International Journal of Business and Society, Vol. 19 No.1, 2018, 139-160

ASSESSING REFLECTIVE MODELS IN MARKETING RESEARCH: A COMPARISON BETWEEN PLS AND PLSc ESTIMATES

Jun-Hwa Cheah♣
Universiti Teknologi Malaysia

Mumtaz Ali Memon
Universiti Teknologi PETRONAS

Francis Chuah
Universiti Utara Malaysia

Hiram Ting
Universiti Malaysia Sarawak

T. Ramayah
Universiti Sains Malaysia

ABSTRACT

The present study aims at contributing to the growing discourse on analytical methods in marketing research by highlighting the use of Consistent Partial Least Squares (PLSc) estimation to assess reflective models used in marketing literature. Specifically, it demonstrates the significance of using PLSc and compares it with the traditional PLS. The results show that PLSc is more robust than traditional PLS in estimating convergent validity and path coefficients, and yields better power – coefficient of determination (R^2) and effect size (f^2). It is also found that PLSc generates better holdout results than traditional PLS. This study complements and extends prior research on PLSc, and subsequently serves as a resource for marketing researchers who use variance-based approach in their research. Implications, guidelines and future research directions are discussed.

***Keywords*:** Consistent Partial Least Squares; Traditional PLS; Path Modeling; SEM; Marketing

1. INTRODUCTION

Structural Equation Modeling (SEM) technique has become one of the most powerful statistical techniques across various disciplines in recent years. An increasing number of researchers have begun to recognize its ability to model latent variables, taking into account the various forms of measurement errors, and test the underlying theories in a structural manner (Pakpahan et al., 2017). Thus, they stand to benefit from this technique by acquiring more reliable and valid findings to answer respective research questions with accurate estimation.

♣Corresponding author: Azman Hashim International Business School, Universiti Teknologi Malaysia, 54100 Kuala Lumpur, Malaysia. Email: junhwa@ibs.utm.my/jackycheahjh@gmail.com

There are two types of SEM techniques, namely covariance-based SEM (CB-SEM) (Jöreskog and Wold, 1982) and variance-based SEM (VB-SEM) (Lohmöller 1989; Wold, 1982; Hair et al., 2016; Memon, Ting, Ramayah, Chuah, & Cheah, 2017). Despite complementing each other, it is necessary to know that they differ greatly in their statistical methods, and have distinct goals and requirements (Hair et al. 2011; Henseler et al. 2009). In general, CB-SEM estimates the model parameters by means of the empirical covariance matrix. It is more often the preferred method if the hypothesized model consists of one or more common factors. VB-SEM, however, creates proxies as linear combinations of observed variables, and uses them to estimate the parameters. It is usually the method of choice if the hypothesized model contains composites.

Among the VB-SEM methods available to date, partial least squares (PLS) path modelling is regarded as the "most fully developed and general system" (McDonald 1996, p. 240) and has thus been labelled "silver bullet" in a meaningful manner (Hair et al., 2011). It has become one of the most adopted techniques in business and management disciplines such as accounting (Lee et al., 2011), family business (Sarstedt et al., 2014b), management information systems (Ringle et al., 2012), marketing (Hair et al., 2012a), operations management (Peng & Lai 2012), strategic management (Hair et al., 2012b), psychology (Willaby et al., 2015) and tourism (do Valle & Assaker 2016). It is generally agreed that PLS is capable of handling complex models, requires less demand on data distribution and is preferred when the aim of the study is on theory development and prediction. In line with its rising popularity, researchers continue to call for more rigorous assessment on its estimation procedures (e.g., Hair et al., 2013; Rigdon et al., 2014; Sarstedt et al., 2014a) to address the possible problems such as false positive (Cohen, 1988) and estimation bias (Chang et al., 2010). Subsequently, new estimation techniques and methodologies are constantly developed to address these concerns (see: Dijkstra & Henseler, 2015a, b; Henseler et al., 2015; Vorhees et al., 2016).

Generally, factor model has been the dominant measurement model to measure latent constructs, such as attitude or personality traits. The mechanism of this model is that it advocates for the variance of a given number of indicators to be perfectly explained by the existence of one latent variable (i.e., the common factor) and individual random error. According to McDonald (1996), factor model addresses true score theory and is regarded as an important measurement paradigm in behavioural sciences. Specifically, when a construct has a factor model background and random measurement error is likely to be an issue, researchers are advised to use the common factor model estimation method by means of CB-SEM rather than composite model estimation method by means of PLS.

Nonetheless, recent breakthroughs in PLS estimation method provide researchers with alternative solution to assess common factor model especially when distributional assumption and model identification by means of CB-SEM could not be addressed. This method is known as PLS consistent (PLSc), which produces estimations that mimic CB-SEM result. Nevertheless, Sarstedt et al. (2016) articulate that the introduction of PLSc has resulted in confusion among many researchers. This is evident when data analysis is done using both the traditional PLS and PLSc on the same data, without acknowledging the fundamental differences of the measurement philosophies (Sarstedt, 2016, p. 3999). To address this confusion, a simulation study was conducted to examine the efficacy of both traditional PLS and PLSc in estimating common factor model population, with measurement model specification as an effect indicator (Sarstedt et al., 2016). The results suggests that the use of the former is more likely to result in a better approximation because it entails practically little bias in parameter estimation to estimate the common factor model, regardless of the number of indicators, the quality of loadings (i.e., 0.50) and the number of observations. Even if the number of observations and indicators were to increase, the differences are reportedly marginal in comparison with the latter.

Sarstedt and colleagues thus confirm that the traditional PLS is consistent at large trait. In contrast, PLSc performs better under the circumstances of having more indicators, higher loadings, and relatively bigger sample size (>250). Naturally, such simulation studies would reinforce the use of traditional PLS as the preferable SEM method (i.e., when the data's nature is that of a common factor model).

Conversely, Dijkstra and Henseler (2015a, b) postulate that the use of the traditional PLS may lead to inconsistency of PLS path coefficient estimates in reflective measurement model, thus resulting in adverse consequences for hypothesis testing. They further articulate that the use of the traditional PLS tends to overestimate factor loadings (McDonald, 1996) and approximating latent variables with composites leads to the well-known measurement error bias the estimates of structural relationships (Bollen, 1989). Conversely, PLSc can correct the original latent variable correlations for attenuation. This issue is of utmost importance since it would ultimately influence the final estimation in obtaining the true value of path coefficients. Notably, the underestimation of true parameter leads to the occurrence of Type II error while overestimation of true parameter results in Type I error, thus leading to erroneous conclusions. Therefore, from the theoretical and practical perspectives, it is imperative for researchers who are in the methodological frontier to compare and comprehend the advantages of both the traditional PLS and PLSc approaches using real data scenarios in marketing research to further articulate these assertions.

The main objective of this study is to demonstrate and elucidate the assessment of the PLSc estimation and its capability to generate more robust results than that of the traditional PLS approach. Thus, the paper shows how PLSc can be practically employed for the explanation of the relationships among target constructs, the reflective models. In doing so, this is the first study that does not only highlights a clear distinction between PLSc and traditional PLS when estimating the common factor model but also provides guidelines for researchers to select appropriate PLS-SEM estimation method.

