditional view 2: PLS-PM has advantages over covariance-based estimators in the case of small sample sizes. The application of PLS-PM has often been justified by the size of the investigated sample [26]. It is true that PLS-PM is capable of estimating models with more parameters than observations because it only estimates partial model structures, but as with every other statistical method, the standard errors of the estimates increase as the sample size decreases. Therefore, justifying the use of PLS-PM due to small sample sizes should be considered cautiously. In this sense, claiming that PLS-PM is particularly suitable for small sample sizes can be regarded as problematic [26]. However, in case of pure emergent variable models and small sample size constellations, PLS-PM performs superior focusing on accuracy in the

estimation of path coefficients compared to other variance-based estimators [89], i.e., generalized structured component analysis (GSCA [90],) and regression with sum scores.

Traditional view 3: PLS-PM cannot be used for non-recursive models. Although current user-friendly software packages do not yet implement approaches to analyzing non-recursive models, the assumption of recursivity can be relaxed by estimating the structural model parameters using 2SLS or 3SLS instead of OLS ([42,27]). Another approach to estimate non-recursive structural models involves using the construct scores (in the case of emergent variables) or the disattenuated construct correlation matrix (in the case of latent variables) obtained by PLS-PM as input for the full-information maximum-likelihood (FIML) estimator (e.g. [83],).

Traditional view 4: Model identification plays no role when employing PLS-PM. PLS-PM always employs composites to estimate the model, whether or not the theoretical concepts are operationalized by a measurement model or a composite model. Therefore, the identification rules for composite models must be applied [28,134]. In addition to normalization of the weight vector such as fixing the variance of each composite to one, it must be ensured that each construct is connected (by means of a non-zero path) with at least one other construct in the model to ensure that the weights can be uniquely retrieved from the population indicator variance–covariance matrix. Although all weight vectors are scaled and no construct is isolated in the structural model, the signs of the weights of a block of indicators are still ambiguous. The dominant indicator approach is thus recommended to fix construct scores' orientation and thereby uniquely determine the weights [3]. Additionally, the structural model must be identified. As long as one only considers recursive structural models with uncorrelated error terms, identification is straightforward as they are always identified [1].

Traditional view 5: PLS-PM has greater statistical power than the ML estimator. In estimating latent variable models, Reinartz et al. [91] claim that the power of statistical testing is higher when PLS-PM estimates are employed than when ML estimates are used. However, these findings are highly questionable, as they are based on traditional PLS-PM, which is known to produce inconsistent parameter estimates for latent variable models. In line with Goodhue et al. [30], who show that this alleged higher power goes along with an inflated type I error, we conclude that preferring PLS-PM over the ML estimator due to efficiency is not a valid argument for latent variable models. Similar findings observed for GSCA are applicable to PLS-PM [92]. For emergent variable models, however, traditional PLS-PM has shown favorable properties among variance-based estimators, i.e., GSCA, sum scores, and PLS-PM [89].

Traditional view 6: Mode A can be used to consistently estimate reflective measurement models. In its most modern appearance, PLS-PM can deal with models containing both emergent and latent variables. Because PLS-PM inherently estimates composite models [31,32], it is the estimator of choice for models containing emergent variables only [61]. In PLS-PM, composite models can be consistently estimated by Mode B [28]. To obtain consistent parameter estimates for reflective measurement models, PLSc should be used. Estimates obtained by traditional Mode A, or Mode B and C as well, suffer from the attenuation bias [42]. In contrast, PLSc produces consistent and asymptotically normal estimates for reflective measurement models by combining Mode A estimates with a correction for attenuation [41]. Consequently, the development of PLSc [41] enables PLS-PM to analyze models containing both emergent and latent variables. However, scholars should use Mode B in PLS-PM (instead of PLSc) when they estimate pure emergent variable models as PLSc has shown to produce biased estimates in this situation [61].

That being said, in the case of pure latent variable models, covariance-based estimators are preferred, as they are consistent and asymptotically efficient. However, the availability of asymptotically efficient estimators does not mean that scholars cannot use PLS-PM to estimate models of this kind. First simulation studies have investigated the performance of PLSc to other estimators in this situation and in fact, its usage for pure latent models is considered as acceptable and its bias for finite samples has been evidenced of little practical relevance assuming that the model is correctly specified ([42,61,93]).

Traditional view 7: Mode B can be used to estimate causal–formative measurement models consistently. Mode B, or so-called regression weights, cannot be used to consistently estimate causal–formative measurement models, as this kind measurement model is not identified by its own [94]. However, by the development of PLSc, PLS-PM can consistently estimate causal–formative measurement models by means of the MIMIC model [16].

Traditional view 8: Overall fit of models estimated by PLS-PM cannot be assessed. Due to recent developments in the context of PLS-PM, the overall fit of models estimated by PLS-PM can be assessed in two non-exclusive ways: (1) by a bootstrap-based test for overall model fit [42], and (2) by measures of overall model fit such as the SRMR [22]. Both ways assess the difference between the empirical indicator variance–covariance matrix and the estimated model-implied counterpart. While the empirical indicator variance–covariance matrix contains the variances and covariances of the indicators based on the sample, the estimated model-implied counterpart contains the variances and covariances of the indicators implied by the model structure based on the estimated model parameters. Typically, the discrepancy between the two matrices is measured by the squared Euclidean distance (d_{ULS}), the geodesic distance (d_G), and the SRMR. The bootstrap-based test for overall model fit relies on a bootstrap procedure to obtain the reference distribution of the distance measures under the null hypothesis that the population indicator variance–covariance matrix equals the model-implied counterpart [95]. Assuming a 5% level of significance, a discrepancy value larger than the 95% quantile of the corresponding reference distribution leads to rejection of the null hypothesis. In addition to the bootstrap-based test for overall model fit, the values of the distance measures can be compared to threshold values recommended by the literature to assess overall model fit. Measures of fit are thus based on heuristic rules rather than on statistical inference. Moreover, the suggested thresholds for the measures of overall model fit in the context of PLS-PM, e.g., 0.080 for the SRMR, are preliminary and need to be examined in more detail in future research.

Traditional view 9: Reliability of the construct scores obtained by PLS-PM should be assessed using Cronbach's α and Dillon–Goldstein's ρ (also called Jöreskog's ρ or composite reliability). Traditionally, the literature recommended determining the reliability of PLS-PM construct scores through Cronbach's α and Dillon–Goldstein's ρ [12]. However, considering this recommendation, several aspects of these two measures have been widely neglected. First, Cronbach's α and Dillon–Goldstein's ρ both assess the reliability of sum scores (construct scores obtained by equally weighted indicators) created for the latent variable. However, PLS-PM allows the indicator weights used for the calculation of the construct scores to vary such that indicators with a smaller amount of random measurement error take on greater weight than indicators containing a larger amount of random measurement error. Consequently, the PLS-PM construct scores contain less measurement error and are generally more reliable than sum scores [22]. Second, Cronbach's α assumes tau-equivalence, i.e., equal population covariances among the indicators belonging to one latent variable, an assumption that is rarely met in empirical research [34]. While Cronbach's α can be calculated based on the sample variance–covariance matrix, Dillon–Goldstein's ρ is based on factor loadings. Because traditional PLS-PM is known to produce inconsistent factor loading estimates, Dillon–Goldstein's ρ should be based on consistent factor loading estimates obtained by PLSc. Furthermore, as the assumptions of Cronbach's α and Dillon–Goldstein's ρ are likely to be violated in empirical research, their use cannot be recommended. However, the reliability obtained by Cronbach's α can be regarded as a lower bound [33]. To consistently estimate the reliability of latent variable scores obtained by

PLS-PM's Mode A, Dijkstra–Henseler's ρ_A should be used [41].

Traditional view 10: Discriminant validity should be examined by the Fornell–Larcker criterion. Although the Fornell–Larcker criterion [96] had been long recommended to assess discriminant validity of latent variables [12], it is ineffective in combination with traditional PLS-PM because it relies on consistent factor loading estimates [22]. To overcome this drawback, the HTMT was developed to assess discriminant validity in the case of variance-based estimators [35]. The HTMT can be assessed in two ways: (1) by comparing it to a threshold value, and (2) by constructing a confidence interval to examine whether HTMT is significantly smaller than a certain threshold value ([35] [37],). For the first approach, simulation studies suggest a threshold value of 0.90 if constructs are conceptually very similar or 0.85 if the constructs are conceptually more distinct ([35–37]). For the second approach, prior methodological research has suggested to examine whether HTMT is significantly smaller than 1 [35] or below other smaller values, e.g., 0.85 or 0.90 [37] [37]. conclude that HTMT is a reliable tool for assessing discriminant validity, whereas the Fornell–Larcker criterion has limitations that do not justify its reputation for rigor and its widespread use in empirical research.

3. An illustrative example

3.1. Description of the example

We provide an illustrative IS example to present the latest enhancements of PLS-PM. Fig. 1 displays the proposed research model to be estimated and tested. For this purpose, we use a simulated dataset of 300 observations, where each observation represents a firm – the unit of analysis in the example. Because we use a simulated dataset, the obtained results are not scientifically relevant and any comparison of our results to results of other empirical studies is only made for purely illustrative purposes.

Social executive behavior is the positive/negative behavior of the firm's top managers towards the firm's use of social media for business activities. Social employee behavior is the positive/negative behavior of the firm's employees towards the firm's use of social media for business activities. Social media capability refers to the firm's ability to use and leverage external social media platforms purposefully to execute business activities [77,97]. Business process performance is the firm's

relative performance in key business processes as compared with its key competitors [98]. Fig. 1 presents the research model of the example. Based on prior IS research on social media in organizations [77,99], it is assumed that social executive behavior and social employee behavior positively affect development of a firm's social media capability, which, in turn, may positively influence firm's business processes performance.

