The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/0263-5577.htm

Demystifying the role of causal-predictive modeling using partial least squares structural equation modeling in information systems research

Received 7 October 2019
Revised 18 February 2020
18 May 2020
Accepted 15 June 2020

Wynne Chin
Department of Decision and Information Sciences, C.T. Bauer College of Business, University of Houston, Houston, Texas, USA
Jun-Hwa Cheah
School of Business and Economics, Universiti Putra Malaysia, Serdang, Malaysia
Yide Liu
Macau University of Science and Technology, Taipa, China
Hiram Ting
Faculty of Hospitality and Tourism Management, UCSI University, Kuala Lumpur, Malaysia
Xin-Jean Lim
School of Business and Economics, Universiti Putra Malaysia, Serdang, Malaysia, and
Tat Huei Cham
Faculty of Accountancy and Management, Universiti Tunku Abdul Rahman, Kajang, Malaysia

Abstract

Purpose – Partial least squares structural equation modeling (PLS-SEM) has become popular in the information systems (IS) field for modeling structural relationships between latent variables as measured by manifest variables. However, while researchers using PLS-SEM routinely stress the causal-predictive nature of their analyses, the model evaluation assessment relies exclusively on criteria designed to assess the path model's explanatory power. To take full advantage of the purpose of causal prediction in PLS-SEM, it is imperative for researchers to comprehend the efficacy of various quality criteria, such as traditional PLS-SEM criteria, model fit, PLSpredict, cross-validated predictive ability test (CVPAT) and model selection criteria.

Design/methodology/approach – A systematic review was conducted to understand empirical studies employing the use of the causal prediction criteria available for PLS-SEM in the database of Industrial Management and Data Systems (IMDS) and Management Information Systems Quarterly (MISQ). Furthermore, this study discusses the details of each of the procedures for the causal prediction criteria available for PLS-SEM, as well as how these criteria should be interpreted. While the focus of the paper is on demystifying the role of causal prediction modeling in PLS-SEM, the overarching aim is to compare the performance of different quality criteria and to select the appropriate causal-predictive model from a cohort of competing models in the IS field.

Findings – The study found that the traditional PLS-SEM criteria (goodness of fit (GoF) by Tenenhaus, R2 and Q2) and model fit have difficulty determining the appropriate causal-predictive model. In contrast, PLSpredict, CVPAT and model selection criteria (i.e. Bayesian information criterion (BIC), BIC weight, Geweke–Meese criterion (GM), GM weight, HQ and HQC) were found to outperform the traditional criteria in determining the

The authors would like to thank Marko Sarstedt and Benjamin Dybro Liengaard for guiding us the new PLS prediction techniques (i.e., information theoretic model selection criteria and cross-validated predictive ability test) in this manuscript.

Industrial Management & Data Systems
Vol. 120 No. 12, 2020
pp. 2161-2209

0263-5577
DOI 10.1108/IMDS-10-2019-0529

appropriate causal-predictive model, because these criteria provided both in-sample and out-of-sample predictions in PLS-SEM.

Originality/value – This research substantiates the use of the PLSpredict, CVPAT and the model selection criteria (i.e. BIC, BIC weight, GM, GM weight, HQ and HQC). It provides IS researchers and practitioners with the knowledge they need to properly assess, report on and interpret PLS-SEM results when the goal is only causal prediction, thereby contributing to safeguarding the goal of using PLS-SEM in IS studies.

Keywords Causal prediction, Traditional PLS criteria, Model fit, PLSpredict, CVPAT, Model selection criteria

Paper type Research paper

1. Introduction

As partial least squares structural equation modeling (PLS-SEM) progresses, it is becoming an increasingly visible approach for theory testing in a plethora of academic disciplines (Khan *et al.*, 2019; Hwang *et al.*, 2020), such as accounting (Nitzl *et al.*, 2016; Lee *et al.*, 2011), international management (Richter *et al.*, 2016), operations management (Peng and Lai, 2012), information systems (IS) (Shiau *et al.*, 2019; Hair *et al.*, 2017a) and marketing (Hair *et al.*, 2012). Wold (1982), the pioneer of the method, described PLS-SEM as a method that combines both the econometric prediction and the psychometric modeling of latent variables. Specifically, the main objective of PLS-SEM is to predict and explain a key target construct and/or to identify its relevant antecedent constructs. In other words, this approach generates latent variable scores that maximize within-sample prediction in terms of the dependent latent variable's R^2 value. As such, the estimated coefficients depict the relevance of constructs in a certain model that directly, indirectly and totally contribute to the explanation of a target construct of interest. As highlighted by Wold (1985), the latent variable scores allow prediction of the indicator values of the dependent constructs, thus determining the predictive relevance of the model.

Although PLS-SEM is typically presented as an alternative technique to covariance-based structural equation modeling (CB-SEM) (Jöreskog, 1973; Rigdon *et al.*, 2017), many researchers routinely muddle the dichotomy between explanation and prediction in their research goals. This dichotomy is often not useful, because the measurable data are often not an accurate representation of their underlying constructs. Likewise, a similar concern was noted in Shmueli (2010) study, that is, the operationalization of theories and constructs into statistical models and measurable data "creates a disparity between the ability to explain phenomena at the conceptual level and the ability to generate predictions at the measurement level" (Shmueli, 2010, p. 293). After research models are established, "the "wrong" model can sometimes predict better than the correct one" (Shmueli, 2010, p. 293), while the best predictive model often does not provide much explanation. Hence, it depends on the goal of the research (see Figure 1), as this guides the way the method is applied. For example, if the goal concerns prediction, then the researchers should look into applying machine learning forecasting methods (i.e. random forests and artificial neural networks), where prediction is critical but theoretical consistency [1] may be of secondary concern (Shmueli and Koppius, 2011; Shmueli *et al.*, 2016). Alternatively, if the study focuses on confirmatory/explanatory modeling, researchers should consider either CB-SEM or the newly proposed consistent PLS

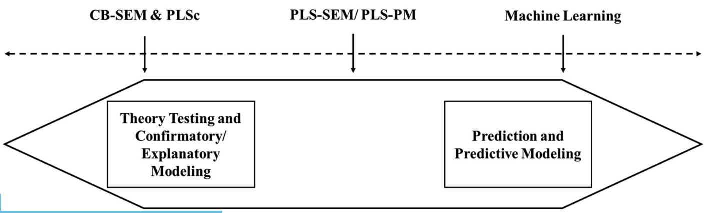


Figure 1.
PLS-SEM as method for confirmatory/ explanatory and predictive modeling

approach (Bentler and Huang, 2014; Dijkstra and Henseler, 2015a, 2015b). Both of these techniques have their own fit measures—usually oriented to in-sample measures—for assessing the explanatory power and specification of the model.

Nevertheless, the goal of a study is usually not restricted to one pole of the characterized continuum but takes a position near the midpoint, which is the basis for developing theoretical and managerial implications (Hair and Sarstedt, 2019). Explanation and prediction are two distinct concepts of statistical modeling and estimation. In particular, "explanatory modeling focuses on minimizing bias to obtain the most accurate representation of the underlying theory, which is grounded in well-developed causal explanations" (Evermann and Tate, 2016; Liengaard *et al.*, 2020). Predictive modeling seeks to minimize the combination of bias and estimation variance, occasionally sacrificing theoretical accuracy for improved empirical precision (Shmueli, 2010, p. 293). Correspondingly, a grossly mis-specified model can yield superior predictions, whereas a correctly specified model can perform extremely poorly in terms of prediction. This perspective corresponds to the work of Jöreskog and Wold (1982, p. 270), who designed PLS-SEM as a "causal-predictive" technique [2] that overcomes the dichotomy between explanatory and predictive modeling. In other words, the use of PLS-SEM as a causal-predictive technique often means the model is expected to exhibit high predictive accuracy, while also being grounded in well-developed causal explanations (see Figure 1). Gregor (2006, p. 626) refers to this interplay as explanation and prediction theory, which "implies both understanding of underlying causes and prediction, as well as description of theoretical constructs and the relationships among them." According to this assumption, researchers who use PLS-SEM may consider both sets of evaluation criteria—for both explanatory modeling and predictive modeling—to different degrees. A successful union between explanation and prediction lends authority to our understanding of the system that relates to the fundamental quest of science (Dublin, 1969).

In a recent work, Hair *et al.* (2017b) highlighted that the latest PLS-SEM toolbox included a broad range of evaluation criteria for assessing the adequacy of a model. However, many researchers still fail to fully utilize the suitability evaluation criteria metrics in PLS-SEM to corroborate the research goal of causal prediction. For example, most of the studies were found to interpret an in-sample prediction of the dependent construct by using model estimates that predicted the case values of the entire sample, such as the coefficient of determination (also known as the R^2) or adjusted R^2 (Hair *et al.*, 2017b; Sarstedt *et al.*, 2013) [3]. In addition, some researchers still chose to report the model's out-of-sample predictive performance, such as the Q^2 or q^2 values that were generated through blindfolding procedures (Geisser, 1974; Stone, 1974). However, Shmueli *et al.* (2016) stressed that the blindfolding procedure has several limitations; hence, they introduced the PLSpredict procedure to assess the PLS path model's predictive quality via true out-of-sample metrics, such as the root mean square error (RMSE) and the mean absolute error (MAE). Along with this advancement, Liengaard *et al.* (2020) also proposed another type of prediction procedure that could be assessed in PLS-SEM, called the cross-validated predictive ability test (CVPAT), which offers an overall inferential test for predictive model comparison. There were also researchers who explored whether the use of in-sample measures, such as the model selection criteria (i.e. Akaike information criterion (AIC), Bayesian information criterion (BIC), Geweke–Meese criterion (GM)) could be a potential substitute for out-of-sample criteria that require a holdout sample (Sharma *et al.*, 2019a, 2019b). That same year, Danks *et al.* (2020) extended the use of the model selection criteria by looking into the AIC weights (AICw), BIC weights (BICw) and GM weights (GMw). These new criteria were intended to assist scholars in overcoming selection uncertainty when selecting an appropriate model over others alternative model based on the model selection criteria [4]. Alternatively, there are other groups of scholars who support the evaluation criteria metrics of goodness of fit (GoF) by Tenenhaus *et al.* (2005). Some now even consider using the model fit criteria in PLS-SEM (i.e.

standardized root mean square residual (SRMR)) [5] for explanatory purposes (Dijkstra and Henseler, 2015a, 2015b; Henseler *et al.*, 2014).

The broad range of evaluation metrics in PLS-SEM makes an evaluation of the causal prediction of a model problematic, because an optimal model for prediction purposes may differ from one obtained in an explanatory modeling context. In other words, a well-fitting model designed in an explanatory context may perform very poorly in terms of out-of-sample prediction (Shmueli, 2010). Therefore, accomplishing highly satisfactory results in both directions can be difficult and complex, especially with respect to the IS field, which has many parts and many possible arrangements with intricate underlying causal interactions and relationships (Sharma *et al.*, 2019a, 2019b).

Grasping this trade-off between explanation and prediction requires researchers to have a sound understanding of each evaluation metric in PLS-SEM. This is important because methodological research in PLS-SEM is quite dynamic in terms of evaluating the method's characteristics (e.g. Henseler *et al.*, 2014; Rigdon, 2012, 2014; 2016; Rönkkö and Evermann, 2013; Sarstedt *et al.*, 2016), hence understanding each evaluation criteria metric allows researchers to further extend their existing theories, combine different theoretical models and compare theories. Focusing on such capabilities also allows researchers to formulate parsimonious models with relatively high predictive relevance to empirically validate their *a priori* hypothesized predictive relationships, and to substantiate and generalize the relevant effects and the theoretical model across time and different datasets. In the IS field, most research goals are to uncover both the significant and the relevant effects that allow researchers to determine the key success factors that the information management and data system field should focus on. Because the use of PLS-SEM in research is increasing, examining its usage is critical to counteracting misapplication, which could otherwise be reinforced over time.

Given the growing concerns about causal prediction practice in PLS-SEM, we first conducted a systematic review using the Emerald database to search for journals that publish in both Industrial Management and Data Systems (IMDS) and Management Information Systems Quarterly (MISQ) under the subject area of causal prediction. Next, we reviewed all the evaluation metrics in PLS-SEM as an integral element of model assessment, with the aim of providing important guidance on the interplay between theory testing and prediction testing (e.g. Shmueli, 2010) and, if necessary, finding opportunities for realignment in future applications. We also illustrated the procedures for each of the metrics using Zhang *et al.* (2018) model of the impact of channel integration on consumer responses in omni-channel retailing. Through this study, we provide recommendations on how a report should be conducted in research that focuses on explanation or prediction, respectively. While our focus is on each evaluation metric procedure, our overarching aim is to encourage the appropriate assessment of the causal prediction goal in PLS-SEM analyses.

1.1 Systematic review of causal prediction

The systematic review of causal prediction was conducted using the Emerald search engine by focusing on two top-tier journals in the IS field (i.e. IMDS and MISQ). The search was focused on identifying articles in the journal's database that adopted a quality criteria assessment when using PLS-SEM. The search included articles published in the last two decades, from 1999 to 2018. With these search queries, we aimed to extract all the academic articles that were published under the keyword of PLS-SEM. The Preferred Reporting Items for Systematic Reviews (PRISMA) flowchart was used to present the search outcome (see Figure 2).

The search resulted in 91 and 40 hits from the IMDS and MISQ databases, respectively, in the first stage of the systematic review process. The review was restricted to research articles published in English, while non-academic publications (i.e. conference papers, book chapters and book series) were excluded. Overall, only one article was excluded based on these criteria.

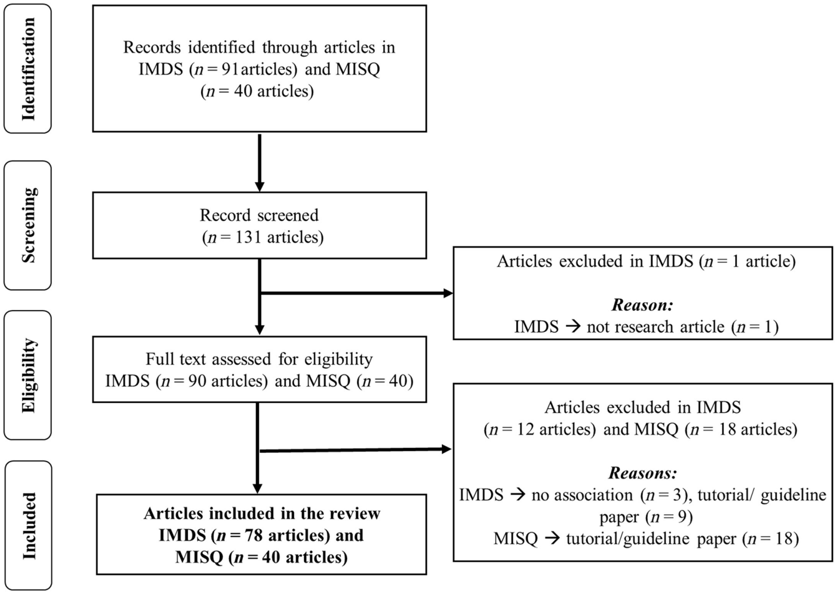


Figure 2.
PRISMA flow chart for PLS-SEM in IMDS

Subsequently, the titles, abstracts and contents of the remaining 130 articles (i.e. IMDS = 90 and MISQ = 40) were examined thoroughly to ensure that they fulfilled the inclusion criteria for the present study. Consequently, nine tutorials/guidelines and three irrelevant articles from the IMDS were excluded, while 18 tutorial/guideline articles that were published in the MISQ were excluded. Details of the 78 articles used in our analysis (see Figure 1) are presented in Tables A1 and A2.

Based on the review of the IMDS articles (see Table A1), we observed that the coefficient of determination (R^2) was used as the main criterion for evaluating research model suitability (100.00%). Only 17.95% of the articles published between 2014 and 2018 reported the blindfolding procedure (i.e. the Q^2 criterion). Interestingly, the usage of Tenenhaus *et al.* (2005) GoF was quite high (12.82%) between 2010 and 2018, despite its deficiency. As for the SRMR, it was used in 14.10% of the studies published between 2016 (when the method was first introduced in the journal IMDS) and 2018. In addition, bootstrap results of the exact model fit (i.e. d_ULS, the squared Euclidean distance and d_G, the geodesic distance) were rarely used between 2016 and 2017 (5.13%). Finally, although PLSpredict was introduced by Shmueli *et al.* (2016) at the same time as the model fit criteria (i.e. SRMR, d_G and d_ULS), it was not adopted in any study between 2016 and 2018. Similarly, no study used the model selection criteria (i.e. AIC or BIC) (Sharma *et al.*, 2019a, 2019b) or the CVPAT (Liengaard *et al.*, 2020) to determine the suitability of a research model. However, surprisingly, R^2 was the only criterion that was reported in MISQ (100.00%) (refer to Table A2).