2. THE IMPORTANCE OF PLSc COMPARE TO THE TRADITIONAL PLS

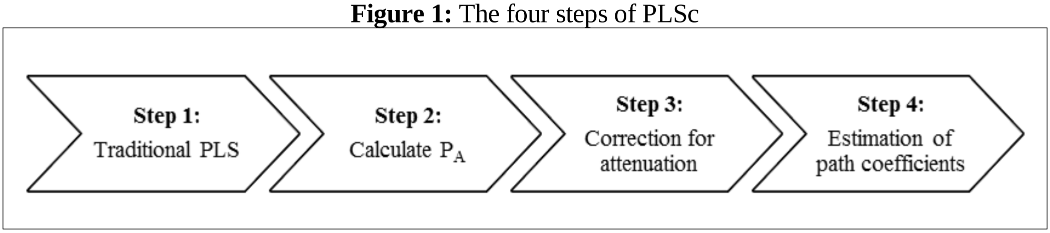
Basically, in the traditional PLS algorithm, once the weights are derived from the independence of the epistemic relationships between constructs and their observed indicators, the method always produces a single specific (i.e., determinate) score for each case per composite (Henseler, 2017a). By using these scores as input, traditional PLS-SEM applies a series of ordinary least squares regressions, which estimate the model parameters such that they maximize the endogenous constructs' explained variance (i.e., R^2 values). Evermann and Tate's (2016) as well as Becker et al. (2013) emphasized that such estimation procedure allows researchers "to work with an explanatory or prediction of a theory-based model, to aid in theory development, evaluation, and selection. In addition, since traditional PLS-SEM-based model estimation always relies on composites, regardless of the measurement model specification, the method can process reflectively and formatively specified measurement models without identification issues (Hair et al., 2016; Ramayah, Cheah, Chuah, Ting & Memon, 2018).

However, one major issue is plethora of researchers today are still confused and perplexed with the depicted direction of arrows in the measurement model to determine a reflective and a formative model. Notably, in traditional form of PLS, the measurement model (reflective or formative) does not indicate whether a factor model or composite model is estimated, but whether correlation weights (Mode A, represented by arrows pointing from a construct to its indicators) or regression weights

(Mode B, represented by arrows pointing from indicators to their construct) shall be used to create the proxy (Becker et al., 2013). Worse still, many researchers still assume that the use of traditional PLS can estimate true reflective measurement model (rather than estimating true reflective/common factor) when in fact, it aggregates the observed variables to form a composite score (Henseler, 2017a). All these eventually create ambiguity among marketing scholars of what is actually a true factor model of reflective measurement.

If the constructs are meant to be a true factor model (reflective measurement), the use of traditional PLS estimation will generate inconsistent estimates, which may lead to flawed theoretical conclusions (Henseler et al., 2014). This is a serious issue for marketing scholars because the misspecification of the estimation procedure in PLS-SEM will definitely lead to erroneous conclusions in marketing studies, as well as unsound marketing implications (van Riel, Henseler, Kemény & Sasovova, 2017). As a remedy, Dijkstra and Henseler (2015a, b) introduced PLSc, which is a better solution to estimating factor model (reflective measurement). PLSc estimation consists of four steps as shown in Figure 1. These steps are: (1) The traditional PLS is employed to identify latent variable scores, and to estimate latent variable correlations and weights, (2) The new reliability coefficient ρA is determined for each reflective construct, (3) ρA can be used to correct the original latent variable correlations for attenuation and may thus obtain consistent latent variable correlations, and (4) consistent path coefficients are estimated in a least squares manner based on the consistent latent variable correlations.

Figure 1: The four steps of PLSc



The use of PLSc is believed to help rectify the inconsistent results, generated by traditional PLS algorithm. These inconsistencies are in relation to the estimations, i.e., disproportionate true values when an increase in sample size is observed. PLSc does not only estimate the path coefficients, inter-construct correlations, and indicator loadings in reflective models consistently, but is also said to have achieve the true values asymptotically. In other words, PLSc corrects inter-construct correlations for attenuation so that the estimates of path coefficients and loadings become consistent as compared to the traditional PLS estimation (Dijkstra 2010; Dijkstra & Henseler 2015a, b).

Dijkstra and Henseler (2015a) argued that the PLSc estimation has the capability to compare statistical power of the traditional PLS as well as CB-SEM. Authors point out that the development of PLSc algorithms (Bentler & Huang, 2014; Dijkstra, 2014; Dijkstra and Henseler, 2015a) has the potential to fully mimic CB-SEM. In their simulation study, they found out that PLSc is only slightly lower in power than full information maximum likelihood in CB-SEM but it can more or less produce identical results of loading, path coefficients and explanatory power, thereby offering an opportunity to fill the gap between factor models and composite models (Goodhue et al., 2006; Lu et al., 2011; Reinartz et al., 2009). Nevertheless, PLSc has the edge of being able to handle non-normally distributed data like the use of traditional PLS.

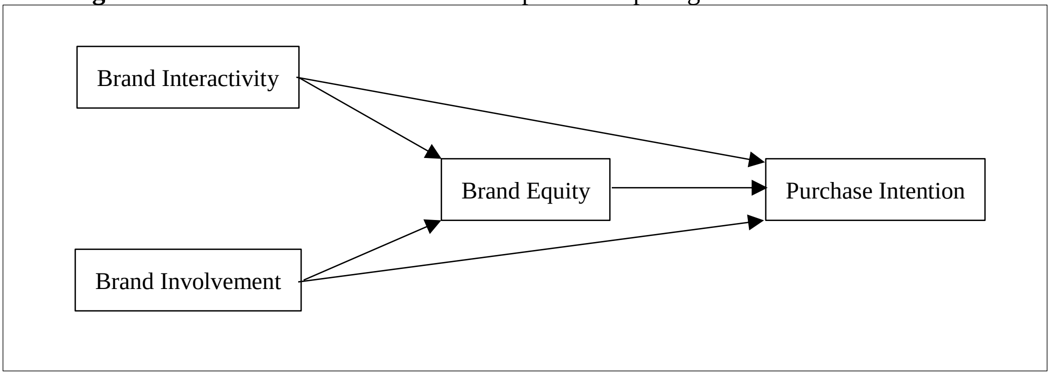
Furthermore, the use of the traditional PLS algorithm "tends to overestimate the loadings in absolute value, and underestimate multiple and bivariate (absolute) correlations between the latent variables" (Djikstra & Henseler, 2015b, p.11). The advantage of using PLSc, however, is that it is 'well calibrated', i.e., "it will produce the true parameter values for the models we discuss when applied to the 'population' (Dijkstra & Schermelleh-Engel 2014, p.586). Besides, the use of the traditional PLS may also underestimate the R-squared value of endogenous latent variables (Dijkstra, 2010), which is regarded as a major error in research. Hence, without any correction of the parameter estimate, it may lead to more errors of probable attenuation of inter-construct correlations thus resulting in inconsistent estimation of results (Goodhue et al., 2012; Dijkstra & Henseler 2015a, b). PLSc, however, is designed with the aim to increase the power and to reduce Type II error in reflective models. In other words, researchers are less likely to reject a true model. Inconsistency of estimates associated with the traditional PLS may also imply that in the case of PLS overestimating these parameters, Type I error is more likely to occur in the traditional PLS than PLSc.

Dijkstra and Henseler (2015b) conclude that the PLSc approach is considered the least (less) problematic for non-recursive reflectively-modeled linear models. Therefore, it is ideally meant for true common factor models where constructs are reflectively measured. Nonetheless, when indicator correlations are not informative for gauging the reliability in reflective and formative models, PLSc is not appropriate, and the use of traditional PLS estimation will be preferred. (Dijkstra & Henseler 2015a, b).

3. AN ILLUSTRATED EXAMPLE TO DEMONSTRATE THE VIABILITY OF PLSc COMPARED TO THE TRADITIONAL PLS: MODEL, MEASURES, AND SAMPLE

Considering the state-of-the-art development of PLSc, the example in the present study comprises data on the constitution, effects of customization and customer equity on telecommunication service in the context of Malaysia (n = 202). The underlying model tests the effect of the antecedents of brand interactively, brand involvement and brand equity on purchase intention as shown in Figure 2. A quantitative research design was employed by means of self-administered questionnaires.

Figure 2: A model that illustrates the example of comparing PLSc and Traditional PLS



The final questionnaire was adapted from the telecommunication services study consisting predominantly a 5-point Likert scale measuring brand involvement (5 items), brand interactivity (5 items) (Merrilees & Fry, 2003; Labrecque, 2014) and brand equity (4 items) (Verhoef et al., 2007).