The research model represents the theory proposed by an author/team to be tested empirically. It illustrates how the theoretical concepts are operationalized, i.e., how the indicators are related to the constructs representing the theoretical concepts, and how these constructs are connected. It usually includes several hypotheses to be tested. Based on prior literature and anecdotal evidence from the real world, authors should explain one by one why the hypothesized relationships are included and state expectations about their signs. These explanations are omitted from this article because theoretical explanation of the relationships included in the example is beyond the paper's scope. In our example, the following three hypotheses are tested:

Hypothesis 1 (H1). *Social executive behavior has a positive impact on the development of social media capability.*

Hypothesis 2 (H2). *Social employee behavior has a positive impact on the development of social media capability.*

Hypothesis 3 (H3). *Social media capability has a positive impact on business process performance.*

Although prior IS studies using PLS-PM have investigated more complex models (e.g., including a greater number of constructs, second-order constructs, moderation effects), the presented research model seems reasonable for our purposes due to the following three reasons: (1) the goal of our study is to provide guidelines for using PLS-PM in causal IS research (confirmatory and explanatory), employing the most recently proposed standards. In sake of brevity, parsimony, and pedagogical illustration for IS scholars, we think, in line with Occam's razor, the simpler the research model is, the better. "Parsimonious yet well-fitting models are more likely to be scientifically replicable, explainable" [100]. "Parsimony is also regarded by many social scientists as an important ingredient in theory development (e.g. [101,102],), precisely because it 'explains much by little' ([103]; p.153)" [100]; (2) the considered model contains both latent variables (ovals) and emergent variables (hexagons), and therefore, presents a situation in which PLS-
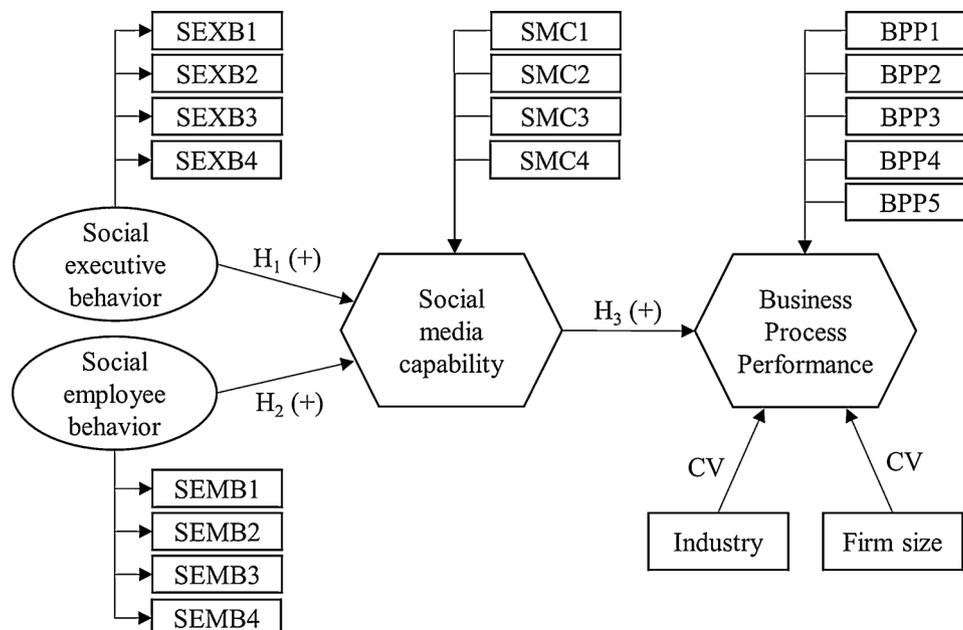

Fig. 1. Research model (CV = Control variables).

J. Benitez, et al.

PM can leverage its full capacities; and (3) the research model is theoretically positioned in IS literature on business value of IT, where the research models are usually parsimonious (e.g. [104,105],).

Theoretical concepts of behavioral research such as personality traits, individual behavior, and individual attitude are usually represented as latent variables [106]. Because social executive behavior and social employee behavior indicate types of individual behavior and attitude, the two theoretical concepts were operationalized by reflective measurement models. The ovals represent the latent variables and the connected rectangles their indicators. Social executive behavior and social employee behavior were each measured by four indicators, (SEXB1-SEXB4) and (SEMB1-SEMB4), respectively. To obtain consistent estimates, the reflective measurement models were estimated by PLSc [41].

In contrast, the theoretical concepts social media capability and business process performance were considered as artifacts designed by firms, executives, and/or employees. To operationalize these theoretical concepts, the composite model was employed. In doing so, social media capability is assumed to be composed of the following ingredients: Facebook, Twitter, corporate blog(s), and LinkedIn capabilities [77], which are the ingredients that shape social media capabilities and are lower-order capabilities. IS scholars and analysts from other contexts (e.g., China) might consider the social media WeChat (capability) as a key ingredient of social media capability and might remove other, less relevant social media capabilities for Chinese firms. This illustrates the potential of including/studying different artifacts to investigate the same phenomenon of interest for firms and society. The artifact business process performance was also operationalized by a composite model. It comprises supplier relations, product and service enhancement, production and operations, marketing and sales, and customer relations (Tallon and Pinsoneault 2011). Fig. 2 illustrates how the artifact social media capability was operationalized. The hexagon represents the construct, i.e., the emergent variable, while the rectangles represent ingredients forming the construct.

Besides the variables of main interest, firm size and industry were included as control variables in the structural model to control for effects of extraneous variables [80,81]. Firm size was modeled as a single-indicator composite to account for the role of different firm sizes in explaining business process performance through the natural logarithm of the number of employees [76]. Due to the skewed distribution, it is advisable to also apply the logarithm when the firm size is measured through sales or total assets. Industry was incorporated as a composite to control for an overall industry effect on business process performance and was shaped by three indicators, i.e., industry groups 1–3. The three industry group dummies indicate whether an observation belongs to industry 1, 2, or 3. Each industry assigns 0 if the observation does not belong to the industry and 1 if the observation does. For example, the variable industry group 1 will have a value of 0 for firms that do not belong to industry group 1 and a value of 1 for firms that belong to industry group 1. Although the dataset consists of four different industries, industry group 4 was not included to avoid perfect multicollinearity. Therefore, group 4 became the reference category. The
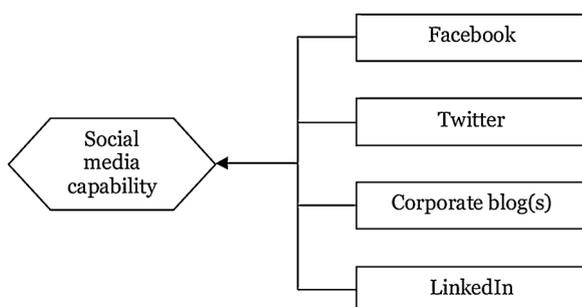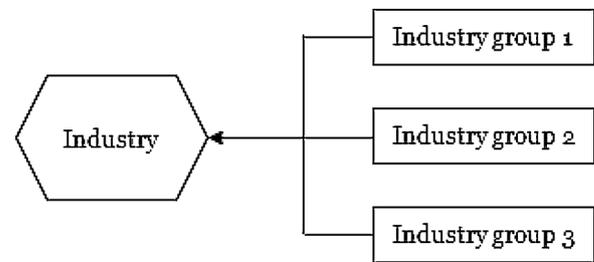

Fig. 3. Modeling a nominal control variable.
Note: Industry group 4 is the reference group.

weights of the industry composite can be interpreted as a simple contrast, i.e., the difference in contribution to the total industry effect between the industry considered and the reference industry. Fig. 3 presents how industry, a nominal control variable, was included in the structural model. IS scholars can use the dominant or the most important industry as the reference group.

3.2. Statistical power analysis

A power analysis should typically be conducted before data collection. It gives insight into the minimum sample size required to obtain sufficient statistical accuracy to detect effects of interest existing in the population. The power of a statistical test is the probability of rejecting the false null hypothesis correctly, that is, of finding an effect in the sample if it indeed exists in the population [107]. Power analysis can be conducted in two ways: (1) using heuristic rules such as Cohen's power tables and the inverse square root method [108,109], and (2) conducting a Monte Carlo simulation study [110]. The 10-times rule [111] or the minimum R^2 rule is no longer recommended to estimate the minimum sample size [26,46,109].

To apply Cohen's power tables for multiple regression analysis, four parameters must be considered: effect size (the extent to which the path coefficient/weight exists in the population), power (probability of rejecting the false null hypothesis correctly), significance level (probability of rejecting the true null hypothesis incorrectly), and the number of independent variables of the equation containing the considered path coefficient/weight. Once these values are determined, Cohen's power tables can be used to approximate the minimum required sample size in order to achieve a certain power level. To determine the number of required observations, analysts can assume a small effect size ($0.020 \leq f^2 < 0.150$) for a more conservative approximation or a medium to large effect size ($0.150 \leq f^2 < 0.350$ or $f^2 \leq 0.350$) for a more optimistic approximation of the required sample size. The statistical power is usually set to 0.8, and a significance level of 0.05 is assumed [107]. Often the equation with the highest number of independent variables is considered to determine the minimum number of observations to reliably detect an effect. In our example, the composite model for business process performance has the highest number of independent variables (supplier relations, product and service enhancement, production and operations, marketing and sales, and customer relations) in an equation. Cohen's power tables suggest a minimum sample size of 91 observations assuming a medium effect size ($f^2 = 0.150$), statistical power of 0.8, and significance level of 0.05 [108]. Considering the outcomes of the power analyses for our example, a sample size of 300 seems adequate to detect the effects of interests. The inverse square root method assumes that the estimates are standard normally distributed, and approximates the standard error using $\sqrt{N}$. Assuming a 5% significance level, the required sample size to obtain a statistically significant effect ($\hat{N}$), if it exists in the population, can be approximated by $\hat{N} > \left(\frac{2.486}{|\beta|_{min}} \right)^2$, where $|\beta|_{min}$ represents the minimum magnitude of the coefficient considered.