Drawing from this review, there is a need to understand the efficacy of each criterion for determining the appropriateness of a model. Some possible reasons that the PLSpredict, the CVPAT and the model selection criteria were not used in most past studies could be that researchers were either unaware of these new and suitable criteria that could bridge the gap between the explanation and prediction goals in PLS-SEM, or they were comfortable with the existing reporting criteria. Alternatively, it could be that a plethora of researchers are interested in explanatory modeling research rather than prediction-oriented modeling research. Therefore, many reviewers or editors may be uncomfortable with the absence of a global index assessment; hence, they ask for the fit measures.

2. Current PLS-SEM evaluation toolkit

2.1 PLS-SEM criteria for in-sample prediction

In the recent PLS-SEM literature, there are a growing number of criteria that allow achievement of the causal-predictive goal in IS research. The simplest and most widely adopted criterion is the R^2. R^2 is typically used as a criterion of predictive power (Hair *et al.*, 2017b; Sarstedt *et al.*, 2013), which indicates the variance explained in each of the endogenous constructs. The higher the R^2 value, the greater the predictive accuracy. As a general guideline for marketing studies, the R^2 values of 0.75, 0.50 and 0.25 can be considered substantial, moderate and weak, respectively (Hair *et al.*, 2011, p. 147). Other scholars have also indicated a different interpretation of R^2 values. In particular, Falk and Miller (1992) recommended that R^2 should be at least greater than 0.10, whereas Chin (1998b) considered R^2 values of 0.67, 0.33 and 0.19 as substantial, moderate and weak, respectively (Roldán and Sánchez-Franco, 2012, p. 205). Despite the guideline, researchers have always been advised to interpret R^2 values according to the context of the related discipline. For example, when measuring a concept that is inherently predictable, such as the technology acceptance model, R^2 values of 0.35–0.51 may be considered strong (see Lee *et al.*, 2003). Focusing solely on the R^2 value to assess the adequacy of a theoretical model is problematic because it may cause researchers to overfit their model to the point that it overly accommodates both the information and the idiosyncratic noise in the data (Hair *et al.*, 2019a, b, c, d) [6]. This limits the prospect of the generalizability of the model results to other samples (Pitt and Myung, 2002; Sharma *et al.*, 2019a, 2019b), which is central to empirical research. In other words, the same model would likely not fit if it is used on another sample drawn from the same population (Sharma *et al.*, 2019a, 2019b). Hence, given that R^2 increases as predictors are added to the model, resulting in a more complex model, researchers have widely used the adjusted R^2, which attempts to correct for model complexity by including a penalty proportional to the number of predictors in the model. However, the adjusted R^2 is deemed to lack formal justification and is not suitable for assessing a model's predictive accuracy (Berk, 2008).

2.2 PLS-SEM criteria for out-of-sample prediction

Because R^2 only provides information regarding in-sample prediction, another frequently used metric is the Stone–Geisser's Q^2 (Geisser, 1974; Stone, 1974). Wold recognized the usefulness of this technique by stating that it fits PLS-SEM "like hand in glove" (Wold, 1982, p. 30). Specifically, Q^2 is obtained by means of the blindfolding procedure, which omits a part of the data matrix, estimates the model parameters and predicts the omitted data by using the previously computed estimates. This process is repeated until every data point has been omitted and the model re-estimated. The smaller the difference between the predicted and the original values, the greater the Q^2 value (or $Q^2 \geq 0$), thus ensuring the model's predictive accuracy and relevance (Chin, 1998). Moreover, recent literature by Hair *et al.* (2019b) proposes another rule of thumb: that Q^2 values higher than 0.00, 0.25 and 0.50 depict small, medium and large predictive relevance of the PLS path model, respectively. Chin (2010, p. 680) also indicated that "in general, a cross-validated redundancy Q2 above 0.50 is indicative of a predictive model." To initiate the blindfolding procedure, researchers need to determine the sequence of data points to be omitted in each run. Chin (1998) suggested using an omission distance between 5 and 10. For example, an omission distance of 7 implies that every seventh data point of the endogenous construct's indicator is eliminated in a single blindfolding run. Furthermore, there are two approaches to calculating the Q^2 value: cross-validated redundancy and cross-validated communality, the former of which is generally recommended to explore the predictive relevance of the PLS path model focus on cross-validated redundancy (Hair *et al.*, 2017b; Wold, 1982). Analogous to the effect size (f^2), researchers can also analyze the q^2 effect size, which indicates the change in the Q^2 value when a specified exogenous construct is omitted from the model. As a relative measure of predictive relevance, q^2 values of 0.02, 0.15 and 0.35 indicate that an exogenous construct on a certain endogenous construct has small, medium or large predictive relevance, respectively.

However, both Q^2 and q^2 are ad hoc metrics that do not provide highly interpretable results in terms of prediction error magnitude (i.e. no clear cutoffs for model comparison), and furthermore, their imputation steps do not take heterogeneity in prediction errors into account (see Shmueli *et al.*, 2016). In addition, these metrics are not true measures of out-of-sample prediction, as blindfolding does not omit entire observations, but only data points. "Hence, both Q^2 and q^2 values can only be partly considered a measure of out-of-sample prediction, because the sample structure remains largely intact in its computation" (Shmueli *et al.*, 2016). Fundamental to a proper predictive procedure is the ability to predict measurable information for new cases (Shmueli *et al.*, 2016, p. 4553). Addressing this concern, Shmueli *et al.* (2016) developed the PLSpredict procedure for generating holdout sample-based point predictions in PLS path models on an item or construct level. When running PLSpredict, researchers need to make a series of choices. Technically, the default PLSpredict choices include (1) the number of k-fold cross-validations ($k = 10$); (2) the number of repetitions ($r = 10$) (see Witten *et al.*, 2016); and (3) the selection of an adequate prediction statistic to quantify the degree of prediction error (the MAE, the RMSE and the Q^2_predict criterion from the PLSpredict assessment (Hair *et al.*, 2019b; Shmueli *et al.*, 2019), assuming the sample size is large enough).

Importantly, when interpreting PLSpredict results, the focus should be on the model's key endogenous construct, as opposed to the prediction errors for all the endogenous constructs' indicators. After the key target construct has been selected, the Q^2_predict statistic should be evaluated first to verify that the predictions outperform the most naïve benchmark, defined as the indicator means from the analysis sample (Shmueli *et al.*, 2019). Researchers then need to compare the RMSE and MAE values (produced by the PLSpredict method) against a naïve benchmark using a linear regression model (LM) (Danks and Ray, 2018; Hair *et al.*, 2019c; Shmueli *et al.*, 2019). In most cases, researchers should use the RMSE because it assigns a greater weight to larger errors via squaring the errors, which makes it particularly useful when large errors are undesirable (Hair *et al.*, 2019b). Yet, if the prediction error distribution is highly non-symmetric, the MAE is the more appropriate prediction statistic (Shmueli *et al.*, 2019). If the empirical comparison of the predictive power of competing models with the same endogenous dependent variable is of interest, both the RMSE and MAE should be investigated at the composite level. Often researchers use PLSpredict to produce a case-specific prediction. However, if researchers are interested in establishing a set of competing models from an out-of-sample prediction perspective, they can subsequently look into the prediction errors of the latent variable scores that focus on the out-of-sample error statistic (e.g. RMSE or MAE) (see Shmueli *et al.*, 2019). If a model gives a low prediction error statistic for the key target endogenous construct compared with the results from a set of competing models, it has a better chance of being scientifically replicable and explainable and of exhibiting higher predictive abilities.

Although PLSpredict improves prediction assessment capabilities to a certain degree, the method does not offer any insight regarding the overall inferential test to assess whether the alternative model's (AM) predictive capabilities are significantly better than the established model's (EM). As a remedy, Liengaard *et al.* (2020) established the CVPAT method, which is non-parametric. The purpose of this new method is to conduct a pairwise comparison between two theoretically derived models regards their ability to predict the indicators for all the dependent latent variables (regardless whether reflective or formative) simultaneously. In particular, the test facilitates researchers in determining whether the AM has significantly better predictive accuracy than the EM (or vice-versa) at a pre-specified significance level (e.g. $\alpha = 0.05$). To estimate the out-of-sample prediction errors, CVPAT relies on the concept of k-fold cross-validation, which randomly splits the overall dataset into a specific number of folds (e.g. $k = 10$) and iterates through all the folds. In particular, the initial iteration reserves the first fold as an independent holdout set and estimates the model on the remaining observations, which act as training set. Using the training parameter estimates, the output variables of the first fold are predicted by their input variables. The out-of-sample prediction error is the difference between the predicted value of the output variables and their actual values. Subsequently, the procedure is repeated for each fold (e.g. $k = 10$) to generate the out-of-sample prediction errors for each case in the dataset (see Liengaard *et al.*, 2020). Then, the loss in predictive power associated with a given model is measured as the average squared prediction error over all indicators associated with the endogenous variables. This average lost difference is a measure of the difference in the average out-of-sample performance between two competing models (EM vs AM) when the indicators of the endogenous latent variable (Liengaard *et al.*, 2020) are predicted. A higher average loss of the EM compared with the AM implies a higher average prediction error, which exhibits an inferior out-of-sample model performance for EM (vice-versa). In order to complement this result, CVPAT also provides significance test results for both the *p*-value and the confidence interval as crucial evidence in favor of one model in terms of its predictive accuracy compared with the other model [7] (Liengaard *et al.*, 2020).

2.3 PLS-SEM criteria of model selection criteria

As an alternative to both the PLSpredict and CVPAT, researchers can revert to model selection criteria derived from information theory (McQuarrie and Tsai, 1998). These criteria strike a balance between model fit and complexity to prevent overfitting so that the model can generalize beyond the particular sample (Myung, 2000). The two most widely used model selection criteria are the AIC (Akaike, 1973) and the BIC (Schwarz, 1978). The AIC and BIC differ somewhat in their conceptual underpinnings and assumptions. Specifically, the BIC provides an estimate of the posterior probability of a model being true and chooses the model that maximizes this probability on a given dataset. In other words, it strives to select a model that is most likely (in the Bayesian sense) to coincide with the underlying data generating model. In contrast, the AIC is designed to estimate the relative amount of information lost (using the Kullback–Leibler divergence measure between distributions) when a given model estimated from data is compared with a "true" but unknown data generating process (Burnham and Anderson, 2002). In addition, there are several variations of the original AIC and BIC criteria that have also been proposed in recent decades, including Mallow's C_p criterion, the unbiased AIC (AICu), the corrected AIC (AICc), the GM, the Hannan–Quinn Criterion (HQ) and the corrected Hannan–Quinn Criterion (HQC) (see Sharma *et al.*, 2019a, 2019b). These criteria are typically written as a function of the maximum likelihood value, but they can be expressed as a function of the model residuals when the error distribution is normal with constant variance (Burnham and Anderson, 2002, p. 63). In other words, these least squares formulation characteristics (i.e. the AIC, BIC, GM and other criteria) using model residuals are merely a special case of the model selection criteria equivalent to that of the likelihood estimation (see Sharma *et al.*, 2019a). Consequently, it makes these criteria suitable for PLS-SEM estimation as it relies on an iterative estimation of piecewise LMs (e.g. Hair *et al.*, 2017b) [8]. Importantly, Sharma *et al.* (2019a, 2019b) highlighted that the model selection criteria (particularly BIC and GM) are known as in-sample criteria that could be a substitute for out-of-sample criteria that require a holdout sample. Such a substitution is advantageous, especially when the researcher does not have the luxury of a holdout sample (using an insufficient sample for the holdout sample causes considerable loss of statistical and predictive power), and the goal is to select correctly specified models with low prediction error. Subsequently, these model selection criteria help compare different model configurations that could result from different theories or research contexts.

However, one possible issue in the application of model selection criteria (i.e. AIC, BIC and GM) is that—in their simple form (i.e. raw values)—they do not provide any in-depth information regarding the relative weights of evidence in favor of the models under consideration (Burnham and Anderson, 2002; Wagenmakers and Farrell, 2004). Specifically, while the differences in the criteria values are useful for ranking and selecting models, such differences can often be small in practice, resulting in a false sense of model selection confidence and uncertainty (Preacher and Merkle, 2012). For example, when comparing two models with similar BIC values, it is difficult to determine how much statistical importance should be attached to the small difference. To overcome this uncertainty, Danks *et al.* (2020) proposed the use of AICw, BICw and GMw to help researchers comprehend how much a selected model is better than others in a given sample (Symonds and Moussalli, 2011). The AICw, BICw and GMw are all derived from the model selection criteria's raw values to compute the relative likelihood of a model, given the data and set of models (see Wagenmakers and Farrell, 2004). The likelihood values allow a researcher to draw more robust inferences by creating an additional measure that can be used to judge the relative strength of each model in the set. This measure gives researchers more information about whether to base the inference on a single superior model when weighing all models equally (Danks *et al.*, 2020). This approach is particularly useful in situations where the models under consideration are close in an information theoretic sense, as evidenced by similar relative model likelihoods (Breiman, 1996). Overall, these model selection criteria strengthen PLS's repertoire by allowing researchers to select correctly specified models with low prediction error.

2.4 PLS-SEM criteria for fit measures

Past methodological developments in PLS have also introduced GoF measures to PLS-SEM. One of the earliest metrics, which uses GoF as an index for validating the PLS model globally, was proposed by Tenenhaus *et al.* (2005). Subsequently, some researchers concluded that such GoF measures could be used for theory testing and confirmation (Tenenhaus *et al.*, 2005, p. 173). The GoF procedure uses the quality of the complete measurement model in terms of average communality (i.e. AVE) and the quality of the complete structural model in terms of average R^2. The average of communality is computed as a weighted average of all of the communalities, using weights as the number of manifest variables in each construct with at least two manifest variables. A study by Henseler and Sarstedt (2013) showed that the GoF was not able to distinguish between valid and invalid models and is thus unsuitable for model selection and not applicable to formatively measured constructs. Evidently, the GoF does not penalize over-parameterization efforts, which means that these indices will almost always favor complex models over parsimonious ones, thus resulting in overfitting.

Given all the drawbacks, the methodological developments in PLS-SEM have focused on developing its explanatory strengths by proposing several model fit measures to overcome the heated debate on the absence (Dijkstra and Henseler, 2015a, 2015b; Henseler *et al.*, 2014) (Rönkkö and Evermann, 2013) as well as the deficiencies in Tenenhaus *et al.* (2005) GoF. Some of the recent model fit measures suggested for the PLS-SEM context include SRMR, the root mean square residual covariance (RMStheta), the normed fit index (NFI; also referred to as Bentler–Bonett index), the non-normed fit index (NNFI; also referred to as Tucker–Lewis index), and the exact model fit test (i.e. geodesic discrepancy, d_G and the unweighted least squares discrepancy, d_ULS) (see Dijkstra and Henseler, 2015a; Lohmöller, 1989; Henseler *et al.*, 2014). These model fit measures are capable of judging how well a hypothesized model structure fits the empirical data, and thus help to identify model misspecifications (Dijkstra and Henseler, 2015a; Henseler *et al.*, 2014). Importantly, each of these model fit measures has its own naïve benchmarks and usage recommendations (see Hair *et al.*, 2017b) [9].

However, there are several reasons that model fit measures must be used with caution (Hair *et al.*, 2019b, 2019c). First, the fit measures (SRMR and exact model fit test) are also known as in-sample measures that are oriented towards assessing a model's explanatory power and specification. As a result, these measures provide no guarantee regarding how well the model will fit to another new dataset, nor regarding how generalizable the inferences and policy recommendations will be to other, similar contexts (Petter, 2018). Second, a comprehensive assessment of these measures has not been conducted thus far. Subsequently, any thresholds (guidelines) advocated in the literature should be considered as very tentative. Third, because the algorithm for obtaining PLS-SEM solutions is not based on minimizing the divergence between observed and estimated covariance matrices, the concept of chi-square-based model fit measures, and their extensions—as used in CB-SEM—is not applicable. In addition, even bootstrap-based model fit assessments (d_G and d_ULS) based on, for example, some distance measure or the SRMR (Henseler, 2017; Henseler *et al.*, 2016), which quantifies the discrepancy between the sample covariance matrix and the implied covariance matrix, should be considered with extreme caution. The reason is that the PLS-SEM method considers error variance in model estimation (or maximizing the explained variance) as well as the prediction-oriented goal. Hence, rejecting and modifying a model on the grounds of thresholds for these fit metrics ignores a central component in the PLS-SEM algorithm's objective function (see Lohmöller, 1989, Chapter 5.5, p. 216–222). Fourth, scholars have questioned whether the concept of model fit, as applied in the context of CB-SEM research, is of value to PLS-SEM applications in general (see Hair *et al.*, 2017a, 2017b; Rigdon *et al.*, 2017; Lohmöller, 1989, Chapter 5.5, p. 216–222). Lastly, the model fit criteria that result in a "misfit" in PLS-SEM indicate that more information can be extracted from the data, rather than that the model is incorrect (Sarstedt *et al.*, 2016). From these uncertainty shortcomings, Sarstedt *et al.* (2017) also concluded that validation using GoF measures is also relevant in a PLS-SEM context, but less so than CB-SEM.