Purchase intention (3 items), on the other hand, were measured using a 7-point Likert scale (Liang & Lai 2002).

The rationale for using different scales of measure in the survey is to minimize the possibility of the respondents having mental map judgment about the response categories provided by the questions within all sections of the instrument using the same scale. It is one of the recommended procedural remedies to address common method variance (CMV) issue (Podsakoff et al., 2003; Podsakoff & Organ 1986).

A pre-test of the instrument using five respondents from the target population was conducted. The purpose was to elicit comments from them about any potential miscomprehension and flaws related to the questionnaire's format, design and wording (Hair et al., 2010; Memon et al. 2017). After the questionnaire was revised based on those comments, it was further pilot-tested with 30 respondents to allow for error identification and preliminary scale optimization.

A non-probability purposive sampling method was employed, in which, the target respondents must subscribe to either one of the three major telecommunication companies, namely Digi Bhd., Maxis Bhd., and Celcom Axiata Bhd. in the telecommunication industry in Malaysia for more than three years. The questionnaire was distributed by means of the mall intercept approach (Bush & Hair 1985). This approach ensured diversity in respondents' characteristics, including gender (Male: 39.6% and Female: 60.4%), ethnicity (Malay: 52.0%, Chinese: 34.7% and Indian: 13.3%), age (21-30: 79.7%, 31-40: 15.3% and 41-50: 5.0%) and level of income (Less than RM2000: 55.0%, RM2001–RM4000: 25.7%, RM4001-RM6000: 16.3% and above RM6001: 3.0%). Notably, the split distribution strategy provided the researchers with access to wider groups without hampering the data (Huang 2006; Lin & Van Ryzin 2012).

Respondents were asked to complete the survey for the brand they were patronizing. Such instruction was necessary to lead their thoughts to the brands of interest, as it would increase the level of participation and return rates and reduce response errors. The completed questionnaires were subsequently collected from 220 respondents, and the data were keyed into SPSS version 25. An examination on missing and erroneous data was conducted and 18 cases were removed due to serious missing data. Thus, the total sample usable for data analysis comprised 202 responses, exceeds the minimum sample size required to achieve a power of $(1-\beta) = 80\%$, margin error of 5%, effect size of 0.15, and a maximum of three number of predictors in a priori power analysis (Faul et al., 2007).

Lastly, in addition to procedural remedy, Harman's Single Factor technique was performed to examine CMV (Podsakoff et al., 2003) before analysing the data with VB-SEM technique in ADANCO 2.0. The result indicated that the largest variance explained by the first factor was 31.182% of the total variance (< 40%; suggested by Babin, Griffin & Hair, 2016). Therefore, it is concluded that no general factor emerged from the factor analysis, which means that the common method bias was not an issue in this data set.

4. ASSESSING AND REPORTING OF PLS ANALYSIS

Notwithstanding a plethora of software (i.e., WarpPLS, SmartPLS, PLS-GUI and XL-STAT) to assess VB-SEM, this study utilizes a relatively new software package, namely ADANCO 2.0.1 (Henseler & Djikstra 2015; Henseler, 2017b), to analyse the data. VB-SEM method promises favourable convergence behaviour (Henseler, 2010) and composite scores of all constructs. It also allows the handling of constructs' reflective nature in PLSc mode (Djikstra & Henseler 2015b). Nonetheless, researcher has to be cautious about model identification of VB-SEM (Henseler et al., 2016a). Every construct requires a nomological net. In other words, every construct is expected to have certain level of relationship with at least another construct in the model. A strong theoretical basis and relevant past empirical evidences for the development of conceptual model are always the keys to justifying the identification.

The overall goodness-of-fit (GoF) of the model should be the starting point of model assessment. If the model does not fit the data, the data contains more information than the model conveys. As a result, the obtained estimates may be meaningless, and the conclusions drawn from them become questionable. The global model fit can be assessed in two non-exclusive ways: by means of inference statistics, i.e. so-called tests of model fit, or using fit indices, i.e. an assessment of approximate model fit. Bootstrap-based tests of the model fit over the unweighted least squares (dULS) and the geodesic discrepancy (dG) between the empirical and the model-implied correlation matrix allow the assessment of the global goodness of the model fit (Dijkstra & Henseler, 2015a). If the discrepancy between these two matrices points to an insignificant result, researchers may not need to reject the model. Furthermore, as a measure of approximate fit, the standardized root mean square residual (SRMR) may help quantify the degree of (mis-)fit (Henseler et al., 2014). The SRMR of well-fitting models typically do not exceed a value of 0.08 (Hu & Bentler, 1999). In addition, both global model fit indices have become customary to determine the model fit both for the estimated model and for the saturated model. Saturation refers to all constructs correlate freely in the structural model. The estimated model is based on a total effect scheme and it takes the model structure into account.

In any quantitative survey, the indicators are expected to demonstrate sufficient reliability. According to Nunnally and Berstein (1994), the recommended reliability value as low as 0.7 indicated proper reliability in early phases of research. A higher value, such as 0.8 or 0.9, should prevail in more advanced research. However, these reliability measures have also been found to be inconsistent. Conventionally, Jöreskog's rho (also known as Composite Reliability) and Cronbach's alpha are used to assess internal consistency. Both reliability measure use sum scores, rather than construct scores (Henseler et al., 2016a). Specifically, Cronbach's alpha tends to underestimates the true reliability, and should therefore be regarded as a lower boundary to the reliability (Sijtsma, 2009). On the other hand, Jöreskog's rho tends to overestimate the internal consistency reliability and should therefore be regarded as an upper boundary to the reliability (Hair et al., 2016). As such, another reliability measure, namely ρA (Henseler et al. 2016a), is gradually considered as the most important and consistent measure of internal consistency reliability when performing VB-SEM.

Apart from reliability, validity assessment is equally important in any business research. There are two types of validity criteria assessment, namely the average variance extracted (AVE) and the discriminant validity. AVE serves as a measure of unidimensionality (Fornell & Larcker 1981). If the first factor, which is extracted from a set of indicators, appears to explain more than one-half of their variance, there may not be any second andequally important factor. Therefore, an AVE of 0.5 or higher is regarded as acceptable (Bagozzi & Yi, 1988; Fornell & Larcker, 1981). In terms of discriminant validity, it is suggested that the heterotrait–monotrait ratio of correlations (HTMT) be adopted as a better-suited criterion to assess discriminant validity (Henseler et al., 2015). Monte Carlo simulations also show that the HTMT outperforms more traditional measures, such as that of the Fornell and Lacker (1981) criterion and the cross-loading technique (Voorhees et al., 2016). There are two ways to assess discriminant validity using HTMT: (1) as a criterion or (2) as a statistical test.

With regard to the criterion, if the HTMT value is greater than $HTMT_{.85}$ value of 0.85 (Kline 2011), or $HTMT_{.90}$ value of 0.90 (Gold et al. 2001), researchers would have to deal with the problem of discriminant validity. The second way is to test the HTMT inference from Henseler et al., (2015) by running the null hypothesis (where H0: HTMT $\geq$ 1) against the alternative hypothesis (where H1: HTMT < 1). If the confidence interval indicates the value of one (*i.e.*, H_0 holds), it may imply that there is a lack of discriminant validity (Henseler et al. 2015).

When assessing structural models, it is vital to assess that there are no collinearity issues in the structural model (or inner model). This typically occurs when two variables, which are hypothesized to have a causal relationship end up measuring the same construct. In other words, predictor-criterion collinearity issue may sometimes subtly mislead the findings as it can weaken the strong causal effects in the structural model. In order to assess the collinearity issue, VIF should not be greater than 5, or else it indicates that there might be a potential collinearity problem (Hair et al., 2011). Researchers can also opt for a more stringent criterion where VIF should be lower than 3.3 (Diamantopoulos & Siguaw, 2006).