In addition to considering heuristic rules, IS scholars can conduct a

Fig. 2. Operationalization of the artifact social media capability.

Monte Carlo simulation to examine the sample size required to reliably detect effects that exist in the population. A population model with the same structure as the estimated model must be specified and all population parameter values need to be determined. In the second step, the model is estimated several times and the rejection rates of the null hypothesis significance test for the coefficients under examination are considered, i.e., the statistical power. The appealing property of this approach is that it can take into account various aspects of the model and indicators incorporated, such as sample size, number of indicators, their distribution, and magnitude of the effect. Moreover, sensitivity analyses can be conducted by changing the assumed population model to see how these changes affect the statistical power. While guidelines have been proposed for pure latent variable models in the context of PLS-PM [110], development of guidelines for models containing emergent variables is still an open issue.

3.3. Estimation

Various software packages – such as PLS-Graph [112], SmartPLS [113], WarpPLS [114], XLSTAT-PLS [115], and ADANCO [116] – can be used to estimate the model with PLS-PM. We used ADANCO 2.0.1 Professional for Windows (http://www.composite-modeling.com/) [116] to estimate the empirical example. In the following, we used Mode B to estimate composite models and PLSc to estimate reflective measurement models. Moreover, we used the factor weighting scheme for inner weighting and statistical inferences were based on the bootstrap procedure, relying on 4999 bootstrap runs.

Prior to model estimation, analysts should set a dominant indicator in each composite and reflective measurement model. As the signs of the weight and factor loading estimates of a block of indicators are ambiguous, the dominant indicator is used to dictate the orientation of a construct. A dominant indicator that is expected to positively correlate with the construct is preferable. Face validity can be used to select the dominant indicator – the indicator that is theoretically most relevant and thus expected to positively correlate with the construct. In our example, we chose SEXB2, SEMB2, SMC1, and BPP4 as dominant indicators.

Before the model assessment, the researcher has to ensure that the estimation is technically valid, i.e., that the estimation is admissible and no Heywood case has occurred [117]. In doing so, he/she needs to investigate whether the PLS-PM algorithm has properly converged. Additionally, in particular in the context of PLSc, he/she needs to ensure that the construct correlation and the model-implied indicator correlation matrix are valid, i.e., positive semi-definite. To assess the definiteness of a matrix, user-written Excel plugins for the calculation of Eigenvalues can be used. A symmetric matrix is positive semi-definite if all Eigenvalues are larger or equal to 0. Finally, all absolute factor loading estimates and reliability estimates must be smaller or equal to 1. For our example, the solution was technically valid.

3.4. Assessment of reflective measurement and composite models

3.4.1. Evaluation of overall fit of the saturated model

Table 4 summarizes the steps to assess reflective measurement and composite models. Joint assessment should begin with the evaluation of the overall fit of a model with a saturated structural model [16,118], that is, with confirmatory factor/composite analysis. The estimated model is as specified by analysts [118]. The saturated model corresponds to a model in which all constructs are allowed to be freely correlated, whereas the concept's operationalization is exactly as specified by the analyst. The evaluation of the overall model fit of the saturated model is useful to assess the validity of the measurement and the composite models, because potential model misfit can be entirely attributed to misspecifications in the composite and/or measurement models. Therefore, empirical support can be obtained for the constructs, i.e., "Does a latent variable exist?", or "Do the indicators form

Table 2
Results of the confirmatory factor/composite analysis.

Discrepancy	Overall saturated model fit evaluation		
	Value	HI$_{95}$	Conclusion
SRMR	0.030	0.049	Supported
d$_{ULS}$	0.210	0.546	Supported
d$_G$	0.049	0.221	Supported

an emergent variable?" Table 2 contains the values of the discrepancy measures and 95% quantiles of their corresponding reference distribution for our example. The value of the SRMR was below the recommended threshold value of 0.080 [22,119]. However, the thresholds for the overall model fit in the context of PLS-PM should be considered cautiously as they are preliminary and need to be examined in more detail in future methodological research. Moreover, all discrepancy measures were below the 95% quantile of their reference distribution (HI$_{95}$). Empirical evidence was thus obtained for the latent variables (social executive behavior and social employee behavior) as well as the emergent variables (social media capability and business process performance) incorporated in the model. In case of contradictory results for the measure of fit (SRMR) and the test of overall model fit (d$_{ULS}$, and d$_G$), the test for overall model fit is preferred, as it is based on statistical inference rather than heuristic rules. Moreover, if none of the discrepancies was below the 95% quantile of the corresponding reference distribution (HI$_{95}$), analysts can evaluate whether the discrepancies are at least below the 99% quantile (HI$_{99}$) before finally rejecting the model. In the next step, each measurement and composite model must be examined separately. Authors of future studies in IS research are encouraged to report a table like Table 2.

Scholars should assess content validity for both kinds of constructs, i.e., latent variables and emergent variables, by carefully considering each type of construct and how the according concept has been operationalized in prior research. In the case of emergent variables, however, it might be desirable to modify the weighting scheme, number of indicators, and content of the indicators as illustrated in the bread and beer example. Finally, construct validity should be assessed. Depending on the concept's operationalization, this can be done in several non-exclusive ways.

3.4.2. Assessment of the reflective measurement model

For reflective measurement models in which latent variables represent behavioral concepts such as social executive behavior and social employee behavior, composite reliability, convergent validity, indicator reliability, and discriminant validity should be evaluated. Dijkstra–Henseler's ρ_A should be considered in assessing composite reliability (the correlation between latent variable and construct scores). A value of Dijkstra–Henseler's ρ_A larger than 0.707 can be regarded as reasonable, as more than 50% of the variance in the construct scores can be explained by the latent variable [120]. Table 3 shows that the values of Dijkstra–Henseler's ρ_A for social executive behavior and social employee behavior. Both are 0.938 and 0.913, and thus above the suggested threshold of 0.707, indicating reliable construct scores.

Convergent validity is the extent to which the indicators belonging to one latent variable actually measure the same construct. The average variance extracted (AVE), typically used to assess convergent validity [121], indicates how much of the indicators' variance can be explained by the latent variable. An AVE larger than 0.5 has been suggested to provide empirical evidence for convergent validity, as the corresponding latent variable explains more than half of the variance in the belonging indicators, and consequently, all other latent variables explain less than a half [96]. In our example, all AVE values are above 0.5 (0.788 and 0.716), indicating convergent validity (see Table 3).

Indicator reliability can be assessed through the factor loading estimates. As factor loading estimates are standardized in PLS-PM, the

Table 3
Measurement model evaluation.

Code	Construct/indicator	ρ_A	AVE	VIF	Weight	Loading
Social executive behavior (1: Strongly disagree, 5: Strongly agree) (reflective measurement model, Mode A consistent (PLSc), SEXB2 as dominant indicator)		0.938	0.788			
SEXB1	Behavior of top business executives towards adoption of social media is positive				0.278***	0.905***
SEXB2	Top business executives are positive in adopting social media for business activities				0.269***	0.877***
SEXB3	Top business executives support adoption of social media for business activities				0.263***	0.856***
SEXB4	Top business executives are willing to support adoption of social media in the firm				0.280***	0.912***
Social employee behavior (1: Strongly disagree, 5: Strongly agree) (reflective measurement model, Mode A consistent, SEMB2 as dominant indicator)		0.913	0.716			
SEMB1	Employee behavior towards adoption of social media is positive				0.301***	0.901***
SEMB2	Employees are positive to adopt social media in the firm				0.274***	0.820***
SEMB3	Employees support adoption of social media in the firm				0.257***	0.769***
SEMB4	Employees are willing to support adoption of social media in the firm				0.296***	0.888***
Social media capability: My firm has deliberately used and leveraged… for business activities (1: Strongly disagree, 5: Strongly agree) (composite model, Mode B, SMC1 as dominant indicator)						
SMC1	Facebook			1.037	0.229***	0.397***
SMC2	Twitter			1.032	0.489***	0.627***
SMC3	Corporate blog(s)			1.059	0.601***	0.751***
SMC4	LinkedIn			1.020	0.333***	0.455***
Business process performance: Relative to your key competitors, what is your performance in last three years in the following business processes (1: Significantly worse, 5: Significantly better than my key competitors) (composite model, Mode B, BPP4 as dominant indicator)						
BPP1	Supplier relations			1.022	0.285**	0.397**
BPP2	Product and service enhancement			1.134	0.553***	0.307**
BPP3	Production and operations			1.105	0.108	0.203†
BPP4	Marketing and sales			1.064	0.609***	0.531***
BPP5	Customer relations			1.063	0.629***	0.591***

Note: †p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001, one-tailed test.

squared factor loading estimate equals the estimated indicator reliability. It is generally advisable for factor loadings to be greater than 0.707, indicating that more than 50% of the variance in a single indicator can be explained by the corresponding latent variable. In this context, also the significance of the factor loading estimates should be investigated. Somewhat lower values are not really problematic as long as the construct validity and reliability criteria are met. Table 3 presents the factor loading estimates from our example. They range from 0.769 to 0.912 and are all significant on a 1‰ level, suggesting that the measures are reliable.