3. Study design

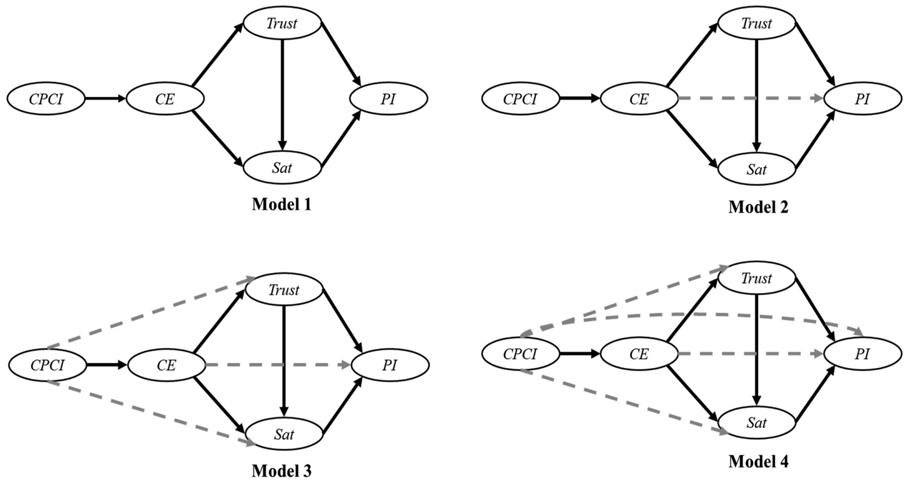
To illustrate the use of the causal prediction procedure with empirical data, we replicate Zhang *et al.* (2018) model on the impact of channel integration on consumer responses in omni-channel retailing [10]. The goal of this model is to explain the effects of customer empowerment (*CE*) (e.g. Prentice *et al.*, 2016; Zhang *et al.*, 2018) on trust (*TRUST*) and satisfaction (*SAT*). Furthermore, the model also includes the dependence of *TRUST* on *SAT* as well as *TRUST* and *SAT* on patronage intention *(PI)*. In addition, the model includes six dimensions (i.e. integrated customer services, integrated information access, integrated order fulfillment, integrated product and price, integrated promotion and integrated transaction information) that configure a higher-order construct of customer perception of channel integration (*CPCI*), as an antecedent construct of omni-channeling (Bendoly *et al.*, 2005; Zhang *et al.*, 2018). The measurement models of *CE*, *TRUST* and *SAT* draw on five reflective items each, whereas *PI* is measured as reflective, with three items. In contrast, the lower-order constructs for *CPCI* (i.e. integrated customer services, integrated information access, integrated order fulfillment, integrated product and price, integrated promotion and integrated transaction information) are measured as reflective with some having four to six items. From these lower-order constructs, *CPCI* is manifest as a formative construct (Zhang *et al.*, 2018), making it a type 2 higher-order (reflective-formative) construct (Sarstedt *et al.*, 2019). All these measurements relied on rating each item on a 7-point Likert scale, with response choices ranging from one (strongly disagree) to seven (strongly agree). In addition, all these variables—regardless whether reflective or formative—that were included in our study were modeled as composites. Composites are formed as linear combinations of sets of indicators or dimensions to represent the concepts in the statistical model, which is well-grounded in the PLS-SEM technique (Hair *et al.*, 2020; Hair and Sarstedt, 2019) [11]. When using these variables in PLS-SEM, the notion of both reflective and formative measurements refers to the epistemic relationship between indicators and constructs as assumed from measurement theory (Sarstedt *et al.*, 2016). In particular, if correlation weights (Mode A) are used for estimating a composite, the arrows typically point away from a construct to the indicators, hence this is often referred to as a reflective measurement model. However, if regression weights are used for compositing a composite, the arrows typically point from the indicators to the construct, hence this is often referred to as a formative measurement.

In addition to looking into the measures, the survey design followed a sequence of steps. A pre-test was conducted on 15 participants who had omni-channel shopping experience, and three experts who had conducted research on multichannel retailing in hypermarkets (i.e. Tesco, Aeon and Carrefour). Prior to administering the formal questionnaire, we modified the inappropriate descriptions and expressions in the items based on the pre-test outcomes (Hulland *et al.*, 2018). Following the pre-test, a pilot test was also conducted with 30 respondents to check for ambiguity in the questionnaire, identify errors and optimize the survey design (Hulland *et al.*, 2018). Like many studies related to retail (see Calvo-Porral and Lévy-Mangín, 2018; Sharma and Lowalekar, 2017), this study followed the non-probabilistic approach of purposive sampling (Sarstedt *et al.*, 2018). Particularly, this study focused on Gen-Y consumers, who are considered to be digital natives, well-educated and sophisticated shoppers (Lissitsa and Kol, 2016; Prasad *et al.*, 2019). Also, respondents were chosen based on their occasional experience with omni-channeling, especially hypermarket visits. To ensure that the participants had a good understanding of omni-channel retailing and were able to provide real, accurate and valid data (Campbell, 1955), we first asked them some screening questions before they started the main survey. For example, only respondents who gave a positive answer to the following question were invited to fill out the survey: "Are you familiar with other channels?" or "Do you often buy products through multiple channels of the retailer?" In addition, when respondents encountered many unknown items in the questionnaire, they were asked to stop the survey. Thus, only those participants who had a good understanding of the omni-channel retailing model finished the questionnaire.

A total of 250 respondents were surveyed for the study. Observations with missing values and straight lining were deleted (Sarstedt and Mooi, 2019; see Chapter 5), leaving a total sample size of 234, which met the minimum sample size suggested by Kock and Hadaya (2018) [12]. The respondents were primarily female ethnic Malays with a master's degree, a monthly income between RM 4,501 and RM 6,500, and a civil servant occupation (Table A3 shows a detailed breakdown of the sample characteristics).

4. Empirical illustration

Our analysis compared five different model configurations of omni-channel retailing with the key target construct of *PI* (see Figure 3) [13]. Model 1 is the theoretically well-established original model that was used by Zhang *et al.* (2018) in prior illustrations of PLS-SEM in the IS field (e.g. Zhang *et al.*, 2018). Model 2 is an extended version of Model 1, in which *CE* influences *PI* directly because consumers who perceive themselves as empowered (i.e. an individual's conviction of self-efficacy determines the initiation of an activity and increases persistence in task performance) are more likely to forge stronger positive behavior (i.e. patronage intention) (Fuchs and Schreier, 2011). Model 3 is a more complex configuration of the original model, in



Note(s): The grey dashed line are the alternative paths that exhibit different plausible theoretical model configurations

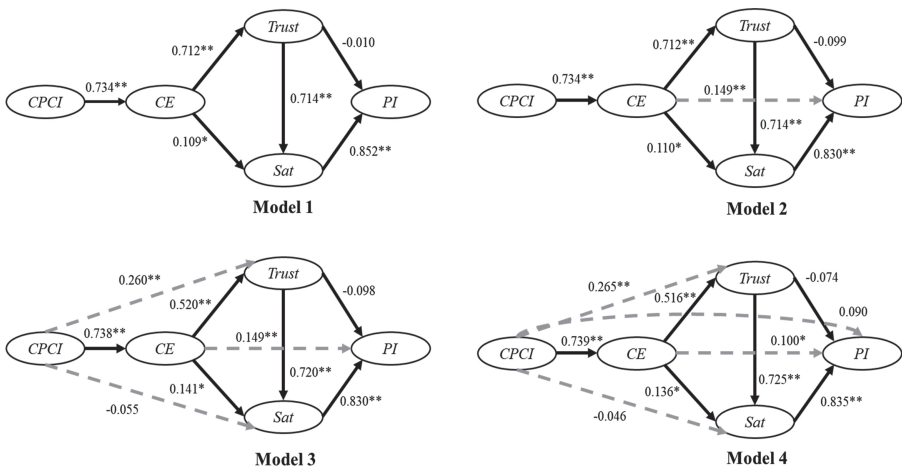
Figure 3.
The four alternative models of channel integration on consumer responses in an omni-channel retailing framework

which the antecedent constructs of *CPCI* also directly influence *TRUST* and *SAT*. Finally, Model 4 is a saturated model, in which the antecedent construct of *CPCI* also directly influences *PI*. Both models 3 and 4 are theoretically plausible when it is assumed that *CE* may only partially mediate the relationship between *CPCI* and endogenous constructs like *TRUST*, *SAT* and *PI*.

5. Model assessment in PLS-SEM

Our analyses utilized SmartPLS 3.3.2 software (Ringle *et al.*, 2015), a prepared Excel table for the computation of the full model selection criteria [14], and the CVPAT package in R statistical software (R Core Team, 2019) [15] to estimate the model parameters. The evaluation of the reflective measurement models by means of standard evaluation criteria (Hair *et al.*, 2019b, 2017b; Sarstedt *et al.*, 2017) supports the measures' reliability and validity (see Table A4). This also holds for the discriminant validity assessment using Henseler *et al.* (2015) recently proposed heterotrait–monotrait ratio of correlations (HTMT) criterion (see Table A5). In particular, the HTMT bias-corrected and accelerated (BCa) confidence intervals—derived from bootstrapping with 5,000 samples and the no sign change option—include the conservative threshold value of 0.90 (Henseler *et al.*, 2015; Franke and Sarstedt, 2019; see Table A5). As for the assessment of lower-order constructs forming the higher-order construct of CPCI, the evaluation of the formative measurement model by means of standard evaluation criteria supports the measure of convergent validity assessment using the redundancy analysis procedure because the threshold value is above 0.7 (Cheah *et al.*, 2018). In addition, the variance inflation factor (VIF) of each dimension is below the threshold value of 3 (Mason and Perreault, 1991; Becker *et al.*, 2015; see Table A6). We also ran bootstrapping with 5,000 samples, using the no sign change option, to test the lower order construct weights' significance based on the 99% BCa confidence interval (Hair *et al.*, 2017b, see Chapter 5). The results showed that all lower order construct weights were significant ($p < 0.01$).

Finally, both Figure 4 and Table A7 illustrate the path models (Model 1 to Model 4) used in our structural model assessment [16]. The bootstrapping results for Model 1 show that most of the path coefficients are statistically significant for all relationships (p-value < 0.01), except



Note(s): i. *p*-value < 0.05; ** *p*-value < 0.01 (one-tailed test)

ii. Model 1: The R^2 results of each endogenous variable are CE: 0.539, Trust: 0.506, Sat: 0.633, and PI: 0.713

iii. Model 2: The R^2 results of each endogenous variable are CE: 0.538, Trust: 0.507, Sat: 0.633, and PI: 0.724

iv. Model 3: The R^2 results of each endogenous variable are CE: 0.545, Trust: 0.537, Sat: 0.634, and PI: 0.724

v. Model 4: The R^2 results of each endogenous variable are CE: 0.546, Trust: 0.539, Sat: 0.633, and PI: 0.727

Figure 4.
Path model result of model 1 to model 4

for *TRUST* on *PI* ($\beta = -0.010$; SE = 0.059). Similarly, Model 2 shows that *TRUST* on *PI* is insignificant ($\beta = -0.099$; SE = 0.075) but the other paths are all significant (*p*-value < 0.05). As for Model 3, the path coefficients are statistically significant for all relationships, except for *CPCI* on *SAT* ($\beta = -0.055$; SE = 0.056) and *TRUST* on *PI* ($\beta = -0.098$; SE = 0.072). Finally, Model 4 shows that all results are statistically significant, except for *CPCI* on *SAT* ($\beta = -0.046$; SE = 0.059), *CPCI* on *PI* ($\beta = 0.090$; SE = 0.072) and *TRUST* on *PI* ($\beta = -0.074$; SE = 0.062).

6. Model comparison results using PLS-SEM criteria

Table 1 shows the results of our study when the robustness of the 4 models was assessed. The PLS-SEM criteria (i.e. GoF by Tenenhaus, R^2 and Q^2) show that the best model is Model 4, which is the most complex model in our set and the least theoretically defensible. In addition, the preference for the saturated model (Model 4) is supported by the model selection criteria results of the asymptotically efficient (i.e. AIC, AICc and Mallow's Cp) and the out-of-sample criteria (i.e. Q^2_predict). Even the adjusted R^2, which is designed to adjust for parsimony, shows greater preference for Model 4. In addition, Model 1 produced a better result in terms of PLSpredict, especially when looking into the value of both the RMSE and MAE metrics. However, Model 1 is still not the best model when compared with other alternatives because of the favorable results of the other criteria.

Furthermore, Models 2 and 3 performed similarly when compared with Models 1 and 4. Interestingly, Models 2 and 3 yielded a better model based on asymptotically consistent

	Criteria	Model 1	Model 2	Model 3	Model 4
PLS Based	GoF by Tenenhaus	0.539	0.542	0.550	0.551
	R^2	0.713	0.724	0.724	0.727
	Adj R^2	0.710	0.720	0.720	0.722
	Q^2	0.602	0.611	0.611	0.616
Asymptotically Efficient	AIC	-287.098	-294.243	-294.243	-294.800
	AICu	-284.079	-290.208	-290.208	-289.746
	AICc	-50.923	-57.980	-57.980	-58.430
	Mallow's Cp	16.949	9.560	9.560	9.000
	FPE	0.293	0.284	0.284	0.284
Asymptotically Consistent	BIC	-276.732	-280.422	-280.422	-277.524
	GM	261.315	257.382	257.382	260.277
	HQ	-282.918	-288.670	-288.670	-287.835
	HQc	-282.696	-288.313	-288.313	-287.311
PLSpredict	RMSE	0.388	0.433	0.498	0.544
	MAE	0.307	0.338	0.389	0.425
	Q^2_Predict	0.122	0.162	0.159	0.235

Note(s): Grey shading represents best value

Table 1.
Criteria values for alternative Models 1–4

selection criterion (i.e. BIC, GM, HQ and HQC) compared with the other models. However, the PLSpredict result for Model 2 exhibits a better value (i.e. RMSE, MAE and Q^2_predict) than Model 3 for the final target construct *PI*, making it a promising alternative to the other models. From these findings, we believe that the more parsimonious research models showed stronger performance for out-of-sample prediction (RMSE and MAE) among a set of competing models, but not necessarily for in-sample model selection criteria.

In order to minimize the uncertainty of the model selection criteria results (AIC, BIC and GM values in Table 1), we performed an assessment of AICw, BICw and GMw. Based on Table 2, AICw firmly disqualifies the parsimonious models (Model 1) with a very low weight of 0.008. AICw also discards Models 2 and 3, but the difference with Model 4 is much less pronounced, with AIC–AIC_{min} = 0.391. On the contrary, BICw clearly disqualifies the more complex Model 1, as Model 1 shows a pronounced difference from Models 2 and 3, which yield the smallest BIC value (BIC–BIC_{min} = 3.690). BICw also discards the saturated Model 4, but the difference from Model 1 is much less pronounced (BIC–BIC_{min} = 2.898). Overall, considering the low weights of Models 1 (0.062) and 4 (0.092), both models offer more compelling evidence of model misspecification. Similarly, the GMw performs better for Models 2 and 3 than for Models 1 and 4 based on the validation data.

With regard to the selection uncertainty (Table 2), AICw strongly favors Model 4, with some support given to Models 2 and 3. Even though the complex models are theoretically least defensible (Sharma *et al.*, 2019a, 2019b), the fact that AIC has a strong tendency to select overparametrized models should give rise to concern regarding this choice. However, BICw and GMw assign the highest weights to Models 2 and 3—the most theoretically defensible models. Overall, we find that the use of both BIC/BICw and GM/GMw complement the uncertainty result generated from the model selection criteria's raw values. In particular, these criteria are able to identify the relative strength of a single model as clearly superior to other sets of competing models.