The path coefficients are the most important results of a structural model. They indicate the change in a dependent variable resulting from a unit change in an independent variable with the condition that all the other independent variables remain constant. Bootstrap percentile confidence intervals of the path coefficients help generalize from the sample to the target population; and are more preferred than the mere null hypothesis significance testing (Hair et al., 2016). Hence, bootstrap percentile confidence intervals are used to examine the relationships between constructs.

In addition, the predictive power of the research model can be evaluated by means of the coefficient of determination score (R^2). The R^2 is a measure of the model's predictive accuracy and it can also be viewed as the combined effect of the exogenous variables on endogenous variables. In other words, R^2 is assessed as the main objective of PLS, which is to maximize the variance explained in the endogenous variables. This effect ranges from 0 to 1 with 1 representing complete predictive accuracy. Since R^2 is embraced by a variety of disciplines, researchers are advised to rely on a "rough" guide in relation to an acceptable R^2 — 0.26, 0.13, 0.02, indicating substantial, moderate, or weak levels of predictive accuracy respectively (Cohen 1988). It is however mandatory that R^2 values should be high enough for the model to achieve a minimum level of explanatory power (Urbach & Ahlemann 2010).

The effect size of the predictor constructs can also be evaluated by means of Cohen's f^2 (Cohen, 1988). The effect size (f^2) is a measure used to assess the relative impact of a predictor construct on an endogenous construct (Cohen 1988). It analyzes how much a predictor construct contributes to the R^2 value of a target construct in a structural model. The R^2 value is estimated with a particular predecessor construct. If one of the predecessor constructs is excluded, the result for R^2 value will become lower. As such, the difference in the R^2 values for estimating the model with and without the predecessor construct is known as the effect size (f^2). According to Cohen (1988), f^2 values of 0.35, 0.15, and 0.02 are considered large, medium, and small effect sizes respectively. If an exogenous construct strongly contributes to explaining an endogenous construct, the difference between R^2 included and R^2 excluded will be relatively higher, eventually leading to a high f^2. The effect size is calculated using the formula as shown in Equation 1.

$$f2 = \frac{R^2\ included - R^2 excluded}{1 - R^2\ included} \qquad (1)$$

To assess the explanatory validity of model results by means of PLS, the hold-out sample method was assessed (Hahn & Ang, 2017; Woodside 2013). The purpose of hold-out sample assessment is to determine how well an explanatory model can perform in practice (Ebbes et al., 2011). It helps researchers to find a balance between model fit and prediction capability (Schorfheide & Wolpin, 2012). When selecting the types of analysis and hold-out samples, the present study adopts the suggestions by Hair et al. (2010), whereby the total sample size is divided randomly so that a two-third of the observations are placed in the training sample and a one-third in the hold-out sample (rule of 60-40).

5. DATA ANALYSIS AND RESULTS

The results show that the conceptual model has an excellent fit for reflective measurement by means of PLSc (Mode A consistent) instead of the traditional PLS (Mode A). Table 1 demonstrates that all discrepancies for PLSc, both saturated and estimated models, do not exceed the 99% percentile of their bootstrap distribution. In other words, the empirical and the model-implied correlation matrices do not indicate any significant differences. As for the traditional PLS (Mode A), the dULS result for both saturated and estimated models are observed to have exceeded the 99% quantiles of their bootstrap distributions. Moreover, the fit values for SRMR (saturated and estimated) indicate that both PLSc (Mode A consistent) and the traditional PLS (Mode A) lie clearly well below the common cut-off thresholds of 0.08 (Hu & Bentler, 1999). However, unlike the PLSc (Mode A consistent), the SRMR value for the traditional PLS (Mode A) fails to fulfil both the bootstrap-based 95% and 99% for both saturated and estimated models respectively. In addition, the SRMR fit value also indicates that PLSc (Mode A consistent) performs better (SRMR = 0.051) compared to the traditional PLS (Mode A) (SRMR=0.074). In general, it can be surmised that PLSc (Mode A consistent) yields better performance for a reflective model (true common factor model) in the assessment of goodness of fit and is thus, less likely to be affected by any misspecification in the model compared to the traditional PLS (Mode A).

Table 1: Global Goodness of Fit and Bootstrap

Reflective Measurement Model	Saturated Model	Value	HI95	HI99	Estimated Model	Value	HI95	HI99
Traditional PLS - Mode A	SRMR	0.074	0.061	0.064	SRMR	0.074	0.061	0.064
	dULS	0.649	0.443	0.491	dULS	0.649	0.443	0.491
	dG	0.282	0.300	0.333	dG	0.282	0.300	0.333
PLSc - Mode A consistent	SRMR	0.051	0.046	0.051	SRMR	0.051	0.046	0.051
	dULS	0.232	0.193	0.240	dULS	0.232	0.193	0.240
	dG	0.126	0.129	0.160	dG	0.126	0.129	0.160

The construct measurement indicates a fair reliability and convergent validity as shown in Table 2. Dijkstra–Henseler's ρ exceeds 0.7 in all instances (the traditional PLS and PLSc), which points to relatively higher internal consistency reliability of the construct scores. However, brand Interactivity for PLSc shows slightly unsatisfactory result as illustrated by Cronbach's alpha (α) (0.685) and Jöreskog's rho (ρc) (0.688). The table also entails the individual item loadings that are retained in this study. The results of the outer loadings show that two items (Interact 5 and Involve 5) have to be deleted for the traditional PLS in comparison with four items (Interact 1, Interact 4, Interact 5 and Involve 5) for the PLSc. As such, it can be inferred that the use of the traditional PLS (Mode A) might result in an inflated estimation of the loadings (Djikstra 1983), and thus lead to inaccurate conclusions pertaining to path coefficients parameters. Since all AVE values for the traditional PLS and PLSc exceed the value of 0.5, thus there is no second factor of equal importance to confound the first factor. The study-wide maximum HTMT for the traditional PLS and PLSc is 0.803 and 0.812 respectively, which is well below the strictest threshold of 0.85 (Kline, 2011), thus confirming the discriminant validity of measurement (see Table 3).

Table 2: Reliability and Convergent Validity

Construct	Items	Outer Loading	AVE	Cronbach's alpha (α)	Dijkstra–Henseler's rho (ρA)	Jöreskog's rho (ρc)
			Traditional PLS (Mode A)			
Brand Interactivity	Interact 1	0.660	0.572	0.749	0.761	0.841
	Interact 2	0.819				
	Interact 3	0.765				
	Interact 4	0.771				
	Interact 5	Item deleted				
Brand Involvement	Involve 1	0.811	0.657	0.826	0.828	0.885
	Involve 2	0.801				
	Involve 3	0.836				
	Involve 4	0.794				
	Involve 5	Item deleted				
Brand Equity	BE1	0.802	0.661	0.829	0.838	0.886
	BE2	0.890				
	BE3	0.777				
	BE4	0.777				
Purchase Intention	PI1	0.845	0.719	0.805	0.807	0.885
	PI2	0.847				
	PI3	0.853				
			PLSc (Mode A Consistent)			
Brand Interactivity	Interact 1	Item deleted	0.525	0.685	0.700	0.688
	Interact 2	0.770				
	Interact 3	0.676				
	Interact 4	Item deleted				
	Interact 5	Item deleted				
Brand Involvement	Involve 1	0.750	0.544	0.826	0.828	0.826
	Involve 2	0.693				
	Involve 3	0.783				
	Involve 4	0.721				
	Involve 5	Item deleted				
Brand Equity	BE1	0.686	0.552	0.829	0.842	0.829
	BE2	0.855				
	BE3	0.789				
	BE4	0.617				
Purchase Intention	PI1	0.724	0.579	0.805	0.807	0.805
	PI2	0.799				
	PI3	0.759				

Table 3: HTMT for Discriminant Validity

Traditional PLS	**Construct**	**Brand Involvement**	**Purchase Intention**	**Brand Equity**	**Brand Interactivity**
Mode A	Brand Involvement				
	Purchase Intention	0.663			
	Brand Equity	0.455	0.737		
	Brand Interactivity	0.711	0.803	0.785	
	Construct	Brand Involvement	Purchase Intention	Brand Equity	Brand Interactivity
PLSc Mode A Consistent	Brand Involvement				
	Purchase Intention	0.663			
	Brand Equity	0.455	0.737		
	Brand Interactivity	0.668	0.812	0.709	

***Note**:* HTMT < 0.85 (Kline, 2011), HTMT < 0.90 (Gold et al. 2001)

Prior to evaluating the structural model, researchers have to ensure that there are no collinearity issues in the inner model. Table 4 presents the outcome of the collinearity test for structural model. The VIF scores for each individual construct in the first and second set for the traditional PLS (Mode A) and PLSc (Mode A consistent) are below the threshold value of 3.3 (Diamantopoulos & Sigouw, 2006), thus implying that there are no inner collinearity issues.