Discriminant validity entails that two latent variables that are meant to represent two different theoretically concepts are statistically sufficiently different. To obtain empirical evidence for discriminant validity, IS scholars should consider the HTMT [35]. The HTMT should be lower than 0.85 (more strict threshold) or 0.90 (more lenient threshold) or significantly smaller than 1 ([36,37]). In our example, the HTMT of social executive behavior to social employee behavior is 0.322, and thus below the recommended threshold of 0.85 (and of 0.90). Moreover, the one-sided 95% percentile confidence interval of HTMT does not cover 1, that is, it is significantly different from 1. Scholars can also follow [37] suggestion to test whether the HTMT is significantly smaller than 0.85 or 0.90.

3.4.3. Assessment of the composite model

The composite model requires an evaluation sui generis – an examination of the composite model with respect to multicollinearity, weights, composite loadings, and their significances [43,76,77,83,122].

As composite models are typically estimated by Mode B (regression weights) in PLS-PM, collinearity among indicators forming an emergent variable should be investigated by means of the variance inflation factor (VIF), as high multicollinearity can lead to insignificant estimates and unexpected signs of the weights. Traditionally, VIF values above 5 are regarded as indications of problematic multicollinearity [38,46]. Yet, typical phenomena of multicollinearity can also occur in case of VIF values far below 5. For weights estimated by Mode A, an assessment of multicollinearity is not necessary as these equal scaled covariances, and therefore, ignores multicollinearity [32].

While weights show the relative contribution of an indicator to its construct, composite loadings represent the correlation between the indicator and the corresponding emergent variable; a loading shows the absolute contribution of an indicator to its construct [122]. As weights show the degree of importance of each indicator (ingredient) to the construct, analysts should examine whether all indicator weight estimates are significant. For indicators with non-significant weight estimates, one must investigate whether composite loading estimates are statistically significant and consider dropping any indicators with non-significant weight and loading estimates. However, content validity must be considered as well, because dropping an indicator may alter the meaning of the emergent variable. IS scholars can thus decide to keep an indicator with non-significant weight and loading to preserve the construct's content validity [46].

Table 3 shows that the VIF values for the indicators of the composite models range from 1.020 to 1.134, suggesting that multicollinearity is not a problem in our data. Moreover, all weight and composite loading estimates show the expected sign and are significant at a 5% significance level except one (estimated weight of the indicator *production and operations* of the construct business process performance). The weight estimate of this indicator is 0.108, and its composite loading estimate is 0.203† (close to be significant). Considering content validity, the indicator production and operations may include some of the firm's key business processes. Therefore, we decided to keep the indicator in the empirical analysis to preserve content validity and avoid altering the meaning of the emergent variable business process performance. In this type of situation, analysts can also repeat the analysis, dropping the questionable indicators to explore whether the decision to keep or drop these indicators affects the results. We dropped BPP3 and repeated the empirical analysis. The results obtained were qualitatively identical, suggesting that this decision does not affect the research findings. One might ask why BPP3 should be included when the results do not change. This is theoretically justified, as it is difficult to imagine a company's business process performance without production and operations processes, as these often are the heart of a company. Because all reflective measurement and composite models from our example show desirable properties, we proceed to evaluate the structural model.

Table 4
Steps to assess common factor and composite models.

Steps	Type of construct	Description	Assessment criterion	Decision criterion	Interpretation
Testing the adequacy of reflective measurement and composite models	Latent and emergent variable	Evaluate the overall fit of the model with a saturated structural model by investigating discrepancy between empirical and model-implied indicator variance–covariance matrix	SRMR	$SRMR < 0.080$ $SRMR < HI_{95}$	A SRMR value smaller than 0.080 indicates an acceptable model fit [22]; however, these thresholds are preliminary and need to be investigated in more detail
			d_{ULS} d_G	$d_{ULS} < HI_{95}$ $d_G < HI_{95}$	The null hypothesis that the population indicator variance–covariance matrix equals the model-implied counterpart is not rejected. Hence, empirical evidence for the model is given when the value of the discrepancy measure is below the 95% quantile of its corresponding reference distribution
Evaluating content validity	Latent and emergent variable	How the corresponding theoretical concepts have been operationalized (measured or built) in prior research			Flexibility in the case of artifacts represented by an emergent variable (bread and beer analogy)
Evaluating reliability of construct scores	Latent variable	Evaluating whether the construct scores reliably represent the underlying construct	ρ_A	$\rho_A > 0.707$	More than 50% of the variance in the construct scores can be explained by the underlying latent variable
Evaluating indicator reliability	Latent variable	Evaluating whether indicators are reliable	Factor loading estimates; Factor loading significance	Factor loading estimates > 0.707; Significant at 5% significance level	More than 50% of the indicator's variance is explained by the latent variable
Evaluating convergent validity	Latent variable	Evaluating the share of variance in the indicators that is explained by the underlying latent variable	AVE	$AVE > 0.5$	More than 50% of indicators' variance is explained by the underlying latent variable
Evaluating discriminant validity	Latent variable	Evaluating whether two latent variables are statistically different	HTMT	$HTMT < 0.85$ (or whether the HTMT is significantly smaller than 1)	Factors are statistically different and thus have discriminant validity
Multicollinearity	Emergent variable (estimated by Mode B)	Evaluating how the standard errors of the weight estimates are affected by the correlations of the indicators	VIF	$VIF < 5$	If the estimates suffer from multicollinearity, weights obtained by Mode A or predetermined weights can be used
Weights	Emergent variable	Evaluating relative contribution of an indicator to its construct	Weights' value and significance	Significant at 5% significance level	Each indicator contributes significantly to the emergent variable
Loadings	Emergent variable	Evaluating absolute contribution of an indicator to its construct	Loading significance	Significant at 5% significance level	Each indicator contributes to the emergent variable in a statistically significant way

3.5. Assessment of the structural model

In evaluating the structural model, the analyst should examine the overall fit of the estimated model, the path coefficient estimates, their significance, the effect sizes (f^2), and the coefficient of determination (R^2, [3,123]). Analysts should focus specifically and primarily on overall model fit in confirmatory research and primarily on R^2, the path coefficient estimates, and the effect sizes in explanatory research [24]. Table 7 summarizes the steps to follow in evaluating the structural model.

3.5.1. Evaluation of the overall fit of the estimated model

First, analysts should evaluate the overall fit of the estimated model through the bootstrap-based test of overall model fit and the SRMR as a measure of approximate fit to obtain empirical evidence for the proposed theory. Analysis in confirmatory research without assessing the overall model would be incomplete as this means ignoring empirical evidence for and also against the proposed model and the postulated theory [124]. Without assessing the model fit, a researcher would not obtain any signal if he or she had incorrectly omitted an important effect in the model. Because the test for overall model fit was introduced only recently in the context of PLS-PM, the vast majority of models estimated by PLS-PM in past IS research has not been evaluated in this respect. However, because the overall model fit can now be tested in the context of PLS-PM, we encourage IS scholars to take this evaluation very seriously in causal research. In our example, all values of discrepancy measures were below the 95% quantile of their corresponding reference distribution (HI_{95}), indicating that the estimated model was not rejected at a 5% significance level (see Table 5). Moreover, the SRMR was below the preliminary suggested threshold of 0.080, indicating an acceptable model fit. This result suggests that the proposed model is well suited for confirming and explaining the development of social media capability and business process performance among firms. While the model fit suggests that there is a possibility that the world functions according to the specified model, the model can still be misspecified in the sense of over-parameterization, i.e., the model contains superfluous zero-paths [22]. Neither the bootstrap-based test of model fit nor the SRMR punishes for unnecessary paths, i.e., neither of them rewards parsimony. Regardless of whether one conducts confirmatory or explanatory research, it remains indispensable to assess all path coefficients and their significance. Table 6 presents the construct correlation matrix.

Table 6
Construct correlation matrix.

	1	2	3	4	5	6
1. Social executive behavior	1.000					
2. Social employee behavior	0.322	1.000				
3. Social media capability	0.550	0.532	1.000			
4. Business process performance	0.216	0.309	0.515	1.000		
5. Firm size	−0.025	−0.048	−0.014	0.016	1.000	
6. Industry	0.069	0.073	0.010	0.036	0.038	1.000

3.5.2. Evaluation of path coefficients and their significance levels

The path coefficient estimates are essentially standardized regression coefficients, whose sign and absolute size can be assessed. These coefficients are interpreted as the change in the dependent construct measured by standard deviations, if an independent construct is increased by one standard deviation while keeping all other explanatory constructs constant (ceteris paribus consideration). For example, increasing social media capability by one standard deviation will increase business process performance by 0.515 standard deviations if all other variables are kept constant. Statistical tests and confidence intervals can be used to draw conclusions about the population parameters. For confidence intervals, the percentile bootstrap confidence interval is recommended [125]. As shown in Fig. 4, the path coefficient estimates for the hypothesized relationships included in the example range from 0.396 to 0.515, and are all significant at a 5% significance level except the effect of the two control variables, firm size and industry. A path coefficient estimate is considered as statistically significant different from zero at a 5% significance level when its p-value is below 0.05 or when the 95% bootstrap percentile confidence interval constructed around the estimate does not cover the zero.

3.5.3. Evaluation of effect sizes

The practical relevance of significant effects should be investigated by considering the effect sizes of the relationships between the constructs. The effect size is a measure of the magnitude of an effect that is independent of sample size. The f^2 values ranging from 0.020 to 0.150, 0.150 to 0.350, or larger or equal to 0.350, indicating weak, medium, or large effect size respectively [108]. Just as all actors in a movie cannot play a leading role, it is unusual and unlikely that most constructs will have a large effect size in the model. We provide this clarification because scholars often expect/self-demand that all/most of their effect

Table 5
Structural model evaluation.