Next, we tested the CVPAT method with the goal of determining whether the AM (Models 2, 3 and 4) offers significantly higher predictive power than the EM (Model 1) (see Table 3). We found that the EM exhibited lower loss than the AM for both comparison 2 [Model 1 (EM):

Criteria	Model 1	Model 2	Model 3	Model 4
AIC	-287.098	-294.243	-294.243	-294.800
Δ (AIC - AIC_{min})	7.702	0.557	0.557	0.000
AIC weights	0.008	0.296	0.296	0.391
BIC	-276.732	-280.422	-280.422	-277.524
Δ (BIC – BIC_{min})	3.690	0.000	0.000	2.898
BIC weights	0.062	0.392	0.392	0.092
GM	261.315	257.382	257.382	260.277
Δ (GM – GM_{min})	3.933	0.000	0.000	2.895
GM weights	0.056	0.398	0.398	0.094

Note(s): Grey shading represents best value

Table 2. AICw, BICw and GMw value for alternative Models 1–4

Comparison	CVPAT results	Averages losses EM	AM	EM-AM	*p*-value*	CI**
1	Model 1 (EM) and Model 2 (AM)	0.523	0.522	0.001	0.045	[0.001; ∞]
2	Model 1 (EM) and Model 3 (AM)	0.516	0.522	−0.007	0.095	[−0.002; ∞]
3	Model 1 (EM) and Model 4 (AM)	0.519	0.522	−0.004	0.251	[−0.005; ∞]

Note(s): (1) EM = established model; AM = alternative model; we use Model 1 as EM to compare with other AM models (Model 2, Model 3 and Model 4). The reason is Model 1 is derived from an established theoretical justification work by Zhang *et al.* (2018), (2) * The null hypothesis is equal predictive ability and the alternative hypothesis is that the AM (column 3) has better predictive ability than the EM (column 2); the *p*-value is based on 10,000 bootstrap samples with a specified seed of 42, (3) ** CI = 95% confidence interval of the one-sided test, (4) A negative average loss value difference between the EM and AM indicates that the EM has a smaller average loss and is therefore preferable. If the average loss value difference is positive, the AM has superior predictive power

Table 3. The CVPAT results for alternative Models 1–4

0.516; Model 3 (AM): 0.522] and comparison 3 [Model 1 (EM): 0.519; Model 4 (AM): 0.522)], which supported retaining the EM as the best predictive model. Subsequently, these comparisons were also supported with insignificant *p*-value and confidence interval results, which indicated that the EM of Model 1 had a lesser average prediction error loss than the AM (Model 3 and Model 4). In contrast, our study also found that comparison 1 exhibited a slightly high EM loss (Model 1: 0.523) than the AM loss (Model 2: 0.522). This comparison was supported by the significant inference result of both the *p*-value and confidence interval; thus, comparison 1 showed that Model 2 (AM) was a better predictive model than Model 1. Because only Model 2 (AM) has higher predictive accuracy for our CVPAT result, we would prefer this model over the original version of Model 1 (EM) proposed by Zhang *et al.* (2018).

Finally, this study also assessed the model fit criteria in PLS-SEM (see Table 4) [17]. The model fit criteria (i.e. SRMR, NFI and chi-squared) favors Model 3 over the other three models. However, the differences in this value for each model fit criteria are relatively small (especially Model 4—the saturated model), and it is difficult to ensure that Model 3 exhibits a better fit. In addition, the use of bootstrap-based tests (d_G and d_ULS) shows that the original value does not fall into the 95% (or 99%) confidence interval across the models, except for the bootstrap-based model fit assessment for SRMR. Particularly, the results in Table 4 show that the values for SRMR do not exceed the threshold value of 0.08 (Hu and Bentler, 1998, 1999) or the upper bound values of both the 95% and 99% quantiles of their reference distribution, thus

Table 4.
Assessment of Model Fit in PLS-SEM for alternative Models 1–4

	Criteria	Model 1	Model 2	Model3	Model 4
Model Fit: Estimated Model	SRMR	0.073 95% [0.074] 99% [0.079]	0.069 95% [0.073] 99% [0.079]	0.055 95% [0.057] 99% [0.060]	0.056 95% [0.058] 99% [0.062]
	d_ULS	1.602 95% [0.630] 99% [0.800]	1.449 95% [0.593] 99% [0.743]	0.918 95% [0.391] 99% [0.455]	0.926 95% [0.388] 99% [0.452]
	d_G	0.677 95% [0.367] 99% [0.407]	0.670 95% [0.367] 99% [0.407]	0.656 95% [0.367] 99% [0.409]	0.663 95% [0.365] 99% [0.405]
	Ch-Square	826.904	818.188	811.119	814.200
	NFI	0.862	0.863	0.864	0.864

Note(s): i. Grey shading represents best value
ii. The 95% and 99% of SRMR, d_ULS and d_G are bootstrapped according to suggested procedure by Dijkstra and Henseler (2015b, 2015a)

demonstrating that model fit is obtained by all four models, except for both the d_G and the d_ULS. Because the SRMR fit measure exhibits an acceptable fit result that is identical on all four models, this fit criterion would still not be sufficient to explain which model has the best explanatory power and specification. In addition, the SRMR fit measure provides no guarantee regarding how well the model fits another dataset. Therefore, the use of the model fit criteria did not clearly show the best fit preference for deciding causal prediction in this empirical study.

In conclusion, comparing the four model configurations in terms of all the criteria values revealed that the original Model 2 variant clearly outperformed models that were more complex (Model 3 and Model 4), as well as models that were too parsimonious (Model 1). In other words, when striking a balance between the explanation and prediction goals of PLS-SEM, the most theoretically well-established model should be Model 2, as it achieved reasonable results for most of the criteria, such as the PLSpredict, the CVPAT and the model selection criteria (i.e. BIC, BICw, GM, GMw, HQ and HQ_C). Therefore, our study shows that Model 2 has a better chance of being scientifically replicable, explainable and exhibiting higher predictive abilities (Shmueli and Koppius, 2011; Bentler and Mooijaart, 1989), because of its balanced result in terms of in-sample and out-of-sample prediction power.

7. Discussion

Since its inception, PLS-SEM has been an exploratory technique for theory building where researchers might want to compare several models (Wold, 1982, 1985). Recent work in the PLS-SEM literature has also highlighted its abilities as a causal-predictive technique (Shmueli *et al.*, 2016, 2019). Because PLS-SEM straddles the divide between causal explanation and prediction, researchers using the method need to ensure that the estimated model adequately maps reality while offering sufficient predictive capabilities (Shmueli *et al.*, 2016). While prior studies have evaluated the efficacy of model selection criteria in CB-SEM for selecting a specific or appropriate model from among a set of competing models (see Gangwar and Date, 2016; Shiau and Chau, 2012; Shiau and Chau, 2016), none have examined the causal-predictive perspective in the IS field, where the goal is to select the appropriate model with both explanatory and predictive power.

Using an empirical study, our results show that the practice of comparing models using the standard model evaluation criteria (i.e. R^2, Adjusted R^2, GoF and Q^2) were inadequate because these criteria display a pronounced preference for the saturated model (Model 4), which is the most complex model in our set and the least theoretically defensible. In other words, solely focusing on these standard model evaluation criteria to assess the adequacy of a theoretical model is problematic because this may cause researchers to overfit their model to the point that it will predict poorly and may not be generalizable or replicable by other researchers (Hair *et al.*, 2019b). On the other hand, the model selection criteria cannot be substituted for each other when selecting the appropriate causal-predictive model from among a cohort of competing models. This technique can help researchers to reduce uncertainty in model selection and to determine the appropriate theoretical research model in IS research (see Cheah *et al.*, 2019). The model selection criteria, particularly BIC and GM, have significantly higher agreement levels over the set of specified models (for Models 1–4) compared with the PLS-SEM criteria (i.e. GOF, R^2, adjusted R^2 and Q^2), therefore the criteria safeguard against the overfitting issue, particularly in situations where researchers build complex PLS models. In addition, our findings echo the results of Sharma *et al.* 's (2019a, 2019b) earlier simulations, where the use of model selection criteria (BIC and GM) enable researchers to select correctly specified models when comparing predictive generalizability.

Our study also encourages the use of BICw and GMw, which can be interpreted as conditional probabilities for models (e.g. Burnham and Anderson, 2002; Wagenmakers and Farrell, 2004), thereby offering stronger evidence for or against each model in the set (Danks *et al.*, 2020). Both these criteria help corroborate BIC and GM results when comparing models that are close in the theoretical information sense, and a single superior model cannot easily be identified (Preacher and Merkle, 2012). In other words, the use of the BICw and GMw criteria facilitate researchers in overcoming the false sense of confidence that occurs when selecting between models with similar BIC and GM values.

Although BIC/BICw and GM/GMw are ideal criteria for prediction-oriented model selection, researchers should not neglect the out-of-sample prediction using PLSpredict criteria (RMSE, MAE or Q^2_predict) and CVPAT. Both the PLSpredict and CVPAT techniques allow researchers to address the long-standing calls for a stronger focus on predictive model assessment, most notably a model's out-of-sample predictive power (Liengaard *et al.*, 2020; Shmueli *et al.*, 2016, 2019). By having low prediction errors (e.g. using the RMSE and MAE statistic) and a high value of Q^2_predict, researchers can identify a parsimonious model that is more likely to predict and be generalizable to other samples. Similarly, having appropriate pairwise comparison results for CVPAT (in our case, this is average prediction error of EM: Model 1 > average prediction error of AM: Model 2) with its overall inferential test enables researchers to statistically compare the predictive strengths of models to judge whether model choice is reliable, and not affected by the chance of sampling error. Therefore, we summarize that the introduction of both the PLSpredict and CVPAT in PLS-SEM can aid researchers in developing and examining theoretical models via comparison, improvement in measures and construct operationalization, as well as benchmarking the predictability of a given phenomenon.

However, our empirical result exhibits misfit for the majority of the fit criteria (i.e. d_ULS, d_G, chi-squared and NFI) when assessing the discrepancy between the empirical correlation matrix and the model-implied correlation matrix of all models (Models 1–4), except for the SRMR. Drawing from this finding, this raises an important question as noted by Hair *et al.* (2019c), about whether researcher should cherry-pick the appropriate result when reporting the model fit criteria? Cherry-picking a result from the model fit criteria—in this case, it will be just the SRMR fit measure—can be harmful, as researchers may be tempted to sacrifice the exploratory and causal-predictive nature of social science research. Hence, researchers should be more cautious when interpreting the fit criteria, because little is known about the viability of the fit measures across a range of data. Model constellations are proposed for PLS-SEM in the literature, and their use can even be harmful, as researchers may be tempted to "torture the data until they confess" (Aguinis *et al.*, 2017, p. 654). In fact, should we actually reject all four models when the majority of the fit measures exhibit misfit? Rejecting and modifying a model on the basis of satisfying these fit metrics ignores a central component in the PLS-SEM algorithm's objective function (Lohmöller, 1989). As an alternative to these fit metrics, researchers can revert to the model selection criteria (Sharma *et al.*, 2019a, 2019b), PLSpredict (Shmueli *et al.*, 2016; 2019) and CVPAT (Liengaard *et al.*, 2020) because these criteria achieve a sound tradeoff between explanation and prediction to avoid overfitting, so that the model generalizes beyond the particular sample in the estimation of PLS path models, which perfectly satisfies the method's causal-predictive nature (Jöreskog and Wold, 1982, p. 270).

Overall, it is worth noting that PLS-SEM has become increasingly popular in IS research for evaluating models with complex relationships, and for making valid inferences from a restricted sample to a larger population. Subsequently, a sharper focus on demystifying the role of causal-predictive modeling can help to connect the subjective and objective realities, as well as assess the distance between theory and practice in the IS field and narrow the range of possibilities to ensure successful policy-making. Given the larger goal of creating generalizable theories in IS and even business research, we argue that among all the PLS criteria, the model selection criteria (i.e. BIC/BICw and GM/GMw), PLSpredict and CVPAT enable researchers to compare competing PLS path models by catering to both theory development and predictions aspects. For example, it is widely known that researchers prefer to add more variables (or paths) to the model and to rely on the statistical significance to imply causal prediction in aiding theory development (Sharma *et al.*, 2019a, 2019b). Hence, the use of the model selection criteria, PLSpredict and CVPAT can provide complementary information regarding whether the inclusion of the variable(s) in theoretical modeling has been successful in improving both the in-sample and out-sample information at hand.

In addition, demystifying the role of causal-predictive modeling is crucial for making decisions under conditions of uncertainty, and in creating opportunities to compete for research. From an academic perspective, most PLS-SEM researchers do not simply choose or specify only one model structure, and then use the data to either confirm or contradict the specific structure for theory development (Zheng and Pavlou, 2010). Rather, they often face choices during the literature review process that require a decision to design, select or extend an appropriate model from a set of competing models (Cheah *et al.*, 2019). While from practitioner's standpoint, the focus is not typically on validating or testing theories from a set of competing models, but rather on finding generalizable approaches or policies that could result in immediate commercial utility or predictive power (Ruddock, 2017). We therefore call for the use and report of model selection criteria, PLSpredict, and CVPAT results in PLS-SEM because these criteria enable researchers to reduce the uncertainty of model choice by ruling out alternative explanations and identification of variables with high predictive power. In the meantime, these criteria enable practitioners to make such decisions with less error by reducing generalization error so that policy decisions will be more likely to work in other settings.

Our research sheds light on the performance of causal prediction criteria in the context of IS research. Future research should benchmark PLS-SEM's criteria performance against other composite-based SEM methods, such as consistent PLS (Dijkstra and Henseler, 2015a, 2015b), weighted PLS (Becker and Ismail, 2016; Cheah *et al.*, 2020), generalized structured component analysis (Hwang *et al.*, 2010) and regularized canonical correlation analysis (Tenenhaus and Tenenhaus, 2011). In addition, it is suggested that future studies extend our empirical design by considering more complex model structures, such as interaction terms (Becker *et al.*, 2018), mediating effects (Nitzl *et al.*, 2016) and non-linear effects (Ahrholdt *et al.*, 2019; Hair *et al.*, 2018), experimental studies (Streukens and Leroi-Werelds, 2016), longitudinal studies (Roemer, 2016) and discrete choice modeling (Hair *et al.*, 2019a). While the use of these modeling elements is recently becoming more in vogue in IS research, nothing is known about which criteria perform the best in achieving the causal prediction goal of PLS-SEM.

Notes

1. Theoretical consistency refers to a well-developed causal explanation for the suggested theoretical constructs, propositions, hypothesis, boundary conditions and assumptions that are logically consistent and relate to each other. As a result, it guarantees the generalizability of the inference.

2. Causal-predictive technique assumes that when the structural theory of a model is strong and well-developed (i.e. a well-established research domain, where extensive previous studies and insights have been performed to investigate a phenomenon), the path relationships can be interpreted as causal in predictive modeling, where a model predicts unseen or new data (Ringle *et al.*, 2020; Liengaard *et al.*, 2020).

3. In-sample predictions (often called fitted values) can be useful for explanatory modeling efforts, in that their computation draws on the entire dataset. In other words, their computation requires estimating the parameters of a path model, and then using the model to predict values for cases from the same sample (see Shmueli *et al.*, 2016).

4. Danks *et al.* (2020) study illustrated that selecting one model over another based on model selection criteria may lead to a false sense of confidence, as the differences in the criteria values are often small.

5. Standardized root mean square residual (SRMR) was originally developed by Bentler (1995) for the covariance-based structural equation modeling (CB-SEM) technique.

6. Overfitting refers to a research model that models the training data or in-sample data too well during its computation of the entire data. The computation requires estimating the parameters of a PLS model and then using the model to predict values for cases from the same sample. An overfitting scenario occurs because the analysis corresponds too closely to a particular set of data (i.e. in-sample data), thus it fails to fit additional data to the research model, or to predict future observations reliably (Shmueli *et al.*, 2016).

7. In CVPAT-based model comparison, a one-sided test is often used to determine whether the AM offers significantly higher predictive power than the EM. However, it is also possible to carry out a two-sided test when researchers consider equally suitable models of both EM and AM. A significant CVPAT result could provide evidence in favor of the one or the other model; if not, we cannot reject their having equal predictive accuracy.

8. A piecewise linear regression in PLS-SEM runs separate regressions of each dependent construct in the structural model on its associated independent constructs. In other words, the algorithm maximizes the endogenous latent variable's explained variance in the structural model stage by estimating partial model relationships in an iterative sequence of OLS regressions.

9. Note that Hair *et al.*'s (2017b, p. 193–194) book contains all the compiled threshold values of the model fit criteria (i.e. SRMR, NFI, RMS, d_G, d_ULS and RMStheta).