Table 4: Collinearity Assessment for Structural Model

	First Set (Brand Equity)		**Second Set (Purchase Intention)**	
	Constructs	**VIF**	**Constructs**	**VIF**
Traditional PLS (Mode A)	Brand Interactivity	1.444	Brand Interactivity	1.884
	Brand Involvement	1.444	Brand Involvement	1.942
			Brand Equity	1.909
PLSc (Mode A consistent)	Brand Interactivity	1.632	Brand Interactivity	2.017
	Brand Involvement	1.632	Brand Involvement	1.449
			Brand Equity	1.638

Note: VIF < 3.3; First set are tested on dependent variables of brand equity and second set are tested on dependent variable of purchase intention

The direct effect of the traditional PLS (Mode A) and PLSc (Mode A consistent) exhibits similar results except for Brand Equity on Brand Involvement as shown in Table 5. However, when looking at the path-coefficient results for the structural model relationships, it is evident that PLSc is able to highlight the true values of the relationships in the PLS path model compared to the traditional PLS. With regard to the R^2, the results also infer that PLSc (Mode A consistent) has the capacity to perform better than the traditional PLS (Mode A). Lastly, there is difference in effect sizes, whereby PLSc (Mode A consistent) tends to perform better than the traditional PLS (Mode A). Therefore, it can be concluded that using PLSc (Mode A consistent) would help researchers draw more credible conclusions because it addresses issues related to both over and under estimation (i.e., Type 1 and Type 2 errors) of the path-coefficient results.

Finally, Table 6 reports the R^2 values of the hold-out sample and compares them with the R^2 values obtained in the training sample. Both R^2 values are fairly similar for Traditional PLS (Mode A) and PLSc (Mode A consistent). Cross-validations of this kind substantiate how the statistical analysis results could be generalized to another independent data set, as well as how well an explanatory model may perform in practice. Interestingly, the results further explain that the traditional PLS (Mode A)

Table 5: Structural Model Results, R^2 and f^2

Relationship	Standard bootstrap results					Percentile bootstrap quantiles				R^2	f^2
	Std. coefficient	**Std. error**	**t-value**	**p-value (2-sided)**	**p-value (1-sided)**	**0.50%**	**2.50%**	**97.50%**	**99.50%**		
Traditional PLS (Mode A)											
Brand Interactivity → Brand Equity	0.592	0.059	10.012**	0.000	0.000	0.420	0.470	0.706	0.728	0.390	0.397
Brand Involvement → Brand Equity	0.056	0.057	0.972	0.332	0.166	-0.086	-0.055	0.177	0.202		0.004
Brand Interactivity → Purchase Intention	0.271	0.069	3.936**	0.000	0.000	0.106	0.142	0.417	0.456	0.521	0.076
Brand Involvement → Purchase Intention	0.256	0.051	5.061**	0.000	0.000	0.133	0.161	0.362	0.389		0.095
Brand Equity → Purchase Intention	0.347	0.079	4.404**	0.000	0.000	0.124	0.183	0.492	0.531		0.154
PLSc (Mode A consistent)											
Brand Interactivity → Brand Equity	0.711	0.146	4.863**	0.000	0.000	0.381	0.464	1.012	1.000	0.498	0.568
Brand Involvement → Brand Equity	-0.009	0.132	-0.066	0.947	0.474	-0.435	-0.307	0.212	0.273		0.000
Brand Interactivity → Purchase Intention	0.415	0.222	1.872*	0.061	0.031	-0.127	0.052	0.881	1.327	0.746	0.244
Brand Involvement → Purchase Intention	0.225	0.116	1.940*	0.053	0.026	-0.274	-0.020	0.411	0.459		0.113
Brand Equity → Purchase Intention	0.349	0.167	2.094*	0.036	0.018	-0.383	-0.001	0.631	0.760		0.241

tends to perform better in hold-out sample compared to training sample. In contrast, PLSc (Mode A consistent) performs better in training sample compared to hold-out sample but in more identical explanatory power.

Table 6: Holdout Sample

	Endogenous	Training Sample R^2 (137 sample size)	Holdout Sample R^2 (67 sample size)
Traditional PLS (Mode A)	Brand Equity	0.332	0.477
	Purchase Intention	0.453	0.661
PLSc (Mode A consistent)	Brand Equity	0.423	0.358
	Purchase Intention	0.687	0.557

6. DISCUSSION AND CONCLUSION

The present study discusses manifold findings to serve its purpose. Firstly, fit measures, such as the SRMR and bootstrap-based tests of the model fit (dULS and dG), play an important role in guiding researchers to assess whether the data follow a common factor model or composite model (Djikstra & Henseler, 2015b; Sarstedt et al., 2016). If the specific measurement does not meet the required level of goodness of fit, this denotes that the data may exhibit the characteristics of a composite model. The findings also imply that when a common factor (reflective) model is present, PLSc (Mode A consistent) is very likely to produce better goodness of fit than that of the traditional PLS (Mode A). Conversely, if the specific measurement model does not meet the required level criterion of the SRMR and the test of exact model fit, this result suggests that the data follow a composite model. Hence, when researchers use a common factor (reflective) model that focuses on behavioural constructs, i.e., Brand Involvement, Brand Interactivity, Brand Equity and Purchase Intention in this study, they may need to consider using PLSc estimation (Mode A consistent) in their study. Thus, the use of global model fit can indicate whether the data are coherent with a true common factor model or composite model.

Secondly, the findings provide substantial evidence that the use of the traditional PLS algorithm (Mode A) is inclined to produce inaccurate convergent validity results. One plausible reason for this could be that the traditional PLS algorithm (Mode A) is likely to overestimate the loadings in absolute value. Hence, the results of AVE may also be overestimated. PLSc, however, employs a correction factor to obtain consistent indicator loadings if the common factor model holds true (Djiksra & Henseler 2015b). Hence, the deletion of indicators in models using the traditional PLS (Mode A) will always result in minimal condition compared to the deletion of indicators for PLSc (Mode A consistent). In addition, since PLS algorithm (Mode A) tends to overestimate the loadings, the results of AVE will in turn also be equally overestimated.