Relationship	Path coefficient	
Social executive behavior → Social media capability (H1)	0.422*** (8.830) [0.327, 0.512]	
Social employee behavior → Social media capability (H2)	0.396*** (8.052) [0.300, 0.490]	
Social media capability → Business process performance (H3)	0.515*** (10.232) [0.426, 0.609]	
Firm size → Business process performance (control variable)	0.022 (0.305) [-0.128, 0.160]	
Industry → Business process performance (control variable)	0.030 (0.312) [-0.161, 0.174]	
Endogenous variable	R^2	
Social media capability	0.443	
Business process performance	0.267	
Overall fit of the estimated model	Value	HI_{95}
SRMR	0.032	0.049
d_{ULS}	0.232	0.558
d_G	0.052	0.222
Effect size	f^2	
Social executive behavior → Social media capability (H1)	0.286	
Social employee behavior → Social media capability (H2)	0.252	
Social media capability → Business process performance (H3)	0.362	
Firm size → Business process performance (control variable)	0.001	
Industry → Business process performance (control variable)	0.001	

Note: t-values (one-tailed test) are presented in parentheses. Percentile bootstrap confidence intervals are presented in brackets.

Table 7
Steps to follow in performing structural model evaluation.

Steps	Description	Criterion	Suggested threshold	Interpretation
Overall fit of estimated model	Evaluating overall fit of the estimated model by evaluating discrepancy between the empirical indicator variance–covariance matrix and its model-implied counterpart	SRMR, d_{ULS}, d_G	SRMR < 0.080 SRMR $< HI_{95}$; $d_{ULS} < HI_{95}$; $d_G < HI_{95}$	Value of discrepancy measure below the 95% quantile of the corresponding reference distribution provides empirical evidence for the postulated model. In other words, it is possible that the empirical data stem from a world that functions as theorized by the model
Consider path coefficient estimates and their significance levels	Standardized regression coefficients are interpreted as change in standard deviations of the dependent variable if an independent variable is increased by one standard deviation while all other independent variables in the equation remain constant	Path coefficient estimates and their significance level	Significant at 5% significance level, i.e., p-value $< 5\%$	Effect of independent variables on dependent variables is statistically significant
Consider effect sizes (f^2)	Measure of the magnitude of an effect that is independent of sample size. Give an indication about the practical relevance of an effect	f^2 value	$f^2 < 0.020$: no substantial effect $0.020 \leq f^2 < 0.150$: weak effect size $0.150 \leq f^2 < 0.350$: medium effect size $f^2 \geq 0.350$: large effect size	Degree of strength of an effect
Evaluate R^2	Explained variance of an dependent construct	R^2	When the phenomena are already quite well understood, one would expect a high R^2. When the phenomena are not yet well understood, a lower R^2 is acceptable	Degree of variance explained for phenomenon under investigation

magnitude be large – an unrealistic expectation. This cautionary note extends to supervisors' expectations for their Ph.D. students (as illustrated by [126]). In our sample, the f^2 values for the hypothesized relationships range from 0.252 to 0.362 (medium to large).

3.5.4. Evaluation of R^2

R^2 is used to assess goodness of fit in regression analysis [87]. In the case of models estimated by OLS, the R^2 value gives the share of variance explained in a dependent construct. Thus, it provides insights into a model's in-sample predictive power [127]. Moreover, R^2 forms the basis for several innovative model selection criteria ([37,100]). Reporting R^2 makes PLS-PM research future-proof in this regard, because the new model selection criteria can still be calculated ex post as long as the R^2 values are given.

The expected magnitude of R^2 depends on the phenomenon investigated. As some phenomena are already quite well understood, one would expect a relatively high R^2. For phenomena that are less well understood, a lower R^2 is acceptable. The R^2 values should be judged relative to studies that investigate the same dependent variable. In our example, the R^2 values for social media capability and business process performance are 0.443 and 0.267, respectively. The study of social media in organizations is in its initial stages [77]. Braojos et al. [128] report an R^2 value of 0.541 for social media capability. In our example, social executive behavior and social employee behavior explain 44.3% of variance in development of social media capability, using two unexplored exogenous variables for social media capability (social executive behavior and social employee behavior). Considering explained variance in prior IS research and the originality of our two exogenous variables in influencing social media capability, an R^2 of 0.443 seems to be an excellent value.

The models of [129,130]) explain 49% and 43.9% of the variance in business process outcomes. In our example, social media capability, firm size, and industry explain 26.7% of the variance in business process performance. Although this R^2 value is somewhat smaller than those obtained by [129,130]), it can be considered as satisfactory because our model is the first using social media capability to explain business process performance individually. The independent variables explaining business process outcomes in [129,130]) work refer to other IT resources (e.g., IT assets, enterprise resource planning capabilities) different from social media capability. This subsection illustrates by our fictive example how analysts can report and compare their R^2 values.

4. Discussion and conclusions

IS research often tackles complex research problems and questions that require conceptualization and operationalization of different types of theoretical concepts, i.e., behavioral concepts and artifacts, as well as the estimation of their relationships. PLS-PM is a suitable estimator for this purpose. How can one perform and report an impactful analysis using PLS-PM in IS research following the recent improvements in PLS-PM? This study provides thorough guidelines on PLS-PM in the framework of causal (confirmatory and explanatory) research, employing the latest standards recommended. In doing so, it addresses the why and how to perform and report a PLS-PM estimation in confirmatory and explanatory IS research, illustrated by a fictive example on business value of social media. This is the key contribution of this paper to the methodological literature in IS empirical research.

In the last five years, methodologists have overcome major weaknesses of traditional PLS-PM, such as its inconsistency for latent variable models and lack of a test for overall model fit. To benefit from all these enhancements, IS scholars need new guidelines for empirical studies that incorporate all these recent new developments and insights, as most of the guidelines papers on PLS-PM in the IS research were published before 2013 (e.g., [12,43–45]). Although several recent scholarly textbooks and articles (e.g. [46–48],) have provided guidelines for causal research that cover some of the latest enhancements to
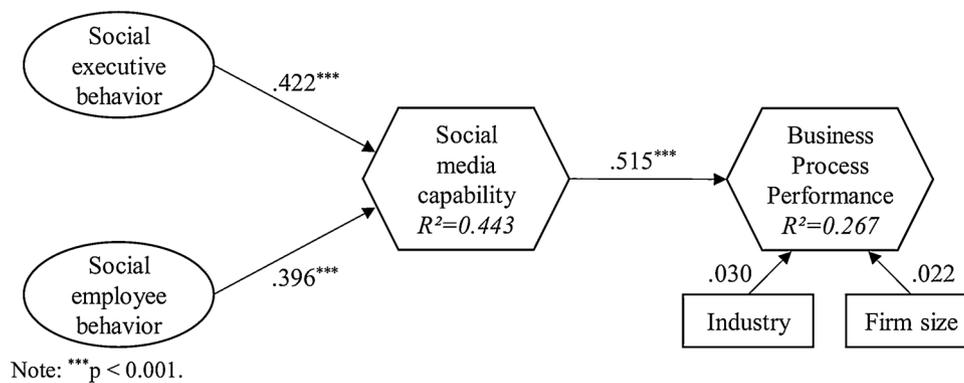
Fig. 4. Results of the structural model.
Note: ***p < 0.001.

PLS-PM, neither of these PLS-PM guidelines for causal research covered the full range of recent developments, nor did they introduce any new framework for applying PLS-PM and reporting its outcomes. To address this shortcoming in the existing IS literature, this paper provides updated guidelines on the use of PLS-PM in assessment of reflective measurement models, composite models, and structural models. To the best of our knowledge, the proposed guidelines take into account all recent enhancements. An application of the guidelines is illustrated using a parsimonious IS research example on business value of social media.

In contrast to prior guidelines [11,12], our article introduces the artifact – a human-made/firm-made object – as a new kind of theoretical concept and shows how this type of theoretical concept can be operationalized by means of the composite model. Because a significant proportion of theoretical concepts in IS research are human-made/firm-made, one can expect the composite model to become the dominant conceptualization in IS research in the coming years. Against this background, we highlight the usefulness of model testing in confirmatory and explanatory research using PLS-PM. Without considering its results, it is hardly possible to obtain empirical evidence for or against a scholar's proposed theory. Finally, we strongly recommend that scholars employ consistent estimators, using PLSc when the theoretical concept is operationalized by a measurement model.

As our article about the use of PLS-PM for causal research is limited to linear, recursive models containing only first-order constructs, future IS research should develop additional updated guidelines incorporating recent developments for more complex models, such as models containing moderation effects, second-order emergent variables of emergent variables, and for composite models that account for more complex relationships between the indicators and the emergent variable. Although some steps have been made using PLS-PM to deal with endogeneity in the form of omitted variables (e.g., [27,55]), the problem of endogeneity requires more attention in the field of IS and in the context of PLS-PM.

Although this study focuses only on PLS-PM for confirmatory and explanatory purposes, PLS-PM can be used for different types of research [24]. A further promising application of PLS-PM in IS research is predictive modeling, which aims to produce accurate forecasts [23]. These assessment criteria differ from those employed in confirmatory and explanatory research [131,132], and models of causal research do not necessarily perform well when it comes to prediction purposes and vice versa.

As part of the progress in methodology-related research, scholars continuously suggest new approaches and new validity criteria. For instance, [133] recommend conducting redundancy analysis to assess the convergent validity in the context of causal–formative measurement and composite models. As soon as there is sufficient evidence (e.g., by means of Monte Carlo simulations) for the efficacy of the new suggestion, scholars may consider adding this suggestion to their methodological toolbox.