10. Zhang *et al.* (2018) work was published in Electronic Commerce Research and Applications.

11. The use of composite modeling in PLS-SEM is a suitable and advantageous choice when estimating relationships among conceptual variables that are both reflective and formative. This is particularly true because PLS-SEM enables researchers (1) to estimate complex models comprised of many latent and observed variables, (2) to obtain solutions with small sample sizes without identification or convergence concerns and (3) to provide an aspect of predictive modeling that complements the retrospective nature of causal-explanatory modeling (Hair and Sarstedt, 2019; Hair *et al.*, 2020).

12. The reasonable value for minimum sample size in PLS-SEM by the inverse square root method is about 160, and by the gamma exponential methods is about 146 (Kock and Hadaya, 2018).

13. The indicator correlation matrix lead result can be obtained upon request from the corresponding author.
14. The computations of the model selection criteria were estimated using the Excel file provided at the following webpage: https://www.pls-sem.net/downloads/sample-projects-and-various/.
15. The CVPAT code for the statistical software R and technical instructions for its application are available for download at the following webpage: https://github.com/ECONshare/CVPAT/.
16. See appendix, Table A7 for the specific result of each of the path coefficients, t-value, *p*-value, confidence interval and the effect size result (f2).
17. The RMStheta index is only useful for assessing purely reflective models, because outer model residuals for formative measurement constructs are not meaningful (see Hair *et al.*, 2017b; Henseler *et al.*, 2016) Given that the four models contain a higher-order, reflective-formative construct (i.e. CPCI), this index was not used in the study.

References

Addas, S. and Pinsonneault, A. (2018), "E-mail interruptions and individual performance: is there a silver lining?", *MIS Quarterly*, Vol. 42 No. 2, pp. 381-405.

Aguinis, H., Cascio, W.F. and Ramani, R.S. (2017), "Science's reproducibility and replicability crisis: international business is not immune", *Journal of International Business Studies*, Vol. 48, pp. 653-663.

Ahrholdt, D.C., Gudergan, S.P. and Ringle, C.M. (2019), "Enhancing loyalty: when improving consumer satisfaction and delight matters", *Journal of Business Research*, Vol. 94, pp. 18-27.

Ainin, S., Parveen, F., Moghavvemi, S., Jaafar, N.I. and Mohd Shuib, N.L. (2015), "Factors influencing the use of social media by SMEs and its performance outcomes", *Industrial Management and Data Systems*, Vol. 115 No. 3, pp. 570-588.

Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle", in Petrov, B.N. and Csáki, F. (Eds), *Selected Papers of Hirotugu Akaike*, Springer, New York.

Ali, M.H., Zhan, Y., Alam, S.S., Tse, Y.K. and Tan, K.H. (2017), "Food supply chain integrity: the need to go beyond certification", *Industrial Management and Data Systems*, Vol. 117 No. 8, pp. 1589-1611.

Alonso Dos Santos, M., Rejón Guardia, F. and Calabuig Moreno, F. (2018), "Sponsorship image transfer theory in virtual brand communities", *Industrial Management and Data Systems*, Vol. 118 No. 6, pp. 1287-1302.

Andrei, A.G., Zait, A., Vătămănescu, E.M. and Pînzaru, F. (2017), "Word-of-mouth generation and brand communication strategy: findings from an experimental study explored with PLS-SEM", *Industrial Management and Data Systems*, Vol. 117 No. 3, pp. 478-495.

Au, N., Ngai, E.W. and Cheng, T.E. (2008), "Extending the understanding of end user information systems satisfaction formation: an equitable needs fulfillment model approach", *MIS Quarterly*, Vol. 32 No. 1, pp. 43-66.

Bao, H., Li, B., Shen, J. and Hou, F. (2016), "Repurchase intention in the Chinese e-marketplace: roles of interactivity, trust and perceived effectiveness of e-commerce institutional mechanisms", *Industrial Management and Data Systems*, Vol. 116 No. 8, pp. 1759-1778.

Becker, J.M. and Ismail, I.R. (2016), "Accounting for sampling weights in PLS path modeling: simulations and empirical examples", *European Management Journal*, Vol. 34 No. 6, pp. 606-617.

Becker, J.-M., Ringle, C.M., Sarstedt, M. and Völckner, F. (2015), "How collinearity affects mixture regression results", *Marketing Letters*, Vol. 26, pp. 643-659.

Becker, J.-M., Ringle, C.M. and Sarstedt, M. (2018), "Estimating moderating effects in PLS-SEM and PLSc-SEM: interaction term generation*data treatment", *Journal of Applied Structural Equation Modeling*, Vol. 2, pp. 1-21.

Bendoly, E., Blocher, J.D., Bretthauer, K.M., Krishnan, S. and Venkataramanan, M.A. (2005), "Online/in-store integration and customer retention", *Journal of Service Research*, Vol. 7, pp. 313-327.

Bentler, P.M. and Mooijaart, A.B. (1989), “Choice of structural model via parsimony: a rationale based on precision”, *Psychological Bulletin*, Vol. 106, pp. 315-317.

Bentler, P.M. (1995), *EQS. Structural Equations Program Manual*, Multivariate Statistical Software, Los Angeles, CA.

Bentler, P.M. and Huang, W. (2014), “On components, latent variables, PLS and simple methods: reactions to Rigdon’s rethinking of PLS”, *Long Range Planning*, Vol. 47, pp. 138-145.

Berk, R.A. (2008), *Statistical Learning from a Regression Perspective*, Springer, New York.

Breiman, L. (1996), “Heuristics of instability and stabilization in model selection”, *The Annals of Statistics*, Vol. 24, pp. 2350-2383.

Bryan Foltz, C., Schwager, P.H. and Anderson, J.E. (2008), “Why users (fail to) read computer usage policies”, *Industrial Management and Data Systems*, Vol. 108 No. 6, pp. 701-712.

Burnham, K. and Anderson, D. (2002), *Model Selection and Multi-Model Inference*, Springer, Heidelberg.

Calvo-Porral, C. and Lévy-Mangín, J.-P. (2018), “Pull factors of the shopping malls: an empirical study”, *International Journal of Retail and Distribution Management*, Vol. 46, pp. 110-124.

Campbell, D.T. (1955), “The informant in quantitative research”, *American Journal of Sociology*, Vol. 60, pp. 339-342.

Chan, A.L.C. and Wang, W.Y. (2012), “The causal relationships between aspects of customer capital”, *Industrial Management and Data Systems*, Vol. 112 No. 6, pp. 848-865.

Cheah, J.-H., Sarstedt, M., Ringle, C.M., Ramayah, T. and Ting, H. (2018), “Convergent validity assessment of formatively measured constructs in PLS-SEM: on using single-item versus multi-item measures in redundancy analyses”, *International Journal of Contemporary Hospitality Management*, Vol. 30, pp. 3192-3210.

Cheah, J.-H., Ting, H., Cham, T.H. and Memon, M.A. (2019), “The effect of selfie promotion and celebrity endorsed advertisement on decision-making processes: a model comparison”, *Internet Research*, Vol. 29, pp. 552-577.

Cheah, J.H., Roldán, J.L., Ciavolino, E., Ting, H. and Ramayah, T. (2020), “Sampling weight adjustments in partial least squares structural equation modeling: guidelines and illustrations”, *Total Quality Management and Business Excellence*, Forthcoming.

Chen, M. and Qi, X. (2015), “Members’ satisfaction and continuance intention: a socio-technical perspective”, *Industrial Management and Data Systems*, Vol. 115 No. 6, pp. 1132-1150.

Chen, J.V., Chen, Y. and Capistrano, E.P.S. (2013), “Process quality and collaboration quality on B2B e-commerce”, *Industrial Management and Data Systems*, Vol. 113 No. 6, pp. 908-926.

Chi, M., Zhao, J. and George, J.F. (2015), “Mediation and time-lag analyses of e-alignment and e-collaboration capabilities”, *Industrial Management and Data Systems*, Vol. 115 No. 6, pp. 1113-1131.

Chin, W.W. (1998), “The partial least squares approach to structural equation modeling”, in Marcoulides, G.A. (Ed.), *Modern Methods for Business Research*, Erlbaum, Mahwah.

Chin, W.W. (2010), “How to write up and report PLS analyses”, in Esposito Vinzi, V., Chin, W.W., Henseler, J. and Wang, H. (Eds), *Handbook of Partial Least Squares*, Springer-Verlag, Berlin Heidelberg, pp. 655-690.

Chuah, S.H.W., Rauschnabel, P.A., Tseng, M.L. and Ramayah, T. (2018), “Reducing temptation to switch mobile data service providers over time: The role of dedication vs constraint”, *Industrial Management and Data Systems*, Vol. 118 No. 8, pp. 1597-1628.

Clark, J.W., Toms, L.C. and Green, K.W. (2014), “Market-oriented sustainability: moderating impact of stakeholder involvement”, *Industrial Management and Data Systems*, Vol. 114 No. 1, pp. 21-36.

Compeau, D.R. and Higgins, C.A. (1995), “Computer self-efficacy: development of a measure and initial test”, *MIS Quarterly*, Vol. 19 No. 2, pp. 189-211.

Curtis, A.M., Dennis, A.R. and Hilmer, K.M. (2017), "From Monologue to Dialogue: performative Objects to Promote Collective Mindfulness in Computer-Mediated Team Discussions", *MIS Quarterly*, Vol. 41 No. 2, pp. 559-581.

Dalvi-Esfahani, M., Ramayah, T. and Rahman, A.A. (2017), "Moderating role of personal values on managers' intention to adopt Green IS: Examining norm activation theory", *Industrial Management and Data Systems*, Vol. 117 No. 3, pp. 582-604.

Danks, N.P. and Ray, S. (2018), "Predictions from partial least squares models", in Ali, F., Rasoolimanesh, S.M. and Cobanoglu, C. (Eds), *Applying Partial Least Squares in Tourism and Hospitality Research*, Emerald Publishing.

Danks, N.P., Sharma, P.N. and Sarstedt, M. (2020), "Model selection uncertainty and multimodel inference in partial least squares structural equation modeling (PLS-SEM)", *Journal of Business Research*, Vol. 113, pp. 13-24.

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

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

Ding, Y. (2018), "I hope and I continue: Integrating the concept of hope into the expectancy-disconfirmation framework", *Industrial Management and Data Systems*, Vol. 118 No. 4, pp. 728-744.

Dublin, R. (1969), *Theory Building: A Practical Guide to the Construction and Testing of Theoretical Model*, The Free Press, New York, NY.

Enns, H.G., Huff, S.L. and Higgins, C.A. (2003), "CIO lateral influence behaviors: Gaining peers' commitment to strategic information systems", *MIS Quarterly*, Vol. 27 No. 1, pp. 155-176.

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

Fan, J., Shao, M., Li, Y. and Huang, X. (2018), "Understanding users' attitude toward mobile payment use: A comparative study between China and the USA", *Industrial Management and Data Systems*, Vol. 118 No. 3, pp. 524-540.

Falk, R.F. and Miller, N.B. (1992), *A Primer for Soft Modeling*, University of Akron Press, Akron, OH.

Fang, Y., Qureshi, I., Sun, H., McCole, P., Ramsey, E. and Lim, K.H. (2014), "Trust, satisfaction, and online repurchase intention: the moderating role of perceived effectiveness of e-commerce institutional mechanisms", *MIS Quarterly*, Vol. 38 No. 2, pp. 407-427.

Franke, G.R. and Sarstedt, M. (2019), "Heuristics versus statistics in discriminant validity testing: a comparison of four procedures", *Internet Research*, Vol. 29, pp. 430-447.

Fuchs, C. and Schreier, M. (2011), "Customer empowerment in new product development", *Journal of Product Innovation Management*, Vol. 28, pp. 17-32.

Furneaux, B. and Wade, M.R. (2011), "An exploration of organizational level information systems discontinuance intentions", *MIS Quarterly*, Vol. 35 No. 3, pp. 573-598.

Gangwar, H. and Date, H. (2016), "Critical factors of cloud computing adoption in organizations: an empirical study", *Global Business Review*, Vol. 17, pp. 886-904.

Geisser, S. (1974), "A predictive approach to the random effect model", *Biometrika*, Vol. 61, pp. 101-107.

Gil-Saura, I., Ruiz-Molina, M.E. and Arteaga-Moreno, F. (2011), "Value, supplier dependence and long-term orientation: outcomes for B2B commerce in the travel industry", *Industrial Management and Data Systems*, Vol. 111 No. 5, pp. 791-808.

Gold, A.H., Malhotra, A. and Segars, A.H. (2001), "Knowledge management: an organizational capabilities perspective", *Journal of Management Information Systems*, Vol. 18, pp. 185-214.

Gómez-Cedeño, M., Castán-Farrero, J.M., Guitart-Tarrés, L. and Matute-Vallejo, J. (2015), "Impact of human resources on supply chain management and performance", *Industrial Management and Data Systems*, Vol. 115 No. 1, pp. 129-157.

Gonzalez, R., Gasco, J.L. and Llopis, J. (2015), "Information systems outsourcing satisfaction: some explanatory factors", *Industrial Management and Data Systems*, Vol. 115 No. 6, pp. 1067-1085.

Goo, J., Kishore, R., Rao, H.R. and Nam, K. (2009), "The role of service level agreements in relational management of information technology outsourcing: an empirical study", *MIS Quarterly*, Vol. 33 No. 1, pp. 119-145.

Hair, J.F. Jr and Sarstedt, M. (2019), "Factors versus composites: guidelines for choosing the right structural equation modeling method", *Project Management Journal*, Vol. 50, pp. 619-624.

Hair, J.F., Ringle, C.M. and Sarstedt, M. (2011), "PLS-SEM: indeed a silver bullet", *Journal of Marketing Theory and Practice*, Vol. 19, pp. 139-151.

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

Hair, J., Hollingsworth, C.L., Randolph, A.B. and Chong, A.Y.L. (2017a), "An updated and expanded assessment of PLS-SEM in information systems research", *Industrial Management and Data Systems*, Vol. 117, pp. 442-458.

Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2017b), *A Primer on Partial Least Squares Structural Equation Modeling*, PLS-SEM)Sage, Thousand Oaks, CA.

Hair, J.F., Sarstedt, M., Ringle, C.M. and Gudergan, S.P. (2018), *Advanced Issues in Partial Least Squares Structural Equation Modeling*, PLS-SEM)Sage, Thousand Oaks, CA.

Hair, J.F., Ringle, C.M., Gudergan, S.P., Fischer, A., Nitzl, C. and Menictas, C. (2019a), "Partial least squares structural equation modeling-based discrete choice modeling: an illustration in modeling retailer choice", *Business Research*, Vol. 12, pp. 115-142.

Hair, J.F., Risher, J.J., Sarstedt, M. and Ringle, C.M. (2019b), "When to use and how to report the results of PLS-SEM", *European Business Review*, Vol. 31, pp. 2-24.

Hair, J.F., Risher, J.J., Sarstedt, M. and Ringle, C.M. (2019c), "When to use and how to report the results of PLS-SEM", *European Business Review*, Vol. 3, pp. 2-14.

Hair, J.F., Sarstedt, M. and Ringle, C.M. (2019d), "Rethinking some of the rethinking of partial least", *European Journal of Marketing*, Vol. 53, pp. 566-584.

Hair, J.F. Jr, Howard, M.C. and Nitzl, C. (2020), "Assessing measurement model quality in PLS-SEM using confirmatory composite analysis", *Journal of Business Research*, Vol. 109, pp. 101-110.

Han, S. and Yang, H. (2018), "Understanding adoption of intelligent personal assistants: a parasocial relationship perspective", *Industrial Management and Data Systems*, Vol. 118 No. 3, pp. 618-636.

Henseler, J. and Sarstedt, M. (2013), "Goodness-of-fit indices for partial least squares path modeling", *Computational Statistics*, Vol. 28, pp. 565-580.

Henseler, J., Dijkstra, T.K., Sarstedt, M., Ringle, C.M., Diamantopoulos, A., Straub, D.W., Ketchen, D.J., Jr, Hair, J.F., Hult, G.T.M. and Calantone, R.J. (2014), "Common beliefs and reality about PLS: comments on Rönkkö and Evermann (2013)", *Organizational Research Methods*, Vol. 17, pp. 182-209.

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

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

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

Hew, J.J., Lee, V.H., Ooi, K.B. and Wei, J. (2015), "What catalyses mobile apps usage intention: an empirical analysis", *Industrial Management and Data Systems*, Vol. 115 No. 7, pp. 1269-1291.