Thirdly, the difference in consistency estimation of construct reliability can only occur when researchers use PLSc (Mode A consistent) and not the traditional PLS (Mode A) as shown in Table 2. This is consistent with the assertions made by Djikstra and Henseler's (2015b) that the results produced using Cronbach's coefficient alpha (Cronbach, 1951) and composite reliability (Chin, 2010) may not be consistent as both refer to sum scores and not construct scores. In particular, Cronbach's alpha (α) typically underestimates the reliability results when the tau-equivalence is not met or if the sample size is relatively smaller (Sijtsma, 2009; Yuan & Bentler, 2002). Thus, it can be regarded as the lower boundary to the reliability. As for Jöreskog's rho (ρc), it can be an appropriate measure of reliability if the assumption that the parameter estimates are accurate (Chin 1998). However, the assumption of parameter accuracy is unlikely to hold, because the indicator loadings provided by PLS are known to be upward-biased (Dijkstra, 1983). As a result, it is likely to overestimate the actual reliability of construct scores. Thus, if researchers are interested in the reliability of PLS construct scores on a common factor model (reflective), they would have to consider using Dijkstra–Henseler's rho (ρA), which is the only consistent PLS-based reliability estimate in this current state of PLS assessment. It is able to evaluate the probability limits (approximate values obtained for the population) of the construct weights obtained by means of Mode A that are proportional to the true loadings (Dijkstra, 2010).

In terms of discriminant validity using the HTMT criterion and lateral collinearity, there is no major difference observed in results regardless of using either the traditional PLS (Mode A) or PLSc (Mode A consistent). Interestingly, a major highlight of this study is the results of coefficients of determination (R^2). PLSc is found to be able to explain more than half of the variation in the endogenous variable while the majority of variance remains unexplained in the traditional PLS estimation. This result is in line with Djisktra and Henseler’s (2015b) study that the R^2 is indeed more explanatory oriented in PLSc than that of the traditional PLS. This further helps clarify Hair et al.’s (2016) claim that the use of goodness-of-fit may not be a harmful guidance for researchers because achieving better “fit” does not compromise the predictive power.

In addition to the difference in R^2, the different conclusions drawn from path-coefficient results (structural model stage) is another major issue when comparing the traditional PLS with PLSc. The results indicate that the traditional PLS tends to overestimate the relationship parameters, which would most likely lead to Type I error. For instance, the relationship between Brand Involvement and Brand Equity is found significant only for the traditional PLS estimation. This again corresponds to the same concern raised by Djikstra and Henseler (2015b) regarding the inflated Type 1 and Type II errors when estimating the VB-SEM. This matter is particularly highlighted as inconsistent estimates that may lead to wrongful hypothesis testing and erroneous conclusions. Notably, researchers should not hold the belief that an inconsistency of estimates is unproblematic when they are only interested in the existence of prediction effect in business studies administered through survey. Therefore, for hypothesis testing, the present study reinforces the assertion to consider using PLSc when common factor model (reflective) is appropriated.

Apart from considering the statistical significance (p-value), the findings of substantive significance (f^2) also points out that PLSc estimation is able to produce more consistent results than that of the traditional PLS, specifically, when a path coefficient is significant, the f^2 result from PLSc would show a disposition to produce higher estimation. Similarly, when the path coefficients are insignificant, the f^2 results from PLSc would produce lower estimation results. In such circumstances, the accurate estimation from PLSc would better serve as a guide for researchers than traditional PLS to make better decisions of whether an exogenous construct has a substantive impact on the endogenous construct (Cohen, 1988).

Lastly, the results show that the traditional PLS (Mode A) performs better in hold-out sample than training sample. It is because when the composite model holds in the data nature, marketing researchers can take this opportunity to look into prediction type of study (research goals) instead of looking into causal relationships. In contrast, the result of PLSc (Mode A consistent) performs better in training sample than hold-out sample and it demonstrates more consistent explanatory relevance estimation compared to the traditional PLS because the R^2 value comparison between training and hold-out sample data is smaller for PLSc. This is consistent with Hahn and Ang's (2017) contention that both R^2 values (training and hold-out) estimation must be fairly similar. A relatively smaller difference means that both estimations have explanatory relevance in practice and it is suitable for confirmatory theory testing and model comparison research goals, if the common factor model holds in research.

In general, the present study complements and extends prior research on examining PLS's performance on common factor model data (Djikstra & Henseler, 2015a, b). Findings from the present study suggest that PLSc estimations have the potential to play a greater role in future SEM applications when estimating common factor models (true reflective model). As a concluding note, this paper would further propose the conditions when PLSc and traditional PLS are more appropriate to use in marketing studies. Figure 3 exhibits the guideline that could help researchers to practice suitable estimation of PLSc and traditional PLS when reflective model are used in the study (i.e., brand involvement, brand interactivity, brand equity, and purchase intention). *Firstly*, researchers must consider the nature of the measurement model (reflective or formative), which expresses how to measure the construct by means of a set of indicators (Jarvis et al., 2003). This can be done by considering the conceptualization or operationalization of the construct. Since a reflective measurement model dictates that all items reflect the same construct, indicators associated with a construct should be highly correlated with each other (Edwards & Bagozzi, 2000). Also, individual items should be interchangeable, and any single item can generally be left out without changing the meaning of the construct, if the construct has sufficient reliability (Jarvis et al., 2003). The fact that the relationship goes from the construct to its indicators, it implies that if the evaluation of the latent trait changes (e.g., because of a change in the standard of comparison), all indicators will change simultaneously (e.g., Diamantopoulos & Winklhofer, 2001). In addition, researcher must be aware that the indicators are error-prone manifestations of an underlying construct with relationships going from the construct to its indicators (Bollen, 1989).

If all these assumptions are fulfilled, researchers can then assess the saturated model fit of the reflective model. If the results indicate SRMR to be less than 0.08 (Hu & Bentler, 1998) and the test of exact model fit is not significant (d_ULS and d_G must < 95% or < 99% bootstrap quantile) (Dijkstra & Henseler. 2015a), then researchers should opt for the estimation of consistent PLS that focuses on common factor modeling. If the criteria of model fit do not achieve satisfactory result, there are two alternatives solution researchers should consider;

i. Firstly, if the measurement model fit is deemed low or not established, researchers should first look into the residual correlation matrix as a diagnostic or guide to examine the "wellness" of research model. This residual correlation matrix is the discrepancy (difference) between the empirical correlation matrix and the model-implied correlation matrix (Henseler, 2017b). The result of this residuals value can be either positive or negative, depending on whether the model-implied correlation is under or over corresponding empirical correlation matrix. According to Hair et al. (2010), if the result of the residuals is less than 2.5, it exhibits a problem but if it is greater than 4.0 suggests a potentially unacceptable degree of error that does not contribute to estimating the respective of factor model. Hence it may call for the deletion of the offending item. Note, the smaller the result of the residuals is, the better the fit is. Researchers should be aware that some large residuals may occur due to sampling error. They can still accept one or two of these large residuals in many instances. What is of concern is a consistent large pattern of large residuals, associated either with a single variable and a number of other variables or residuals for several variables within a construct (Hair et al. 2010). In addition, residuals between 2.5 and 4.0 deserve some attention, but may not suggest any changes to the model if no other problems are associated with those items. Therefore, researchers could employ the residual estimates as a diagnostic to drop one of the items associated with a residual greater than 2.5 and 4.0.