The adjusted coefficient of determination (adjusted R^2) has been proposed by prior methodological research [3,46,123] as a criterion to assess the structural model in explanatory research [24]. Using a Monte Carlo study [100], find that R^2, adjusted R^2, goodness-of-fit index, and Q^2 are not appropriate criteria to compare competing latent variable models based on the same dataset (same sample). Their argument is that these criteria improve with greater model complexity, and therefore, favor more complex models (e.g., the saturated model). To avoid the shortcoming, they propose to employ the Geweke–Meese criterion (GM) and Bayesian information criteria (BIC) to compare alternative latent variable models, and find support for their proposal in a simulation study. "BIC and GM should be used due to their high model selection accuracy and ease of use" ([100], p. 7). Future methodological research should study the performance of adjusted R^2 in comparing different estimated models based on different samples as well as extending the simulation study of [100] to emergent variable models.

PLS-PM has become a valuable statistical tool for empirical research in IS and many other disciplines of business and social sciences research. To meet the ever-growing scholarly demand in terms of scientific rigor, methodologists have continuously been improving PLS-PM. This paper helps disseminate these improvements, and enables users of PLS-PM to be aware of and fulfill the contemporary methodological standards.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.im.2019.05.003.

References

[1] K. Bollen, A new incremental fit index for general structural equation models, Sociol. Methods Res. 17 (3) (1989) 303–316.
[2] H. Wold, Path models with latent variables: the NIPALS approach, in: H. Blalock, A. Aganbegian, F. Borodkin, R. Boudon, V. Capecchi (Eds.), Quantitative Sociology: International Perspectives on Mathematical and Statistical Modeling,

ARTICLE IN PRESS

J. Benitez, et al.   Information & Management xxx (xxxx) xxx–xxx

Academic, New York, USA, 1975, pp. 307–357.

[3] J. Henseler, G. Hubona, P. Ray, Using PLS path modeling in new technology research: updated guidelines, Ind. Manage. Data Syst. 116 (1) (2016) 1–20.

[4] J. Hair, M. Sarstedt, T. Pieper, C. Ringle, The use of partial least squares structural equation modeling in Strategic Management research: a review of past practices and recommendations for future applications, Long Range Plann. 45 (5/6) (2012) 320–340.

[5] J. Hair, M. Sarstedt, C. Ringle, J. Mena, An assessment of the use of partial least squares structural equation modeling in Marketing research, J. Acad. Mark. Sci. 40 (3) (2012) 414–433.

[6] D. Peng, F. Lai, Using partial least squares in Operations Management research: a practical guideline and summary of past research, J. Oper. Manage. 30 (6) (2012) 467–480.

[7] C. Ringle, M. Sarstedt, R. Mitchell, S. Gudergan, Partial least squares structural equation modeling in HRM research, Int. J. Hum. Resour. Manage. (2018) 1–27.

[8] N. Avkiran, C. Ringle, Partial Least Squares Structural Equation Modeling, Springer, Cham, Switzerland, 2018.

[9] T. Müller, F. Schuberth, J. Henseler, PLS path modeling: a confirmatory approach to study tourism technology and tourist behavior, J. Hosp. Tour. Technol. 9 (3) (2018) 249–266.

[10] M. Sarstedt, C. Ringle, D. Smith, R. Reams, J. Hair, Partial least squares structural equation modeling (PLS-SEM): a useful tool for family business researchers, J. Fam. Bus. Strategy 5 (1) (2014) 105–115.

[11] G. Marcoulides, C. Saunders, PLS: a silver bullet? Mis Q. 30 (2) (2006) iii–ix.

[12] C. Ringle, M. Sarstedt, D. Straub, A critical look at the use of PLS-SEM in MIS Quarterly, Mis Q. 36 (1) (2012) iii–xiv.

[13] G. Polites, N. Roberts, J. Thatcher, Conceptualizing models using multidimensional constructs: a review and guidelines for their use, Eur. J. Inf. Syst. 21 (1) (2012) 22–48.

[14] P. Lowry, J. Zhang, C. Wang, M. Siponen, Why do adults engage in cyberbullying on social media? An integration of online disinhibition and deindividuation effects with the social structure and social learning model, Inf. Syst. Res. 27 (4) (2016) 962–986.

[15] L. Rueda, J. Benitez, J. Braojos, From traditional education technologies to student satisfaction in Management education: a theory of the role of social media applications, Inf. Manage. 54 (8) (2017) 1059–1071.

[16] J. Henseler, Bridging Design and Behavioral research with variance-based structural equation modeling, J. Advert. 46 (1) (2017) 178–192.

[17] H. Reichenbach, The Direction of Time, University of California at Los Angeles, Los Angeles, USA, 1956.

[18] A. Diamantopoulos, Incorporating formatives measures into covariance-based structural equation models, Mis Q. 35 (2) (2011) 335–358.

[19] H. Simon, The Sciences of the Artificial, MIT Press, Cambridge, USA, 1969.

[20] D. Cole, S. Maxwell, R. Arvey, E. Salas, Multivariate group comparisons of variable systems: MANOVA and structural equation modeling, Psychol. Bull. 114 (1) (1993) 174–184.

[21] G. Hancock, F. Lawrence, J. Nevitt, Type I error and power of latent mean methods and MANOVA in factorially invariant and noninvariant latent variable systems, Struct. Equ. Model. 7 (4) (2000) 534–556.

[22] J. Henseler, T. Dijkstra, M. Sarstedt, C. Ringle, A. Diamantopoulos, D. Straub, D. Ketchen, F. Hair, G. Hult, R. Calatone, Common beliefs and reality about PLS: comments on Rönkkö and Evermann (2013), Organ. Res. Methods 17 (2) (2014) 182–209.

[23] G. Shmueli, S. Ray, J. Estrada, S. Chatla, The elephant in the room: predictive performance on PLS models, J. Bus. Res. 69 (10) (2016) 4552–4564.

[24] J. Henseler, Partial least squares path modeling: quo vadis? Qual. Quant. 52 (1) (2018) 1–8.

[25] D. Goodhue, W. Lewis, R. Thompson, Does PLS have advantages for small sample size or non-normal data? Mis Q. 36 (3) (2012) 981–1001.

[26] E. Rigdon, Choosing PLS path modeling as analytical method in European Management research: a realist perspective, Eur. Manage. J. 34 (6) (2016) 598–605.

[27] J. Benitez, J. Henseler, J. Roldan, How to address endogeneity in partial least squares path modeling, Proceedings of the 22th Americas Conference on Information Systems (2016) 1–10.

[28] T. Dijkstra, A perfect match between a model and a mode, in: H. Latan, R. Noonan (Eds.), Partial Least Squares Path Modeling: Basic Concepts, Methodological Issues and Applications, Springer, Cham, Switzerland, 2017, pp. 55–80.

[29] F. Schuberth, J. Henseler, T. Dijkstra, Partial least squares path modeling using ordinal categorical indicators, Qual. Quant. 52 (1) (2018) 9–35.

[30] D. Goodhue, W. Lewis, R. Thompson, A multicollinearity and measurement error statistical blind spot: correcting for excessive false positives in regression and PLS, Mis Q. 41 (3) (2017) 667–684.

[31] T. Dijkstra, J. Henseler, Linear indices in nonlinear structural equation models: best fitting proper indices and other composites, Qual. Quant. 45 (6) (2011) 1505–1518.

[32] E. Rigdon, Rethinking partial least squares path modeling: in praise of simple methods, Long Range Plann. 45 (5/6) (2012) 341–358.

[33] L. Guttman, A basis for analyzing test-retest reliability, Psychometrika 10 (4) (1945) 255–282.

[34] K. Sijtsma, On the use, the misuse, and the very limited usefulness of Cronbach's alpha, Psychometrika 74 (1) (2009) 107–120.

[35] J. Henseler, C. Ringle, M. Sarstedt, A new criterion for assessing discriminant validity in variance-based structural equation modeling, J. Acad. Mark. Sci. 43 (1) (2015) 115–135.

[36] C. Voorhees, M. Brady, R. Calantone, E. Ramirez, Discriminant validity testing in

marketing: an analysis, causes for concern, and proposed remedies, J. Acad. Mark. Sci. 44 (1) (2016) 119–134.

[37] G. Franke, M. Sarstedt, Heuristics versus statistics in discriminant validity testing: a comparison of four procedures, Internet Res. (in press), pp. 1-32.

[38] J. Hair, C. Ringle, M. Sarstedt, PLS-SEM: indeed a silver bullet, J. Mark. Theory Pract. 19 (2) (2011) 139–151.

[39] M. Rönkkö, J. Evermann, A critical examination of common beliefs about partial least squares path modeling, Organ. Res. Methods 16 (3) (2013) 425–448.

[40] T. Dijkstra, K. Schermelleh-Engel, Consistent partial least squares for nonlinear structural equation models, Psychometrika 79 (4) (2014) 585–604.

[41] T. Dijkstra, J. Henseler, Consistent partial least squares path modeling, Mis Q. 39 (2) (2015) 297–316.

[42] T. Dijkstra, J. Henseler, Consistent and asymptotically normal PLS estimators for linear structural equations, Comput. Stat. Data Anal. 81 (1) (2015) 10–23.

[43] W. Chin, Issues and opinion on structural equation modeling, Mis Q. 22 (1) (1998) vii–xvi.

[44] D. Gefen, D. Straub, A practical guide to factorial validity using PLS-Graph: tutorial and annotated example, Commun. Assoc. Inf. Syst. 16 (1) (2005) 91–109.