Ho, C.T. and Wei, C.L. (2016), "Effects of outsourced service providers' experiences on perceived service quality: a signaling theory framework", *Industrial Management and Data Systems*, Vol. 116 No. 8, pp. 1656-1677.

Hsieh, C.T., Lai, F. and Shi, W. (2006), "Information orientation and its impacts on information asymmetry and e-business adoption: evidence from China's international trading industry", *Industrial Management and Data Systems*, Vol. 106 No. 6, pp. 825-840.

Hu, L.T. and Bentler, P.M. (1998), "Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification", *Psychological Methods*, Vol. 3, pp. 424-453.

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

Huang, L.C. and Shiau, W.L. (2017), "Factors affecting creativity in information system development: insights from a decomposition and PLS-MGA", *Industrial Management and Data Systems*, Vol. 117 No. 3, pp. 496-520.

Hulland, J., Baumgartner, H. and Smith, K.M. (2018), "Marketing survey research best practices: evidence and recommendations from a review of JAMS articles", *Journal of the Academy of Marketing Science*, Vol. 46, pp. 92-108.

Hwang, H., Ho, M.-H.R. and Lee, J. (2010), "Generalized structured component analysis with latent interactions", *Psychometrika*, Vol. 75, pp. 228-242.

Hwang, H., Sarstedt, M., Cheah, J.H. and Ringle, C.M. (2020), "A concept analysis of methodological research on composite-based structural equation modeling: bridging PLSPM and GSCA", *Behaviormetrika*, Vol. 47, pp. 219-241.

Igbaria, M., Zinatelli, N., Cragg, P. and Cavaye, A.L. (1997), "Personal computing acceptance factors in small firms: a structural equation model", *MIS Quarterly*, Vol. 21 No. 3, pp. 279-305.

Ilmudeen, A. and Bao, Y. (2018), "Mediating role of managing information technology and its impact on firm performance: Insight from China", *Industrial Management and Data Systems*, Vol. 118 No. 4, pp. 912-929.

Jiang, Y. and Zhao, J. (2014), "Co-creating business value of information technology", *Industrial Management and Data Systems*, Vol. 114 No. 1, pp. 53-69.

Jöreskog, K.G. (1973), "A general method for estimating a linear structural equation system", in Goldberger, A.S. and Duncan, O.D. (Eds), *Structural Equation Models in the Social Sciences*, Seminar Press, New York, NJ, pp. 255-284.

Jöreskog, K.G. and Wold, H.O.A. (1982), *The ML and PLS Techniques for Modeling with Latent Variables: Historical and Comparative Aspects*, North-Holland, Amsterdam.

Kadipasaoglu, S.N., Peixoto, J.L. and Khumawala, B.M. (1999), "Global manufacturing practices: an empirical evaluation", *Industrial Management and Data Systems*, Vol. 99 No. 3, pp. 101-108.

Kemény, I., Simon, J., Nagy, Á. and Szucs, K. (2016), "Measuring quality perception in electronic commerce: a possible segmentation in the Hungarian market", *Industrial Management and Data Systems*, Vol. 116 No. 9, pp. 1946-1966.

Khan, G.F., Sarstedt, M., Shiau, W.-L., Hair, J.F., Ringle, C.M. and Fritze, M. (2019), "Methodological research on partial least squares structural equation modeling (PLS-SEM): an analysis based on social network approaches", *Internet Research*, Vol. 29, pp. 407-429.

Ko, D.G., Kirsch, L.J. and King, W.R. (2005), "Antecedents of knowledge transfer from consultants to clients in enterprise system implementations", *MIS Quarterly*, Vol. 29 No. 1, pp. 59-85.

Kock, N. and Hadaya, P. (2018), "Minimum sample size estimation in PLS-SEM: the inverse square root and gamma-exponential methods", *Information Systems Journal*, Vol. 28, pp. 227-261.

Komiak, S.Y. and Benbasat, I. (2006), "The effects of personalization and familiarity on trust and adoption of recommendation agents", *MIS Quarterly*, Vol. 30 No. 4, pp. 941-960.

Koohang, A., Paliszkiewicz, J. and Goluchowski, J. (2017), "The impact of leadership on trust, knowledge management, and organizational performance: a research model", *Industrial Management and Data Systems*, Vol. 117 No. 3, pp. 521-537.

Koohang, A., Paliszkiewicz, J. and Goluchowski, J. (2018), "Social media privacy concerns: trusting beliefs and risk beliefs", *Industrial Management and Data Systems*, Vol. 118 No. 6, pp. 1209-1228.

Lee, G. and Xia, W. (2010), "Toward agile: An integrated analysis of quantitative and qualitative field data on software development agility", *Mis Quarterly*, Vol. 34 No. 1, pp. 87-114.

Lee, Y., Kozar, K.A. and Larsen, K.R. (2003), "The technology acceptance model: past, present, and future", *Communications of the Association for Information Systems*, Vol. 12, p. 50.

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

Lee, A.B.S., Chan, F.T. and Pu, X. (2018), "Impact of supplier development on supplier's performance", *Industrial Management and Data Systems*, Vol. 118 No. 6, pp. 1192-1208.

Leon, S. (2018), "Service mobile apps: a millennial generation perspective", *Industrial Management and Data Systems*, Vol. 118 No. 9, pp. 1837-1860.

Li, M., Wang, Z. and Zhao, X. (2018), "The role of indigenous technological capability and interpersonal trust in supply chain learning", *Industrial Management and Data Systems*, Vol. 118 No. 5, pp. 1052-1070.

Liang, H., Saraf, N., Hu, Q. and Xue, Y. (2007), "Assimilation of enterprise systems: the effect of institutional pressures and the mediating role of top management", *MIS Quarterly*, Vol. 31 No. 1, pp. 59-87.

Liengaard, B., Sharma, P.N., Hult, T.M.H., Jensen, M.B., SarstedtHair, M.J.F. and Ringle, C.M. (2020), *Prediction: Coveted, yet Forsaken? Introducing a Cross-Validated Predictive Ability Test in Partial Least Squares Path Modeling*, Decision Sciences, in press.

Lin, C. and Lekhawipat, W. (2014), "Factors affecting online repurchase intention", *Industrial Management and Data Systems*, Vol. 114 No. 4, pp. 597-611.

Lin, C.Y., Chao, Y.C. and Tang, T.W. (2017), "Why not be "smarter"? Examining the factors that influence the behavioral intentions of non-smartphone users", *Industrial Management and Data Systems*, Vol. 117 No. 1, pp. 32-49.

Lin, H.F. (2014), "Contextual factors affecting knowledge management diffusion in SMEs", *Industrial Management and Data Systems*, Vol. 114 No. 9, pp. 1415-1437.

Lissitsa, S. and Kol, O. (2016), "Generation X vs. Generation Y–A decade of online shopping", *Journal of Retailing and Consumer Services*, Vol. 31, pp. 304-312.

Liu, Y. and Li, H. (2010), "Mobile internet diffusion in China: an empirical study", *Industrial Management and Data Systems*, Vol. 110 No. 3, pp. 309-324.

Liu, T.C., Chen, J.V. and Hsu, P.Y. (2014), "Opportunity recognition and cooperation flexibility of entrepreneurial franchisees", *Industrial Management and Data Systems*, Vol. 114 No. 4, pp. 506-525.

Liu, S., Yang, Y., Qu, W.G. and Liu, Y. (2016), "The business value of cloud computing: the partnering agility perspective", *Industrial Management and Data Systems*, Vol. 116 No. 6, pp. 1160-1177.

Lohmöller, J.-B. (1989), *Latent Variable Path Modeling with Partial Least Squares*, Springer-Verlag, New York.

Majchrzak, A., Beath, C.M., Lim, R.A. and Chin, W.W. (2005), "Managing client dialogues during information systems design to facilitate client learning", *MIS Quarterly*, Vol. 29 No. 4, pp. 653-672.

Majchrzak, A., Wagner, C. and Yates, D. (2013), "The impact of shaping on knowledge reuse for organizational improvement with wikis", *MIS Quarterly*, Vol. 37 No. 2, pp. 455-469.

Mandal, P. and Bagchi, K. (2016), "Strategic role of information, knowledge and technology in manufacturing industry performance", *Industrial Management and Data Systems*, Vol. 116 No. 6, pp. 1259-1278.

Martínez-Martínez, D., Herrera Madueno, J., Larran Jorge, M. and Lechuga Sancho, M.P. (2017), "The strategic nature of corporate social responsibility in SMEs: a multiple mediator analysis", *Industrial Management and Data Systems*, Vol. 117 No. 1, pp. 2-31.

Mason, C.H. and Perreault, W.D. Jr (1991), "Collinearity, power, and interpretation of multiple regression analysis", *Journal of Marketing Research*, Vol. 28, pp. 268-280.

McQuarrie, A.D. and Tsai, C.-L. (1998), *Regression and Time Series Model Selection*, World Scientific, Singapore, Vol. 43.

Medlin, B., Green, K.W. and Wright, A.D. (2016), "Comprehensive management practices and policies performance model", *Industrial Management and Data Systems*, Vol. 116 No. 5, pp. 1043-1060.

Meena, P.L. and Sarmah, S.P. (2012), "Development of a supplier satisfaction index model", *Industrial Management and Data Systems*, Vol. 112 No. 8, pp. 1236-1254.

Molinillo, S., Muñoz-Leiva, F. and Pérez-García, F. (2018), "The effects of human-game interaction, network externalities, and motivations on players' use of mobile casual games", *Industrial Management and Data Systems*, Vol. 118 No. 9, pp. 1766-1786.

Moneva, J.M. and Ortas, E. (2010), "Corporate environmental and financial performance: a multivariate approach", *Industrial Management and Data Systems*, Vol. 110 No. 2, pp. 193-210.

Moores, T.T. and Chang, J.C.J. (2006), "Ethical decision making in software piracy: initial development and test of a four-component model", *MIS Quarterly*, Vol. 30 No. 1, pp. 167-180.

Myung, I.J. (2000), "The importance of complexity in model selection", *Journal of Mathematical Psychology*, Vol. 44, pp. 190-204.

Nancy Vargas, M. and Begoña Lloria, M. (2014), "Dynamizing intellectual capital through enablers and learning flows", *Industrial Management and Data Systems*, Vol. 114 No. 1, pp. 2-20.

Nfuka, E.N. and Rusu, L. (2011), "The effect of critical success factors on IT governance performance", *Industrial Management and Data Systems*, Vol. 111 No. 9, pp. 1418-1448.

Nitzl, C., Roldán, J.L. and Cepeda Carrión, G. (2016), "Mediation analysis in partial least squares path modeling: helping researchers discuss more sophisticated models", *Industrial Management and Data Systems*, Vol. 119, pp. 1849-1864.

Omigie, N.O., Zo, H., Rho, J.J. and Ciganek, A.P. (2017), "Customer pre-adoption choice behavior for M-PESA mobile financial services: extending the theory of consumption values", *Industrial Management and Data Systems*, Vol. 117 No. 5, pp. 910-926.

Park, J.G., Park, K. and Lee, J. (2014), "A firm's post-adoption behavior: loyalty or switching costs?", *Industrial Management and Data Systems*, Vol. 114 No. 2, pp. 258-275.

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

Petter, S. (2018), "'Haters gonna hate': PLS and information systems research", *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, Vol. 49, pp. 10-13.

Pitt, M.A. and Myung, I.J. (2002), "When a good fit can be bad", *Trends in Cognitive Sciences*, Vol. 6, pp. 421-425.

Prasad, S., Garg, A. and Prasad, S. (2019), "Purchase decision of generation Y in an online environment", *Marketing Intelligence and Planning*, Vol. 37, pp. 372-385.

Preacher, K.J. and Merkle, E.C. (2012), "The problem of model selection uncertainty in structural equation modeling", *Psychological Methods*, Vol. 17, p. 1.

Prentice, C., Han, X.-Y. and Li, Y.-Q. (2016), "Customer empowerment to co-create service designs and delivery: scale development and validation", *Services Marketing Quarterly*, Vol. 37, pp. 36-51.

Puklavec, B., Oliveira, T. and Popovič, A. (2018), “Understanding the determinants of business intelligence system adoption stages: an empirical study of SMEs”, *Industrial Management and Data Systems*, Vol. 118 No. 1, pp. 236-261.

Ram, J., Corkindale, D. and Wu, M.L. (2013), “Examining the role of system quality in ERP projects”, *Industrial Management and Data Systems*, Vol. 113 No. 3, pp. 350-366.

Ravichandran, T. and Rai, A. (2000), “Quality management in systems development: an organizational system perspective”, *MIS Quarterly*, Vol. 24 No. 3, pp. 381-415.

R Core Team (2019), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, available at: https://www.R-project.org.

Richter, N.F., Sinkovics, R.R., Ringle, C.M. and Schlaegel, C. (2016), “A critical look at the use of SEM in international business research”, *International Marketing Review*, Vol. 33, pp. 376-404.

Rigdon, E.E., Sarstedt, M. and Ringle, C.M. (2017), “On comparing results from CB-SEM and PLS-SEM. Five perspectives and five recommendations”, *Marketing ZFP*, Vol. 39, pp. 4-16.

Rigdon, E.E. (2012), “Rethinking partial least squares path modeling: in praise of simple methods”, *Long Range Planning*, Vol. 45, pp. 341-358.

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

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

Ringle, C.M., Wende, S. and Becker, J.-M. (2015), *SmartPLS 3*, SmartPLS, Bönningstedt.

Ringle, C.M., Sarstedt, M., Mitchell, R. and Gudergan, S.P. (2020), “Partial least squares structural equation modeling in HRM research”, *The International Journal of Human Resource Management*, Vol. 31 No. 12, pp. 1617-1643.

Roldán, J.L. and Sánchez-Franco, M.J. (2012), “Variance-based structural equation modeling: guidelines for using partial least squares in information systems research”, in Mora, M., Gelman, O., Steenkamp, A. and Raisinghani, M. (Eds), *Research Methodologies, Innovations and Philosophies in Software Systems Engineering and Information Systems*, IGI Global, Hershey, pp. 193-221.

Rönkkö, M. and Evermann, J. (2013), “A critical examination of common beliefs about partial least squares path modeling”, *Organizational Research Methods*, Vol. 16, pp. 425-448.

Roemer, E. (2016), “A tutorial on the use of PLS path modeling in longitudinal studies”, *Industrial Management and Data Systems*, Vol. 116, pp. 1901-1921.

Ruddock, R. (2017), “Statistical significance: why it often doesn’t mean much to marketers”, available at https://medium.com/@RonRuddock/statistical-significance-why-it-often-doesnt-mean-much-to-marketers-d5bec3e1ed (accessed 4 April 2019).

Ryu, H.S. (2018), “What makes users willing or hesitant to use Fintech?: the moderating effect of user type”, *Industrial Management and Data Systems*, Vol. 118 No. 3, pp. 541-569.

Salleh, N.A.M., Jusoh, R. and Isa, C.R. (2010), “Relationship between information systems sophistication and performance measurement”, *Industrial Management and Data Systems*, Vol. 110 No. 7, pp. 993-1017.

Sarker, S. and Valacich, J.S. (2010), “An alternative to methodological individualism: a non-reductionist approach to studying technology adoption by groups”, *MIS Quarterly*, Vol. 34 No. 4, pp. 779-808.

Sarstedt, M. and Mooi, E.A. (2019), *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics*, Springer, Heidelberg *et al.*

Sarstedt, M., Wilczynski, P. and Melewar, T. (2013), “Measuring reputation in global markets—a comparison of reputation measures’ convergent and criterion validities”, *Journal of World Business*, Vol. 48, pp. 329-339.

Sarstedt, M., Hair, J.F., Ringle, C.M., Thiele, K.O. and Gudergan, S.P. (2016), "Estimation issues with PLS and CBSEM: where the bias lies!", *Journal of Business Research*, Vol. 69, pp. 3998-4010.

Sarstedt, M., Ringle, C.M. and Hair, J.F. (2017), "Partial least squares structural equation modeling", in Homburg, C., Klarmann, M. and Vomberg, A. (Eds), *Handbook of Market Research*, Springer, Heidelberg.

Sarstedt, M., Bengart, P., Shaltoni, A.M. and Lehmann, S. (2018), "The use of sampling methods in advertising research: a gap between theory and practice", *International Journal of Advertising*, Vol. 37, pp. 650-663.

Sarstedt, M., Hair, J.F. Jr, Cheah, J.-H., Becker, J.-M. and Ringle, C.M. (2019), "How to specify, estimate, and validate higher-order constructs in PLS-SEM", *Australasian Marketing Journal (AMJ)*, Vol. 27, pp. 197-211.

Schwarz, G. (1978), "Estimating the dimension of a model", *Annals of Statistics*, Vol. 6, pp. 461-464.