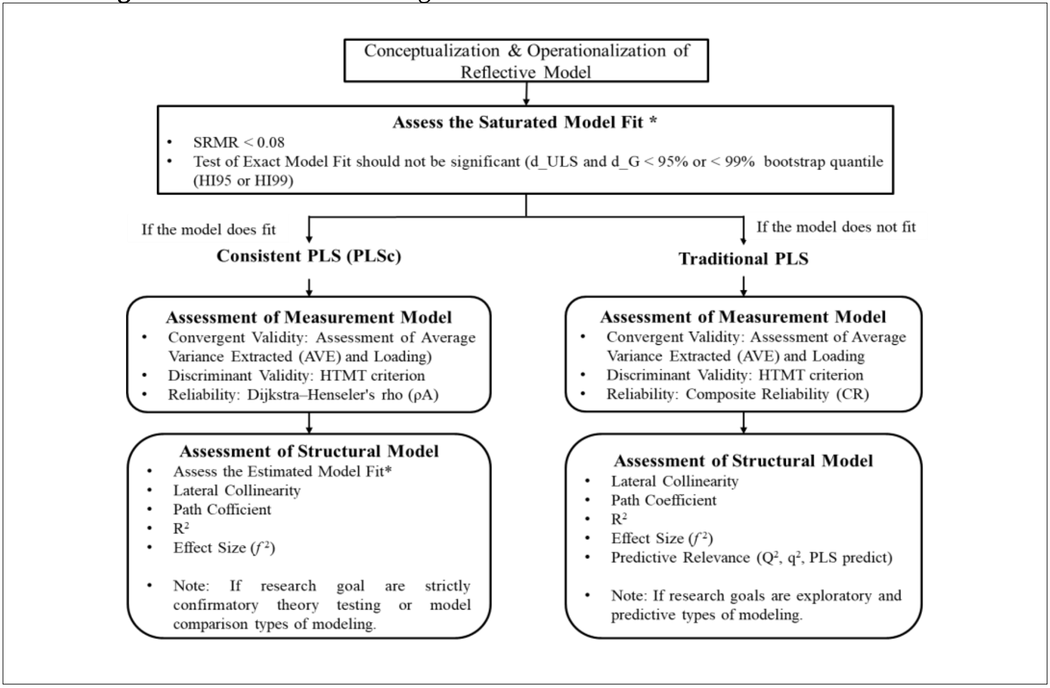
ii. Secondly, researchers should opt for the estimation of traditional PLS that focus on composite modeling. The reason of “misfit” does not necessarily mean that the researcher’s model is incorrect rather it is an indication that more information can be extracted from the data (Jöreskog, 1969, p. 201). Subsequently, when the result indicates poor model fit, one might expect the data's nature to be composite-based data population which is likely to occur in applied research studies as compared to common-based data population (Rigdon 2016; Schonemann & Wang, 1972). Evidently, Sarstedt et al. (2016) findings shows that the traditional PLS is optimal for estimating composite-based population in reflective model as compare to PLSc in a situation when the data of the nature shown a composite-based data population. In other words, the result produced by PLSc will have strong biases across composite-based population even with the increase of sample size (Sarstedt et al. 2016). Thus, one should make use of traditional PLS to measure reflective measurement as the preferred SEM method to avoid any bias estimation based on the unknown nature of the data (Sarstedt et al. 2016). Notably, when the purpose is to estimate data from common factor populations, PLSc performs well in the situation of having more indicators ($\geq$ 4 indicators) in a reflective measurement model and also having larger sample sizes ($n \geq 250$) in a study, to avoid peak result in terms of coefficients' mean absolute error (Sarstedt et al., 2016).

Next, the researchers should proceed to the assessment of measurement model. The criteria of assessing the convergent validity and discriminant validity between PLSc and traditional PLS are the same. Note, the Fornell and Larcker’s (1981) criterion on assessing discriminant validity still able to perform suboptimal in marketing study (Voorhees et al., 2016), if researcher opt to use the PLSc estimation. The disattenuation correction ρA from the PLSc estimation is able to generate consistent loadings and AVE, which these results are able to reduce the bias of measurement error for reflective model when assessing the discriminant validity criterion. However, in terms of reliability assessment, Dijkstra–Henseler's rho (ρA) should be reported when PLSc is used, whereas composite reliability is to be reported in traditional PLS estimation.

Finally, in the assessment of structural model, if the researchers use PLSc for estimation, they need to assess the estimated model fit, lateral collinearity, path coefficient, R^2 and effect size. Importantly, these assessment criteria are suitable to claim confirmatory theory testing and model comparison with the use of model fit criteria. In contrast, if traditional PLS is used for estimation, they need to assess the lateral collinearity, path coefficient, R^2, effect size and the predictive relevance (Q^2, q^2, and PLS predict). Notably, these assessment criteria are suitable to claim exploratory and predictive modeling.

Arguably the understanding of this approach and its purpose are still at the developmental stage; however, it is an important groundwork to open up potential and scholarly discussions on the subject matter to provide further consideration in social science methodology and analysis. Therefore, future research is required to broaden the knowledge of the relative comparison estimation using a real data scenario in marketing research so as to further perpetuate and articulate the assertions made by Dijkstra and Henseler (2015a, b). For example, the estimation of the traditional PLS and PLSc can be compared with a broader range of model constellations and more complex model structures, such

Figure 3: Guideline on the usage of PLSc and traditional PLS for reflective model



Note: * refer to the same threshold value of model fit criteria of SRMR and Test of Exact Model Fit

as hierarchical latent variable model when using Reflective-Composite Type or Reflective-Reflective Type (Ringle et al., 2012), mediation analysis (Nitzl et al., 2016), moderation analysis (Henseler & Chin 2010), multi-group Analysis (Henseler et al., 2016b; Sarstedt et al. 2011) and analysis with nonlinear effects (Djikstra & Henseler 2011). Such assessments would help make known to the researchers the efficacy of different methods for different situations in marketing research. Future investigations may also consider conducting studies to compare estimation of the traditional PLS and PLSc with different number of observations and mixed types of loading results as well as different number of indicators using real data scenario. Such a design will be more likely provide supplementary insights into the comparison between the traditional PLS and PLSc, which may in turn help clarify the estimation modes' efficacy without any parameter bias under different model specifications.

REFERENCES

Babin, B.J., Griffin, M. and Hair, J.F. (2016). Heresies and sacred cows in scholarly marketing publications. *Journal of Business Research*, *69*(8), 3133-3138.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74-94.

Becker, J.-M., Rai, A., & Rigdon, E. E. (2013). Predictive validity and formative measurement in structural equation modeling: Embracing practical relevance. *Proceedings of the International conference on information systems*, Milan.

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

Bollen, K.A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research*, *17*(3), 303-316.

Bush A. J., & Hair Jr, J. F. (1985). An assessment of the mall intercept as a data collection method. *Journal of Marketing Research*, *22*(2), 158-167.

Chang S. J., Van Witteloostuijn A., Eden, L. (2010). From the editors: Common method variance in international business research. *Journal of International Business Studies, 41*(2), 178-184.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Eds.), *Methodology for business and management. Modern Methods for Business Research* (pp. 295-336). Mahwah, NJ: Lawrence Erlbaum Associates.

Chin, W. W. (2010). How to write up and report PLS analyses. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields* (pp. 655–690). Berlin: Springer.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah: Lawrence Erlbaum Associates.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334.

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

Diamantopoulos, A., Siguaw, J. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management, 17*(4), 263-282.

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

Dijkstra T. K., & Henseler, J. (2015a). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics & Data Analysis, 81*, 10-23.

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

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

Dijkstra, T. K. (2010). Latent variables and indices: Herman Wold's basic design and partial least squares. *In Handbook of Partial Least Squares* (pp. 23-46). Berlin: Springer Berlin Heidelberg.

Dijkstra, T. K. (2014). PLS' Janus Face. *Long Range Planning*, *47*(3), 146–153.

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

do Valle, P. O., & Assaker, G. (2016). Using partial least squares structural equation modeling in tourism research: a review of past research and recommendations for future applications. *Journal of Travel Research, 55*(6), 695-708.

Ebbes, P., Papies, D., & VanHeerde, H. J. (2011). The sense and non-sense of holdout sample validation in the presence of endogeneity. *Marketing Science, 30*(6), 1115-1122.

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

Evermann, J., & Tate, M. (2016). Assessing the predictive performance of structural equation model estimators. *Journal of Business Research, 69*(10), 4565-4582.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191.

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

Gold, A. H., Malhotra, A., & Segars, A. H. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems, 18*(1), 185-214.

Goodhue, D. L., Lewis, W., & Thompson, R. (2012). Comparing PLS to regression and LISREL: A response to Marcoulides, Chin, and Saunders. *MIS Quarterly, 36*(3), 703-716.