[45] G. Marcoulides, W. Chin, C. Saunders, A critical look at partial least squares modeling, Mis Q. 33 (1) (2009) 171–175.

[46] J. Hair, G. Hult, C. Ringle, M. Sarstedt, A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), 2nd ed., Sage, Thousand Oaks, USA, 2017.

[47] j. hair, c. hollingsworth, a. randolph, a. chong, an updated and expanded assessment of pls-sem in information Systems research, Ind. Manage. Data Syst. 117 (3) (2017) 442–458.

[48] M. Sarstedt, J. Hair, C. Ringle, Partial least squares, in: C. Homburg, M. Klarmann, A. Vomberg (Eds.), Handbook of Market Research, Springer, Berlin, Germany, 2017.

[49] A. Van Riel, J. Henseler, I. Kemeny, Z. Sasovova, Estimating hierarchical constructs using consistent partial least squares: the case of second-order composites of common factors, Ind. Manage. Data Syst. 117 (3) (2017) 459–477.

[50] J. Becker, I. Ismail, Accounting for sampling weights in PLS path modeling: simulations and empirical examples, Eur. Manage. J. 34 (6) (2016) 606–617.

[51] M. Rademaker, F. Schuberth, T. Dijkstra, Measurement error correlation within blocks of indicators in consistent partial least squares: Issues and remedies," Internet Res. (in press), pp. 1-40.

[52] S. Jung, J. Park, Consistent partial least squares path modeling via regularization, Front. Psychol. 9 (1) (2018) 1–10.

[53] M. Sarstedt, J. Henseler, C. Ringle, Multigroup analysis in partial least squares (PLS) path modeling: alternative methods and empirical results, Measurement and Research Methods in International Marketing, Emerald, Enschede, The Netherlands, 2011, pp. 195–218.

[54] M. Klesel, F. Schuberth, J. Henseler, B. Niehaves A test for multigroup comparison in partial least squares path modeling," Internet Res. (in press), pp. 1-40.

[55] T. Hult, J. Hair, D. Proksch, M. Sarstedt, A. Pinkwart, C. Ringle, Addressing endogeneity in International Marketing applications of partial least squares structural equation modeling, J. Int. Mark. 26 (3) (2018) 1–21.

[56] C. Ringle, M. Sarstedt, Gain more insight from your PLS-SEM results: the importance-performance map analysis, Ind. Manage. Data Syst. 116 (9) (2016) 1865–1886.

[57] G. Khan, M. Sarstedt, W. Shiau, J. Hair, C. Ringle, M, M. Fritze, Methodological research on partial least squares structural equation modeling (PLS-SEM): An analysis based on social network approaches,#xp#, Internet Res. (2019) 1–40.

[58] E. Rigdon, M. Sarstedt, C. Ringle, On comparing results from CB-SEM and PLS-SEM: five perspectives and five recommendations, Mark. Zfp 39 (3) (2017) 4–16.

[59] K. Bollen, Evaluating effect, composite, and causal indicators in structural equation models," Mis Q. (35:2) (2011) 359–372.

[60] R. Bagozzi, Y. Yi, Specification, evaluation, and interpretation of structural equation modeling, J. Acad. Mark. Sci. 40 (1) (2012) 8–34.

[61] M. Sarstedt, J. Hair, C. Ringle, K. Thiele, S. Gudergan, Estimation issues with PLS and CBSEM: where the bias lies!, J. Bus. Res. 69 (10) (2016) 3998–4010.

[62] D. Borsboom, G. Mellenbergh, J. Van Heerden, The theoretical status of latent variables, Psychol. Rev. 110 (2) (2003) 203–219.

[63] J. Henseler, Is the Whole More Than the Sum of Its Parts? On the Interplay of Marketing and Design Research, University of Twente, Enschede, The Netherlands, 2015.

[64] K. Bollen, R. Hoyle, Latent variables in structural equation modeling, in: R. Hoyle (Ed.), Handbook of Structural Equation Modeling, Guilford, New York, USA, 2012, pp. 56–67.

[65] C. Spearman, General intelligence, objectively determined and measured, Am. J. Psychol. 15 (2) (1904) 201–292.

[66] K. Jöreskog, A general approach to confirmatory maximum likelihood factor analysis, Psychometrika 34 (2) (1969) 183–202.

[67] P. Lazarsfeld, Latent structure analysis, Psychol.: Study Sci. 3 (1) (1959) 476–543.

[68] A. Diamantopoulos, Formative indicators: introduction to the special issue, J. Bus. Res. 61 (12) (2008) 1201–1202.

[69] K. Bollen, A. Diamantopoulos, In defense of causal-formative indicators: a minority report, Psychol. Methods 22 (3) (2017) 581–596.

[70] M. Aguirre-Urreta, M. Rönkkö, G. Marakas, Omission of causal indicators: consequences and implications for measurement, Meas. Interdiscip. Res. Perspect. 14 (3) (2016) 75–97.

[71] K. Bollen, R. Lennox, Conventional wisdom on measurement: a structural equation perspective, Psychol. Bull. 110 (2) (1991) 305–314.

[72] M. Rönkkö, C. McIntosh, J. Antonakis, On the adoption of partial least squares in Psychological research: caveat emptor, Pers. Individ. Dif. 87 (1) (2015) 76–84.

[73] F. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology, Mis Q. 13 (3) (1989) 319–340.

[74] V. Venkatesh, S. Brown, L. Maruping, H. Bala, Predicting different conceptualizations of system use: the competing roles of behavioral intention, facilitating conditions, and behavioral expectation, Mis Q. 32 (3) (2008) 483–502.

[75] H. Barki, R. Titah, C. Boffo, Information system use-related activity: an expanded behavioral conceptualization of individual-level information system use, Inf. Syst. Res. 18 (2) (2007) 173–192.

[76] J. Benitez, J. Llorens, J. Braojos, How information technology influences opportunity exploration and exploitation firm's capabilities, Inf. Manage. 55 (4) (2018) 508–523.

[77] J. Benitez, A. Castillo, J. Llorens, J. Braojos, IT-enabled knowledge ambidexterity and innovation performance: the moderator role of social media capability, Inf. Manage. 55 (1) (2018) 131–143.

[78] J. Benitez, Y. Chen, T. Teo, A. Ajamieh, Evolution of the impact of e-business technology on operational competence and firm profitability: a panel data investigation, Inf. Manage. 55 (1) (2018) 120–130.

[79] J. Benitez, R. Walczuch, Information technology, the organizational capability of proactive corporate environmental strategy and firm performance: a resource-based analysis, Eur. J. Inf. Syst. 21 (6) (2012) 664–679.

[80] Y. Chen, Y. Wang, S. Nevo, J. Benitez, G. Kou, IT capabilities and product innovation performance: the roles of corporate entrepreneurship and competitive intensity, Inf. Manage. 52 (6) (2015) 643–657.

[81] A. Ajamieh, J. Benitez, J. Braojos, C. Gelhard, IT infrastructure and competitive aggressiveness in explaining and predicting performance, J. Bus. Res. 69 (10) (2016) 1–31.

[82] Y. Chen, Y. Wang, S. Nevo, J. Benitez, G. Kou, Improving strategic flexibility with information technologies: insights for firm performance in an emerging economy, J. Inf. Technol. 32 (1) (2017) 10–25.

[83] J. Benitez, G. Ray, J. Henseler, Impact of information technology infrastructure flexibility on mergers and acquisitions, Mis Q. 42 (1) (2018) 25–43.

[84] N. Melville, K. Kraemer, V. Gurbaxani, Information technology and organizational performance: an integrative model of IT business value, Mis Q. 28 (2) (2004) 283–322.

[85] A. Bharadwaj, A resource-based perspective on information technology and firm performance: an empirical investigation, Mis Q. 24 (1) (2000) 169–196.

[86] R. Santhanam, E. Hartono, Issues in linking information technology capability to firm performance, Mis Q. 27 (1) (2003) 125–153.

[87] J. Wooldridge, Introductory Econometrics: a Modern Approach, 5th ed., Cengage Learning, Mason, USA, 2013.

[88] J. Lohmöller, Predictive vs. Structural modeling: PLS vs. ML, Latent Variable Path Modeling With Partial Least Squares, Physica-Verlag, Hidelberg, Germany, 1989, pp. 199–226.

[89] J. Hair, G. Hult, C. Ringle, M. Sarstedt, K. Thiele, Mirror, mirror on the wall: a comparative evaluation of composite-based structural equation modeling methods, J. Acad. Mark. Sci. 45 (5) (2017) 616–632.

[90] H. Hwang, Y. Takane, Generalized structure component analysis, Psychometrika 69 (1) (2004) 81–99.

[91] W. Reinartz, M. Haenlein, J. Henseler, An empirical comparison of the efficacy of covariance-based and variance-based SEM, Int. J. Res. Mark. 26 (4) (2009) 332–344.

[92] J. Henseler, Why generalized structured component analysis is not universally preferable to structural equation modeling, J. Acad. Mark. Sci. 40 (3) (2012) 402–413.

[93] Y. Takane, H. Hwang, Comparisons among several consistent estimators of structural equation models, Behaviormetrika 45 (1) (2018) 157–188.

[94] A. Diamantopoulos, P. Riefler, K. Roth, Advancing formative measurement models, J. Bus. Res. 61 (12) (2008) 1203–1218.

[95] R. Beran, M. Srivastava, Bootstrap tests and confidence regions for functions of a covariance matrix, Ann. Stat. 13 (1) (1985) 95–115.