Shang, D. and Wu, W. (2017), "Understanding mobile shopping consumers' continuance intention", *Industrial Management and Data Systems*, Vol. 117 No. 1, pp. 213-227.

Shao, Z., Wang, T. and Feng, Y. (2016), "Impact of chief information officer's strategic knowledge and structural power on enterprise systems success", *Industrial Management and Data Systems*, Vol. 116 No. 1, pp. 43-64.

Sharma, H. and Lowalekar, V. (2017), "Vehicle ownership as an influence on retail shoppers' perception of Indian traditional grocery stores", *International Journal of Marketing and Business Communication*, Vol. 6.

Sharma, P., Sarstedt, M., Shmueli, G., Kim, K.H. and Thiele, K.O. (2019a), "PLS-based model selection: the role of alternative explanations in information systems research", *Journal of the Association for Information Systems*, Vol. 20, p. 4.

Sharma, P.N., Shmueli, G., Sarstedt, M., Danks, N. and Ray, S. (2019b), "Prediction-oriented model selection in partial least squares path modeling", *Decision Sciences*, Forthcoming.

Shi, S. and Chow, W.S. (2015), "Trust development and transfer in social commerce: prior experience as moderator", *Industrial Management and Data Systems*, Vol. 115 No. 7, pp. 1182-1203.

Shiau, W.-L. and Chau, P.Y. (2012), "Understanding blog continuance: a model comparison approach", *Industrial Management and Data Systems*, Vol. 112, pp. 663-682.

Shiau, W.-L. and Chau, P.Y. (2016), "Understanding behavioral intention to use a cloud computing classroom: a multiple model comparison approach", *Information and Management*, Vol. 53, pp. 355-365.

Shiau, W.-L., Sarstedt, M. and Hair, J.F. (2019), "Internet research using partial least squares structural equation modeling (PLS-SEM)", *Internet Research*, Vol. 29, pp. 398-406.

Shmueli, G. and Koppius, O.R. (2011), "Predictive analytics in information systems research", *MIS Quarterly*, Vol. 35, pp. 553-572.

Shmueli, G., Ray, S., Velasquez Estrada, J.M. and Chatla, S.B. (2016), "The elephant in the room: evaluating the predictive performance of PLS models", *Journal of Business Research*, Vol. 69, pp. 4552-4564.

Shmueli, G., Sarstedt, M., Hair, J.F., Cheah, J.-H., Ting, H., Vaithilingam, S. and Ringle, C.M. (2019), " Predictive model assessment in PLS-SEM: guidelines for using PLSpredict", *European Journal of Marketing*, Vol. 53, pp. 2322-2347.

Shmueli, G. (2010), "To explain or to predict?", *Statistical Science*, Vol. 25, pp. 289-310.

Song, H., Yu, K., Ganguly, A. and Turson, R. (2016), "Supply chain network, information sharing and SME credit quality", *Industrial Management and Data Systems*, Vol. 116 No. 4, pp. 740-758.

Sternad, S., Gradisar, M. and Bobek, S. (2011), "The influence of external factors on routine ERP usage", *Industrial Management and Data Systems*, Vol. 111 No. 9, pp. 1511-1530.

Stone, M. (1974), "Cross-validatory choice and assessment of statistical predictions", *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 36, pp. 111-133.

Streukens, S. and Leroi-Werelds, S. (2016), "PLS FAC-SEM: an illustrated step-by-step guideline to obtain a unique insight in factorial data", *Industrial Management and Data Systems*, Vol. 116, pp. 1922-1945.

Suen, H.Y. (2018), "The effects of employer SNS motioning on employee perceived privacy violation, procedural justice, and leave intention", *Industrial Management and Data Systems*, Vol. 118 No. 6, pp. 1153-1169.

Swaim, J.A., Maloni, M., Bower, P. and Mello, J. (2016), "Antecedents to effective sales and operations planning", *Industrial Management and Data Systems*, Vol. 116 No. 6, pp. 1279-1294.

Symonds, M.R. and Moussalli, A. (2011), "A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion", *Behavioral Ecology and Sociobiology*, Vol. 65, pp. 13-21.

Tai, Y.M. (2011), "Perceived value for customers in information sharing services", *Industrial Management and Data Systems*, Vol. 111 No. 4, pp. 551-569.

Tenenhaus, A. and Tenenhaus, M. (2011), "Regularized generalized canonical correlation analysis", *Psychometrika*, Vol. 76, pp. 257-284.

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

Teo, H.H., Wei, K.K. and Benbasat, I. (2003), "Predicting intention to adopt interorganizational linkages: an institutional perspective", *MIS Quarterly*, Vol. 27 No. 1, pp. 19-49.

Teo, A.C., Tan, G.W.H., Ooi, K.B., Hew, T.S. and Yew, K.T. (2015), "The effects of convenience and speed in m-payment", *Industrial Management and Data Systems*, Vol. 115 No. 2, pp. 311-331.

Thompson, R.L., Higgins, C.A. and Howell, J.M. (1991), "Personal computing: toward a conceptual model of utilization", *MIS Quarterly*, Vol. 15 No. 1, pp. 125-143.

Trunk Širca, N., Babnik, K. and Breznik, K. (2013), "Towards organisational performance: understanding human resource management climate", *Industrial Management and Data Systems*, Vol. 113 No. 3, pp. 367-384.

Tsao, W.C., Hsieh, M.T. and Lin, T.M. (2016), "Intensifying online loyalty! The power of website quality and the perceived value of consumer/seller relationship", *Industrial Management and Data Systems*, Vol. 116 No. 9, pp. 1987-2010.

Türkyılmaz, A. and Özkan, C. (2007), "Development of a customer satisfaction index model: An application to the Turkish mobile phone sector", *Industrial Management and Data Systems*, Vol. 107 No. 5, pp. 672-687.

Turkyilmaz, A., Akman, G., Ozkan, C. and Pastuszak, Z. (2011), "Empirical study of public sector employee loyalty and satisfaction", *Industrial Management and Data Systems*, Vol. 111 No. 5, pp. 675-696.

Wagenmakers, E.-J. and Farrell, S. (2004), "AIC model selection using Akaike weights", *Psychonomic Bulletin and Review*, Vol. 11, pp. 192-196.

Wang, J. and Dai, J. (2018), "Sustainable supply chain management practices and performance", *Industrial Management and Data Systems*, Vol. 118 No. 1, pp. 2-21.

Wang, E.T., Tai, J.C. and Grover, V. (2013), "Examining the relational benefits of improved interfirm information processing capability in buyer-supplier dyads", *MIS Quarterly*, Vol. 37 No. 1, pp. 149-173.

Wang, F., Zhao, J., Chi, M. and Li, Y. (2017), "Collaborative innovation capability in IT-enabled inter-firm collaboration", *Industrial Management and Data Systems*, Vol. 117 No. 10, pp. 2364-2380.

Watanuki, H.M. and Moraes, R.D.O. (2016), "Does size matter? An investigation into the role of virtual team size in IT service provisioning", *Industrial Management and Data Systems*, Vol. 116 No. 9, pp. 1967-1986.

Witten, I.H., Frank, E., Hall, M.A. and Pal, C.J. (2016), *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, Burlington, MA.

Wixom, B.H. and Watson, H.J. (2001), "An empirical investigation of the factors affecting data warehousing success", *MIS Quarterly*, Vol. 25 No. 1, pp. 17-41.

Wold, H. (1982), "Soft modeling: the basic design and some extensions", in Jöreskog, K.G.H.W.E. (Ed.), *Systems under Indirect Observation: Causality, Structure, Prediction, Part II*, North Holland, Amsterdam.

Wold, H.O.A. (1985), "Partial least squares", in Kotz, S. and Johnson, N.L. (Eds), *Encyclopedia of Statistical Sciences*, Wiley, New York.

Wong, C.H., Tan, G.W.H., Loke, S.P. and Ooi, K.B. (2014), "Mobile TV: a new form of entertainment?", *Industrial Management and Data Systems*, Vol. 114 No. 7, pp. 1050-1067.

Wu, J.J., Kung, H.Y. and Lin, T.M. (2017), "Influence of customer participation on information technology services", *Industrial Management and Data Systems*, Vol. 117 No. 6, pp. 1077-1092.

Xu, D., Huo, B. and Sun, L. (2014), "Relationships between intra-organizational resources, supply chain integration and business performance: an extended resource-based view", *Industrial Management and Data Systems*, Vol. 114 No. 8, pp. 1186-1206.

Yang, H., Lee, H. and Zo, H. (2017), "User acceptance of smart home services: an extension of the theory of planned behavior", *Industrial Management and Data Systems*, Vol. 117 No. 1, pp. 68-89.

Yim, B. and Leem, B. (2013), "The effect of the supply chain social capital", *Industrial Management & Data Systems*, Vol. 113 No. 3, pp. 324-349.

Yu, P.L., Balaji, M.S. and Khong, K.W. (2015), "Building trust in internet banking: a trustworthiness perspective", *Industrial Management and Data Systems*, Vol. 115 No. 2, pp. 235-252.

Zhang, H. and Yang, F. (2016), "On the drivers and performance outcomes of green practices adoption: an empirical study in China", *Industrial Management and Data Systems*, Vol. 116 No. 9, pp. 2011-2034.

Zhang, S., Wang, Z., Zhao, X. and Zhang, M. (2017), "Effects of institutional support on innovation and performance: roles of dysfunctional competition", *Industrial Management and Data Systems*, Vol. 117 No. 1, pp. 50-67.

Zhang, M., Ren, C., Wang, G.A. and He, Z. (2018), "The impact of channel integration on consumer responses in omni-channel retailing: the mediating effect of consumer empowerment", *Electronic Commerce Research and Applications*, Vol. 28, pp. 181-193.

Zhao, X., Tang, Q., Liu, S. and Liu, F. (2016), "Social capital, motivations, and mobile coupon sharing", *Industrial Management and Data Systems*, Vol. 116 No. 1, pp. 188-206.

Zheng, Z. and Pavlou, P.A. (2010), "Research note—toward a causal interpretation from observational data: a new Bayesian networks method for structural models with latent variables", *Information Systems Research*, Vol. 21, pp. 365-391.

Zhou, W., Tsiga, Z., Li, B., Zheng, S. and Jiang, S. (2018), "What influence users'e-finance continuance intention? The moderating role of trust", *Industrial Management and Data Systems*, Vol. 118 No. 8, pp. 1647-1670.

Appendix

Table A1. Empirical studies employing the use of causal prediction criteria of PLS-SEM in IMDS

No	Empirical study	Research topic	Observation of previous causal prediction criteria: GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
1	Ryu (2018)	What makes users willing or hesitant to use Fintech?: the moderating effect of user type		√						
2	Fan *et al.* (2018)	Understanding users' attitude toward mobile payment use: A comparative study between China and the USA		√						
3	Puklavec *et al.* (2018)	Understanding the determinants of business intelligence system adoption stages: An empirical study of SMEs		√						
4	Han and Yang (2018)	Understanding adoption of intelligent personal assistants: A parasocial relationship perspective		√						
5	Ilmudeen and Bao (2018)	Mediating role of managing information technology and its impact on firm performance Insight from China		√						
6	Suen (2018)	The effects of employer SNS motioning on employee perceived privacy violation, procedural justice, and leave intention		√	√					

(continued)

No	Empirical study	Research topic	Observation of previous causal prediction criteria: GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
7	Koohang *et al.* (2018)	Social media privacy concerns: trusting beliefs and risk beliefs		√						
8	Lee *et al.* (2018)	Impact of supplier development on supplier's performance		√	√					
9	Alonso Dos Santos *et al.* (2018)	Sponsorship image transfer theory in virtual brand communities		√	√				√	
10	Zhou *et al.* (2018)	What influence users' e-finance continuance intention? The moderating role of trust		√						
11	Leon (2018)	Service mobile apps: a millennial generation perspective		√						
12	Wang and Dai (2018)	Sustainable supply chain management practices and performance		√						
13	Chuah *et al.* (2018)	Reducing temptation to switch mobile data service providers over time The role of dedication vs constraint		√						
14	Molinillo *et al.* (2018)	The effects of human-game interaction, network externalities, and motivations on players' use of mobile casual games		√	√					

(*continued*)

Table A1.

No	Empirical study	Research topic	Observation of previous causal prediction criteria							
			GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
15	Li *et al.* (2018)	The role of indigenous technological capability and interpersonal trust in supply chain learning		√						
16	Ding (2018)	I hope and I continue Integrating the concept of hope into the expectancy-disconfirmation framework	√	√						
17	Shang and Wu (2017)	Understanding mobile shopping consumers' continuance intention		√						
18	Yang *et al.* (2017)	User acceptance of smart home services: an extension of the theory of planned behavior		√						
19	Andrei *et al.* (2017)	Word-of-mouth generation and brand communication strategy Findings from an experimental study explored with PLS-SEM		√					√	√
20	Koohang *et al.* (2017)	The impact of leadership on trust, knowledge management, and organizational performance A research model		√						

(continued)

Table A1.

Table A1.

No	Empirical study	Research topic	Observation of previous causal prediction criteria: GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
21	Martinez-Martinez *et al.* (2017)	The strategic nature of corporate social responsibility in SMEs: a multiple mediator analysis		√	√				√	√
22	Zhang *et al.* (2017)	Effects of institutional support on innovation and performance: roles of dysfunctional competition		√					√	
23	Dalvi-Esfahani *et al.* (2017)	Moderating role of personal values on managers' intention to adopt Green IS Examining norm activation theory		√						
24	Lin *et al.* (2017)	Why not be smarter? Examining the factors that influence the behavioral intentions of non-smartphone users		√					√	
25	Omigie *et al.* (2017)	Customer pre-adoption choice behavior for M-PESA mobile financial services Extending the theory of consumption values		√						
26	Ali *et al.* (2017)	Food supply chain integrity: the need to go beyond certification		√						

(*continued*)

No	Empirical study	Research topic	Observation of previous causal prediction criteria GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
27	Wang *et al.* (2017)	Collaborative innovation capability in IT-enabled inter-firm collaboration	√	√						
28	Huang and Shiau (2017)	Factors affecting creativity in information system development Insights from a decomposition and PLS-MGA		√						
29	Wu *et al.* (2017)	Influence of customer participation on information technology services		√						
30	Zhang and Yang (2016)	On the drivers and performance outcomes of green practices adoption: An empirical study in China		√	√				√	√
31	Zhao *et al.* (2016)	Social capital, motivations and mobile coupon sharing		√						
32	Bao *et al.* (2016)	Repurchase intention in the Chinese e-marketplace Roles of interactivity, trust and perceived effectiveness of e-commerce institutional mechanisms		√	√					
33	Tsao *et al.* (2016)	Intensifying online loyalty! The power of website quality and the perceived value of consumer/seller relationship		√					√	√

(continued)

Table A1.

Table A1.

No	Empirical study	Research topic	Observation of previous causal prediction criteria GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
34	Kemény *et al.* (2016)	Measuring quality perception in electronic commerce A possible segmentation in the Hungarian market		√					√	
35	Swaim *et al.* (2016)	Antecedents to effective sales and operations planning		√	√					
36	Shao *et al.* (2016)	Impact of chief information officer's strategic knowledge and structural power on enterprise systems success		√						
37	Song *et al.* (2016)	Supply chain network, information sharing and SME credit quality		√					√	
38	Ho and Wei (2016)	Effects of outsourced service providers' experiences on perceived service quality: A signaling theory framework		√	√					
39	Mandal and Bagchi (2016)	Strategic role of information, knowledge and technology in manufacturing industry performance		√						
40	Liu *et al.* (2016)	The business value of cloud computing: the partnering agility perspective		√						
41	Watanuki and Moraes (2016)	Does size matter? An investigation into the role of virtual team size in IT service provisioning		√					√	

(continued)

Table A1.

No	Empirical study	Research topic	Observation of previous causal prediction criteria							
			GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
42	Medlin *et al.* (2016)	Comprehensive management practices and policies performance model		√						
43	Yu *et al.* (2015)	Building trust in Internet banking: a trustworthiness perspective		√						
44	Hew *et al.* (2015)	What catalyses mobile apps usage intention: an empirical analysis	√	√	√					
45	Teo *et al.* (2015)	The effects of convenience and speed in m-payment		√						
46	Ainin *et al.* (2015)	Factors influencing the use of social media by SMEs and its performance outcomes		√						
47	Shi and Chow (2015)	Trust development and transfer in social commerce: prior experience as moderator		√						
48	Gómez-Cedeño *et al.* (2015)	Impact of human resources on supply chain management and performance	√	√	√					
49	Chen and Qi (2015)	Members' satisfaction and continuance intention: a socio-technical perspective		√						
50	Chi *et al.* (2015)	Mediation and time-lag analyses of e-alignment and e-collaboration capabilities		√						

(*continued*)

Table A1.