Goodhue, D., Lewis, W., Thompson, R. (2006). PLS, small sample size, and statistical power in MIS research. *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)* 8:202b-202b. IEEE

Hahn, E. D., & Ang, S. H. (2017). From the editors: New directions in the reporting of statistical results in the Journal of World Business. *Journal of World Business*, *52*(2), 125-126.

Hair J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing theory and Practice*, *19*(2), 139-152.

Hair, J. F., Black, W. C., & Babin, B. J., Anderson, R. E., Tatham, R. (2010). *Multivariate Data Analysis: Pearson Education*. Upper Saddle River, New Jersey.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Editorial-partial least squares structural equation modeling: Rigorous applications, better results and higher acceptance. *Long Range Planning,* 46(1-2), 1-12.

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C.M. (2012b). The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications. *Long Range Planning*, *45*(5), 320-340.

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

Hair, Jr. J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (2nd ed.). Thousand Oaks: Sage.

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

Henseler, J. (2010). On the convergence of the partial least squares path modeling algorithm. *Computational Statistics*, *25*(1), 107-120.

Henseler, J. (2017a). Bridging design and behavioral research with variance-based structural equation modeling. *Journal of Advertising*, *46*(1), 178-192.

Henseler, J. (2017b). ADANCO 2.0.1: User Manual. Kleve: Composite Modeling GmbH & Co.

Henseler, J., & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural Equation Modeling, 17*(1), 82-109.

Henseler, J., & Dijkstra, T. K. (2015). ADANCO 2.0. Composite Modeling, Kleve: Composite Modeling GmbH & Co.

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

Henseler, J., Hubona, G., & Ray, P. A. (2016a). Using PLS path modeling in new technology research: updated guidelines. *Industrial Management & Data Systems, 116*(1), 2-20.

Henseler, J., Ringle C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing*, *20*(1), 277-319.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015) A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115-135.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: *A Multidisciplinary Journal*, *6*(1), 1-55.

Huang, H. M. (2006). Do print and Web surveys provide the same results? *Computers in Human Behavior*, *22*(3), 334-350.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, *30*(2), 199-218.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34*(2), 183-202.

Jöreskog, K. G., & Wold, H. O. (1982). *Systems under indirect observation: Causality, structure, prediction.* Amsterdam: North-Holland.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: Guilford Press.

Labrecque, L. I. (2014). Fostering consumer–brand relationships in social media environments: The role of parasocial interaction. *Journal of Interactive Marketing, 28*(2), 134-148

Lee, L., Petter, S., Fayard, D., & Robinson, S. (2011). On the use of partial least squares path modeling in accounting research. *International Journal of Accounting Information Systems*, *12*(4), 305-328.

Liang, T. P., & Lai, H. J. (2002). Effect of store design on consumer purchases: an empirical study of on-line bookstores. *Information & Management*, *39*(6), 431-444.

Lin, W., & Van Ryzin, G. G. (2012). Web and mail surveys: An experimental comparison of methods for nonprofit research. *Nonprofit and Voluntary Sector Quarterly*, *41*(6), 1014-1028.

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

Lu, I. R., Kwan, E., Thomas, D. R., & Cedzynski, M. (2011). Two new methods for estimating structural equation models: An illustration and a comparison with two established methods. *International Journal of Research in Marketing*, *28*(3), 258-268.

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

Memon, M. A., Ting, H., Ramayah, T., Chuah, F., Cheah, J-H. (2017). A Review of the Methodological Misconceptions and Guidelines Related to the Application of Structural Equation Modeling: A Malaysian Scenario. *Journal of Applied Structural Equation Modeling*, *1*(1), i-xiii.

Merrilees, B., & Fry, M. L. (2003). E-trust: the influence of perceived interactivity on e-retailing users. *Marketing Intelligence & Planning*, *21*(2), 123-128.

Nitzl, C., Roldán, J. L., & Cepeda, G. (2016). Mediation Analyses in Partial Least Squares Structural Equation Modeling: Helping Researchers to Discuss More Sophisticated Models. *Industrial Management & Data Systems*, *116*(9), 1849-1864.

Nunnally, J. C., & Bernstein, I. H. (1994). The assessment of reliability. *Psychometric Theory, 3*(1), 248-292.

Pakpahan, E., Hoffmann, R., & Kröger, H. (2017). Statistical methods for causal analysis in life course research: an illustration of a cross-lagged structural equation model, a latent growth model, and an autoregressive latent trajectories model. *International Journal of Social Research Methodology, 20*(1), 1-19.

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

Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, *12*(4), 531-544.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879-903.

Ramayah, T., Cheah, J., Chuah, F., Ting, H., & Memon, M. A. (2018). *Partial least squares structural equation modeling (PLS-SEM) using SmartPLS 3.0: An updated and practical guide to statistical analysis.* (2nd edition) Singapore: Pearson.

Reinartz, W., Haenlein, M., & Henseler., J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of research in Marketing*, *26*(4), 332-344.

Rigdon, E. E. (2014). Rethinking partial least squares path modeling: breaking chains and forging ahead. *Long Range Planning*, *47*(3), 161-167.

Rigdon, E. E. (2016). Choosing PLS path modeling as analytical method in European management research: A realist perspective. *European Management Journal*, *34*(6), 598-605.

Ringle, C. M., Sarstedt, M., & Straub, D. (2012) A critical look at the use of PLS-SEM in MIS Quarterly. *MIS Quarterly, 36*(1), iii-xiv.

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CB SEM: Where the bias lies! *Journal of Business Research, 69*(10), 3998-4010.

Sarstedt, M., Henseler, J., & Ringle, C. M. (2011). Multigroup Analysis in Partial Least Squares (PLS) Path Modeling: Alternative Methods and Empirical Results. Measurement and Research methods in International Marketing, *Advances in International Marketing*, *22*, 195-218.

Sarstedt, M., Ringle, C. M., Henseler, J., & Hair, J. F. (2014a). On the emancipation of PLS-SEM: A commentary on Rigdon (2012). *Long range planning*, *47*(3), 154-160

Sarstedt, M., Ringle, C. M., Smith, D., Reams, R., & Hair, J. F. (2014b). Partial least squares structural equation modeling (PLS-SEM): A useful tool for family business researchers. *Journal of Family Business Strategy, 5*(1), 105-115.

Schönemann, P. H., & Wang, M. M. (1972). Some new results on factor indeterminacy. *Psychometrika*, *37*(1), 61-91.

Schorfheide, F., & Wolpin, K. I. (2012). On the use of holdout samples for model selection. *The American Economic Review*, *102*(3), 477-481.

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

Urbach, N., & Ahlemann, F. (2010). Structural equation modeling in information systems research using partial least squares. *Journal of Information Technology Theory and Application*, *11*(2), 5-40.

van Riel, A.C., Henseler, J., Kemény, I., & Sasovova, Z., (2017). Estimating hierarchical constructs using consistent partial least squares: The case of second-order composites of common factors. *Industrial Management & Data Systems*, *117*(3), 459-477.

Verhoef, P. C., Langerak, F., & Donkers, B. (2007). Understanding brand and dealer retention in the new car market: The moderating role of brand tier. *Journal of Retailing*, *83*(1), 97-113.

Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science, 44*(1), 119-134.

Willaby, H. W., Costa, D. S., Burns, B. D., MacCann, C., & Roberts, R. D. (2015). Testing complex models with small sample sizes: A historical overview and empirical demonstration of what partial least squares (PLS) can offer differential psychology. *Personality and Individual Differences*, *84*, 73-78.

Wold, H. (1982). Soft modelling: the basic design and some extensions. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation* (pp. 36-37). Amsterdam: North-Holland.

Woodside, A. G. (2013). Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. *Journal of Business Research*, *66*(4), 463-472.

Yuan, K. H., & Bentler, P. M. (2002). On robustness of the normal-theory based asymptotic distributions of three reliability coefficient estimates. *Psychometrika, 67*(2), 251-259.