[96] C. Fornell, D. Larcker, Evaluating structural equation models with unobservable variables and measurement error, J. Mark. Res. 18 (1) (1981) 39–50.

[97] J. Braojos, J. Benitez, J. Llorens, How do social commerce-IT capabilities influence organizational performance? Theory and empirical evidence, Inf. Manage. 56 (2) (2019) 155–171.

[98] P. Tallon, A. Pinsonneault, Competing perspectives on the link between strategic information technology alignment and organizational agility: insights from a mediation model, Mis Q. 35 (2) (2011) 463–486.

[99] S. Aral, C. Dellarocas, D. Godes, Social media and business transformation: a framework for research, Inf. Syst. Res. 54 (1) (2013) 3–13.

[100] P. Sharma, M. Sarstedt, G. Shmueli, K. Kim, K. Thiele, PLS-based model selection: The role of alternative explanations in IS Research,", J. Assoc. Inf Syst 20 (4) (2019) 346–397.

[101] S. Gregor, The nature of theory in information systems, Mis Q. 30 (3) (2006) 611–642.

[102] H. Simon, Science seeks parsimony, not simplicity: searching for pattern in phenomena, in: A. Zellner, Keuzenkamp H, M. McAleer (Eds.), Simplicity, Inference and Modelling: Keeping It Sophisticatedly Simple, Cambridge University Press, Cambridge, UK, 2001, pp. 32–72.

[103] M. Friedman, The methodology of positive economics, Philos. Econ.: Anthol. 2 (1) (1994) 180–213.

[104] A. Rai, R. Patnayakuni, N. Seth, Firm performance impacts of digitally enabled supply chain integration capabilities, Mis Q. 30 (2) (2006) 225–246.

[105] N. Saraf, C. Schlueter, S. Gosain, IS application capabilities and relational value in interfirm partnerships, Inf. Syst. Res. 18 (3) (2007) 320–339.

[106] J. Churchill, A paradigm for developing better measures of Marketing constructs, J. Mark. Res. 16 (1) (1979) 64–73.

[107] J. Cohen, A power primer, Psychol. Bull. 112 (1) (1992) 155–159.

[108] J. Cohen, Statistical Power Analysis for Behavioral Sciences (2nd Edition), Erlbaum, Hillsdale, USA, 1988.

[109] N. Kock, P. Hadaya, Minimum sample size estimation in PLS-SEM: the inverse square root and gamma-exponential methods, Inf. Syst. J. 28 (1) (2018) 227–261.

[110] M. Aguirre-Urreta, M. Rönkkö, Sample size determination and statistical power analysis in PLS using R: an annotated tutorial, Commun. Assoc. Inf. Syst. 36 (3) (2015) 33–51.

[111] W. Chin, How to write up and report PLS analyses, in: V. Esposito, W. Chin, J. Henseler, H. Wang (Eds.), Handbook of Partial Least Squares: Concepts, Methods and Applications, Springer, Berlin, Germany, 2010, pp. 655–690.

[112] W. Chin, The partial least squares approach for structural equation modelling, in: G. Marcoulides (Ed.), Modern Methods for Business Research, Lawrence Erlbaum Associates, London, UK, 1998, pp. 295–336.

[113] C. Ringle, S. Wende, J. Becker, SmartPLS 3, SmartPLS, Boenningstedt, Germany, 2015www.smartpls.com.

[114] N. Kock, WarpPLS 3.0 User Manual, ScriptWarp Systems, Laredo, USA, 2012.

[115] Addinsoft, Xlstat 2009, Addinsoft, Paris, France, 2009http://www.xlstat.com/en/products/xlstat-plspm/.

[116] J. Henseler, ADANCO 2.0.1. Professional for Windows, Composite Modeling GmbH & Co., Kleve, Germany, 2017www.compositemodeling.com.

[117] J. Henseler, Partial least squares path modeling, in: P. Leeflang, J. Wieringa, T. Bijmolt, K. Pauwels (Eds.), Advanced Methods for Modeling Markets, Springer, Heidelberg, Germany, 2017, pp. 361–381.

[118] D. Gefen, E. Rigdon, D. Straub, An update and extension to SEM guidelines for Administrative and Social Science research, Mis Q. 35 (2) (2011) iii–xiv.

[119] L. Hu, P. Bentler, Y. Kano, Can test statistics in covariance structure analysis be trusted? Psychol. Bull. 112 (2) (1992) 351–362.

[120] J. Nunnally, I. Bernstein, Psychometric Theory, McGraw-Hill, New York, USA, 1994.

[121] C. Fornell, J. Cha, Partial least squares, Adv. Methods Market. Res. 407 (1) (1994) 52–78.

[122] R. Cenfetelli, G. Bassellier, Interpretation of formative measurement in Information Systems research, Mis Q. 33 (4) (2009) 689–707.

[123] J. Hair, M. Sarstedt, L. Hopkins, V. Kuppelwieser, Partial least squares structural equation modeling (PLS-SEM): an emerging tool in Business research, Eur. Bus. Rev. 26 (2) (2014) 106–121.

[124] L. Hayduk, Shame for disrespecting evidence: the personal consequences of insufficient respect for structural equation model testing, BMC Med. Res. Methodol. 14 (124) (2014) 1–10.

[125] M. Aguirre-Urreta, M. Rönkkö, Statistical inference with PLSc using bootstrap confidence intervals, Mis Q. 42 (3) (2018) 1001–1020.

[126] X. Zhao, J. Lynch, Q. Chen, Reconsidering Baron and Kenny: myths and truths about mediation analysis, J. Consum. Res. 37 (2) (2010) 197–206.

[127] J. Becker, A. Rai, E. Rigdon, Predictive validity and formative measurement in structural equation modeling: embracing practical relevance, Proceedings of the 34th International Conference on Information Systems (2013) 1–19.

[128] J. Braojos, J. Benitez, J. Llorens, How do small firms learn to develop a social media competence? Int. J. Inf. Manage. 35 (4) (2015) 443–458.

[129] J. Karimi, T. Somers, A. Bhattacherjee, The impact of ERP implementation on business process outcomes: a factor-based study, J. Manage. Inf. Syst. 24 (1) (2007) 101–134.

[130] J. Karimi, T. Somers, A. Bhattacherjee, The role of information systems resources in ERP capability building and business process outcomes, J. Manage. Inf. Syst. 24 (2) (2007) 221–260.

[131] J. Evermann, M. Tate, Assessing the predictive performance of structural equation model estimators, J. Bus. Res. 69 (10) (2016) 4565–4582.

[132] G. Cepeda, J. Henseler, C. Ringle, J. Roldan, Prediction-oriented modeling in Business research by means of PLS path modeling, J. Bus. Res. 69 (10) (2016) 4545–4551.

[133] J. Cheah, M. Sarstedt, C. Ringle, T. Ramayah, H. Ting, Convergent validity assessment of formatively measured constructs in PLS-SEM: On using single-item versus multi-item measures in redundancy analyses, Int. J. Contemp Hosp. Manage. 30 (11) (2019) 3192–3210.

[134] F. Schuberth, J. Henseler, T. Dijkstra, Confirmatory composite analysis, Frontiers in Psychology 9 (1) (2018) 1–14.

Jose Benitez is Full Professor of IS at Rennes School of Business, France. Jose also holds the COVIRAN-Prodware Chair of Digital Human Resource Strategy at the University of Granada, Spain. Jose is also Instructor of PLS-PM at the PLS School. His research interests cover the study of how the firm's portfolio of IT capabilities affects organizational capabilities and firm performance, and the development of PLS-PM in the field of IS. His research has been published in top IS journals such as *MIS Quarterly*, *Information & Management*, *European Journal of Information Systems*, *Journal of Information Technology*, and *Journal of Business Research*. He currently serves as Associate Editor for *Information & Management*, *European Journal of Information Systems*, and *Decision Support Systems*, and as Guest Editor for *Decision Sciences*. He is also a member of the Editorial Review Board of the *Journal of the Association for Information Systems*. Jose obtained a PhD in Business Administration (with concentration in IS) from the University of Granada, Spain. Jose can be contacted at jose.benitez@rennes-sb.com and joseba@ugr.es.

Jörg Henseler holds the Chair of Product-Market Relations in the Faculty of Engineering Technology at the University of Twente, the Netherlands, and he is Visiting Professor at NOVA Information Management School, Universidade Nova de Lisboa, Portugal. His

broad-ranging research interests encompass empirical methods of Marketing and Design research as well as the management of design, products, services, and brands. His work has been published in *Computational Statistics and Data Analysis*, *European Journal of Information Systems*, *International Journal of Research in Marketing*, *Journal of the Academy of Marketing Science*, *Journal of Supply Chain Management*, *MIS Quarterly*, *Organizational Research Methods*, and *Structural Equation Modeling-An Interdisciplinary Journal*, among others. He chairs the Scientific Advisory Board of ADANCO, software for variance-based SEM (www.composite-modeling.com).

Ana Castillo is a PhD Student of IS in the Department of Management at the University of Granada, Spain. In her doctoral dissertation, she examines how firms leverage social media capabilities to pursue knowledge management and innovation activities. She holds a master's degree in Management from the University of Granada. Her research has been published in *Information & Management*. She has presented her research in the *Academy of Management Annual Meeting, European Conference on Information Systems, Americas Conference on Information Systems*, and *Pacific Asia Conference on Information Systems*.

Florian Schuberth obtained his PhD in Econometrics in the Faculty of Business Management and Economics at the University of Wuerzburg, Germany. Currently, he is Assistant Professor in the Faculty of Engineering Technology at the University of Twente, the Netherlands. His main research interests are focused on SEM, in particular on composite-based estimators and their enhancement.