No	Empirical study	Research topic	GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
			Observation of previous causal prediction criteria							
51	Gonzalez *et al.* (2015)	Information systems outsourcing satisfaction: some explanatory factors		√	√					
52	Xu *et al.* (2014)	Relationships between intra-organizational resources, supply chain integration and business performance		√						
53	Lin and Lekhawipat (2014)	Factors affecting online repurchase intention		√						
54	Wong *et al.* (2014)	Mobile TV: a new form of entertainment?		√						
55	Lin (2014)	Contextual factors affecting knowledge management diffusion in SMEs		√						
56	Park *et al.* (2014)	A firm's post-adoption behavior: loyalty or switching costs?		√	√					
57	Clark *et al.* (2014)	Market-oriented sustainability: moderating impact of stakeholder involvement		√						
58	Jiang and Zhao (2014)	Co-creating business value of information technology		√						
59	Nancy Vargas and Begoñ a Lloria (2014)	Dynamizing intellectual capital through enablers and learning flows	√	√	√					

(*continued*)

No	Empirical study	Research topic	Observation of previous causal prediction criteria: GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
60	Liu *et al.* (2014)	Opportunity recognition and cooperation flexibility of entrepreneurial franchisees		√						
61	Chen *et al.* (2013)	Process quality and collaboration quality on B2B e-commerce		√						
62	Yim and Leem (2013)	The effect of the supply chain social capital		√						
63	Trunk Širca *et al.* (2013)	Towards organisational performance Understanding human resource management climate	√	√						
64	Ram *et al.* (2013)	Examining the role of system quality in ERP projects	√	√						
65	Meena and Sarmah (2012)	Development of a supplier satisfaction index model	√	√						
66	Chan and Wang (2012)	The causal relationships between aspects of customer capital		√						
67	Turkyilmaz *et al.* (2011)	Empirical study of public sector employee loyalty and satisfaction		√						
68	Nfuka and Rusu (2011)	The effect of critical success factors on IT governance performance		√						
69	Sternad *et al.* (2011)	The influence of external factors on routine ERP usage	√	√						

(*continued*)

Table A1.

No	Empirical study	Research topic	Observation of previous causal prediction criteria GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
70	Tai (2011)	Perceived value for customers in information sharing services		√						
71	Gil-Saura *et al.* (2011)	Value, supplier dependence and long-term orientation Outcomes for B2B commerce in the travel industry		√						
72	Moneva and Ortas (2010)	Corporate environmental and financial performance: a multivariate approach		√						
73	Liu and Li (2010)	Mobile Internet diffusion in China: an empirical study	√	√						
74	Salleh *et al.* (2010)	Relationship between information systems sophistication and performance measurement		√						
75	Bryan Foltz *et al.* (2008)	Why users (fail to) read computer usage policies		√						
76	Türkyılmaz and Özkan (2007)	Development of a customer satisfaction index model-An application to the Turkish mobile phone sector		√						
77	Hsieh *et al.* (2006)	Information orientation and its impacts on information asymmetry and e-business adoption-Evidence from China's international trading industry		√						

(continued)

Table A1.

No	Empirical study	Research topic	Observation of previous causal prediction criteria GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
78	Kadipasaoglu *et al.* (1999)	Global manufacturing practices: an empirical evaluation		√						
Overall usage of the criteria (%)			10/78 (12.82%)	78/78 (100.00%)	14/78 (17.95%)	0/78 (0%)	0/78 (0%)	0/78 (0%)	10/78 (12.82%)	4/78 (5.13%)

Note(s): Compilation by author and further reading of these articles are enclosed in the supplementary reference file

Table A1.

Table A2.
Empirical studies employing the use of causal prediction criteria of PLS-SEM in MISQ

No	Authors	Title	Observation of previous causal prediction Criteria: GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
1	Addas and Pinsonneault (2018)	E-mail interruptions and individual performance: is there a silver lining?		√						
2	Curtis *et al.* (2017)	From monologue to dialogue: performative objects to promote collective mindfulness in computer-mediated team discussions		√						
3	Fang *et al.* (2014)	Trust, satisfaction, and online repurchase intention: the moderating role of perceived effectiveness of e-commerce institutional mechanisms		√						
4	Majchrzak *et al.* (2013)	The impact of shaping on knowledge reuse for organizational improvement with wikis		√						
5	Wang *et al.* (2013)	Examining the relational benefits of improved interfirm information processing capability in buyer-supplier dyads		√						
6	Furneaux and Wade (2011)	An exploration of organizational level information systems discontinuance intentions		√						
7	Lee and Xia (2010)	Toward agile: an integrated analysis of quantitative and qualitative field data on software development agility		√						
8	Sarker and Valacich (2010)	An alternative to methodological individualism: a non-reductionist approach to studying technology adoption by groups		√						

(continued)

No	Authors	Title	GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
			Observation of previous causal prediction Criteria							
9	Goo *et al.* (2009)	The role of service level agreements in relational management of information technology outsourcing: an empirical study		√						
10	Au *et al.* (2008)	Extending the understanding of end-user information systems satisfaction formation: An equitable needs fulfillment model approach		√						
11	Liang *et al.* (2007)	Assimilation of enterprise systems: The effect of institutional pressures and the mediating role of top management		√						
12	Komiak and Benbasat (2006)	The effects of personalization and familiarity on trust and adoption of recommendation agents		√						
13	Moores and Chang (2006)	Ethical decision making in software piracy: Initial development and test of a four-component model		√						
14	Ko *et al.* (2005)	Antecedents of knowledge transfer from consultants to clients in enterprise system implementations		√						
15	Majchrzak *et al.* (2005)	Managing client dialogues during information systems design to facilitate client learning		√						
16	Teo *et al.* (2003)	Predicting intention to adopt interorganizational linkages: An institutional perspective		√						

(*continued*)

Table A2.

Table A2.

No	Authors	Title	Observation of previous causal prediction Criteria GoF Tenanhaus	R^2	Q^2	PLSpredict	CVPAT	Model selection criteria	SRMR	d_ULS and d_G
17	Enns *et al.* (2003)	CIO lateral influence behaviors: Gaining peers' commitment to strategic information systems		√						
18	Wixom and Watson (2001)	An empirical investigation of the factors affecting data warehousing success		√						
19	Ravichandran and Rai (2000)	Quality management in systems development: An organizational system perspective		√						
20	Compeau and Higgins (1995)	Computer self-efficacy – development of a measure and initial test		√						
21	Igbaria *et al.* (1997)	Personal computing acceptance factors in small firms: A structural equation model		√						
22	Thompson *et al.* (1991)	Personal computing-toward a conceptual-model of utilization		√						
Overall Usage of the Criteria (%)			0/22 (0%)	22/22 (100.00%)	0/22 (0%)	0/22 (0%)	0/22 (0%)	0/22 (0%)	0/22 (0%)	0/22 (0%)

Note(s): Compilation by author and further reading of these articles are enclosed in the supplementary reference file

Characteristic	Description	Frequency	Percent
Gender	Male	100	42.7
	Female	134	57.3
Ethnicity	Malay	96	41.0
	Chinese	87	37.2
	Indian	35	15.0
	Others	16	6.8
Level of Education	Bachelor Degree	72	30.8
	Master Degree	117	50.0
	PhD Degree	45	19.2
Occupation	Enterprise Employee	74	31.6
	Civil Servant	139	59.4
	Entrepreneur	21	9.0
Personal Income	Below RM 2,500	47	20.1
	RM 2,500 to RM 4,500	51	21.9
	RM 4,501 to RM 6,500	60	25.6
	RM 6,501 to RM 8,500	56	23.9
	Above RM 8,500	20	8.5

Note(s): $N = 234$

Table A3. Sample demographic

Construct	Item	Loading	CA	rho_A	CR	AVE
Customer Empowerment	CE1	0.889	0.922	0.924	0.941	0.762
	CE2	0.873				
	CE3	0.897				
	CE4	0.839				
	CE5	0.865				
Integrated Customer Service	ICS1	0.867	0.898	0.901	0.929	0.765
	ICS2	0.866				
	ICS3	0.870				
	ICS4	0.895				
Integrated Information Access	IIA1	0.715	0.868	0.891	0.903	0.653
	IIA2	0.775				
	IIA3	0.847				
	IIA4	0.826				
	IIA5	0.868				
Integrated Order Fulfillment	IOF1	0.723	0.912	0.916	0.932	0.697
	IOF2	0.900				
	IOF3	0.834				
	IOF4	0.864				
	IOF5	0.811				
	IOF6	0.865				
Integrated Promotion	IP1	0.664	0.846	0.855	0.891	0.662
	IP2	0.853				
	IP3	0.801				
	IP4	0.770				
	IP5	0.841				

(*continued*)

Table A4. Assessment of indicator reliability, average variance extracted and convergent validity

Construct	Item	Loading	CA	rho_A	CR	AVE
Integrated Product and Price	IPP1	0.844	0.901	0.913	0.926	0.715
	IPP2	0.839				
	IPP3	0.803				
	IPP4	0.849				
	IPP5	0.890				
Integrated Transaction Information	ITI1	0.780	0.858	0.883	0.896	0.635
	ITI2	0.857				
	ITI3	0.750				
	ITI4	0.709				
	ITI5	0.876				
Patronage Intention	Patron1	0.954	0.942	0.947	0.963	0.897
	Patron2	0.928				
	Patron3	0.959				
Satisfaction	Sat1	0.914	0.955	0.956	0.965	0.848
	Sat2	0.924				
	Sat3	0.938				
	Sat4	0.919				
	Sat5	0.910				
Trust	Trust1	0.906	0.914	0.922	0.936	0.747
	Trust2	0.938				
	Trust3	0.900				
	Trust4	0.769				
	Trust5	0.796				

Table A4.

Table A5. Assessment of Discriminant Validity using the Heterotrait–Monotrait (HTMT) Criterion

Construct	1	2	3	4	5	6	7	8	9	10
1. CE										
2. ICS	0.709 [0.623; 0.783]									
3. IIA	0.512 [0.368; 0.632]	0.670 [0.556; 0.765]								
4. IOF	0.743 [0.659; 0.807]	0.822 [0.756; 0.878]	0.716 [0.611; 0.797]							
5. IPP	0.590 [0.471; 0.691]	0.430 [0.292; 0.554]	0.467 [0.328; 0.595]	0.500 [0.375; 0.607]						
6. IP	0.706 [0.607; 0.786]	0.609 [0.494; 0.701]	0.525 [0.391; 0.653]	0.662 [0.561; 0.741]	0.863 [0.798; 0.895]					
7. ITI	0.463 [0.318; 0.589]	0.608 [0.485; 0.715]	0.688 [0.563; 0.791]	0.598 [0.464; 0.710]	0.441 [0.303; 0.570]	0.461 [0.329; 0.596]				
8. Patronage	0.633 [0.532; 0.712]	0.348 [0.219; 0.465]	0.263 [0.134; 0.407]	0.393 [0.270; 0.498]	0.633 [0.530; 0.718]	0.626 [0.513; 0.716]	0.279 [0.167; 0.393]			
9. Satisfaction	0.655 [0.554; 0.739]	0.367 [0.239; 0.476]	0.276 [0.145; 0.414]	0.430 [0.309; 0.533]	0.611 [0.516; 0.695]	0.594 [0.486; 0.683]	0.251 [0.146; 0.373]	0.887 [0.810; 0.892]		
10. Trust	0.773 [0.689; 0.843]	0.607 [0.502; 0.700]	0.444 [0.287; 0.582]	0.606 [0.494; 0.699]	0.593 [0.473; 0.694]	0.658 [0.548; 0.741]	0.350 [0.236; 0.469]	0.715 [0.601; 0.804]	0.844 [0.775; 0.894]	

Note(s): (1) CE = Customer Empowerment, ICS = Integrated Customer Service; IIA = Integrated Information Access; IOF = Integrated Order Fulfillment; IPP = Integrated Product and Price; IP = Integrated Promotion; ITI = Integrated Transaction Information, (2) All the HTMT values below HTMT.90 (Gold *et al.*, 2001), (3) Brackets show lower and upper bound of the 95% bias-corrected and accelerated (BCa) confidence intervals

Table A6.
Assessment of formative measurement models

Hierarchical order-construct	Outer weights	99% BCa CI	Significant of outer weight ($p < 0.01$)	VIF	Convergent validity
Integrated Customer Services - > CPCI	0.206	[LB: 0.185; UB: 0.234]	Yes	2.551	0.907
Integrated Information Access - > CPCI	0.185	[LB: 0.151; UB: 0.251]	Yes	2.095	
Integrated Order Fulfillment - > CPCI	0.299	[LB: 0.274; UB: 0.331]	Yes	2.844	
Integrated Product and Price - > CPCI	0.207	[LB: 0.175; UB: 0.247]	Yes	2.445	
Integrated Promotion - > CPCI	0.201	[LB: 0.175; UB: 0.235]	Yes	2.920	
Integrated Transaction Information - > CPCI	0.166	[LB: 0.133; UB: 0.196]	Yes	1.728	

Model	Relationship	Std Beta	Std error	*t*-value	*p*-value	BCa 95% CI LB	UB	f^2
1	CPCI - > CE	0.734	0.029	25.326	0.000	0.707	0.799	NA
	CE - > Trust	0.712	0.038	18.541	0.000	0.641	0.768	NA
	CE - > Sat	0.109	0.060	1.836	0.033	0.012	0.208	0.016
	Trust - > Sat	0.714	0.057	12.518	0.000	0.609	0.797	0.685
	Trust - > PI	−0.010	0.060	0.163	0.435	−0.102	0.092	0.000
	Sat - > PI	0.852	0.047	18.153	0.000	0.769	0.923	0.944
2	CPCI - > CE	0.734	0.029	25.137	0.000	0.678	0.776	NA
	CE - > Trust	0.712	0.038	18.814	0.000	0.644	0.768	NA
	CE - > Sat	0.110	0.058	1.903	0.029	0.017	0.204	0.016
	Trust - > Sat	0.714	0.056	12.709	0.000	0.613	0.799	0.684
	Trust - > PI	−0.099	0.073	1.346	0.089	−0.215	0.023	0.010
	CE - > PI	0.149	0.053	2.845	0.002	0.065	0.241	0.039
	Sat - > PI	0.830	0.050	16.665	0.000	0.740	0.904	0.917
3	CPCI - > CE	0.738	0.028	26.081	0.000	0.682	0.777	NA
	CPCI - > Trust	0.260	0.054	4.780	0.000	0.166	0.345	0.066
	CE - > Trust	0.520	0.060	8.613	0.000	0.422	0.619	0.266
	CPCI - > Sat	−0.055	0.054	1.014	0.155	−0.142	0.037	0.004
	CE - > Sat	0.141	0.070	2.004	0.023	0.020	0.251	0.020
	Trust - > Sat	0.726	0.057	12.780	0.000	0.624	0.812	0.666
	Trust - > PI	−0.098	0.074	1.336	0.091	−0.227	0.015	0.010
	CE - > PI	0.149	0.054	2.789	0.003	0.059	0.236	0.039
	Sat - > PI	0.830	0.050	16.708	0.000	0.739	0.905	0.916
4	CPCI - > CE	0.739	0.028	26.067	0.000	0.682	0.778	NA
	CPCI - > Trust	0.265	0.056	4.752	0.000	0.167	0.350	0.069
	CE - > Trust	0.516	0.061	8.479	0.000	0.416	0.614	0.262
	CPCI - > Sat	−0.046	0.056	0.820	0.206	−0.140	0.046	0.002
	CE - > Sat	0.136	0.072	1.887	0.030	0.016	0.255	0.018
	Trust - > Sat	0.725	0.057	12.647	0.000	0.618	0.808	0.660
	CPCI - > PI	0.090	0.041	2.179	0.015	0.019	0.156	0.013
	Trust - > PI	−0.124	0.073	1.687	0.046	−0.251	−0.010	0.016
	CE - > PI	0.098	0.058	1.690	0.045	0.007	0.198	0.012
	Sat - > PI	0.835	0.049	17.171	0.000	0.745	0.906	0.936

Note(s): NA means that the effect size is not applicable for single exogenous variable on endogenous variable; * p-value < 0.05; ** p-value < 0.01 (one-tailed test)

Table A7. Assessment of the structural model

Corresponding author
Jun-Hwa Cheah can be contacted at: jackycheahjh@gmail.com