COMPARING THE PREDICTIVE ABILITY OF PLS AND COVARIANCE MODELS

Completed Research Paper

Joerg Evermann
Memorial University of Newfoundland
St. John's, Canada
jevermann@mun.ca

Mary Tate
Victoria University of Wellington
Wellington, New Zealand
mary.tate@vuw.ac.nz

Abstract

Partial Least Squares (PLS) is a statistical technique that is widely used in the Partial Least Squares (PLS) is a popular technique for estimating structural equation models with latent variables. It is frequently perceived as an alternative to covariance analysis of such models. While its proponents recognize the shortcomings of PLS for testing explanatory models in comparison to covariance models, PLS is instead positioned as a tool for prediction and argued to be preferable to covariance analysis for this purpose. In this paper, we present an initial study that compares the predictive ability of PLS and covariance analysis in a range of situations using a simulation study. Our results show that PLS does offer some advantages over covariance models, but that these are not the ones advocated by PLS proponents.

Keywords: Partial Least Squares, Prediction, Simulation Study, Covariance Analysis, Research Methods

Introduction

Explanation and prediction are two main purposes of theories and statistical methods (Gregor, 2006). Explanation is understood as the identification of causal mechanisms underlying a phenomenon. On the statistical level, explanation is perceived to be primarily concerned with *testing* the faithful representation of causal mechanisms by the statistical model and the estimation of true population parameter values from samples (Shmueli and Koppius, 2011). Prediction is viewed as the ability to predict values for individual cases. While predictive models may be based on causal mechanisms, they need not necessarily be (Gregor, 2006). On the statistical level, predictive models may be developed in a more exploratory and data-driven way (Shmueli and Koppius, 2011). The aim is not to test whether models accurately represent the causal mechanisms, but instead to identify the best way to predict observations for specific cases that are similar to those in the sample. Prediction is argued to be an important aspect of information systems research (Shmueli and Koppius, 2011).

Structural equation modeling is an increasingly popular statistical technique that allows researchers to represent latent constructs, observations, and their relationship in a single model. The partial least squares (PLS) technique treats the latent constructs as weighted composites of the corresponding observed variables and estimates the composite model using multiple regression. In contrast, covariance analysis estimates the model by minimizing the difference between the model-implied and the observed covariance matrices. PLS is often argued to be technique that emphasizes the observed data over the theory and has therefore been argued to be preferable to covariance analysis for prediction (Hair, Ringle, and Sarstedt, 2011; Reinartz, Haenlein, and Henseler, 2009). In fact, Herman Wold who originally developed PLS positioned it as a method for prediction (as quoted in Dijkstra, 1983), and Lohmöller (1989), a major contributor to the development of PLS, demonstrated that the population parameter estimates produced by the PLS algorithm are biased, in effect acknowledging that PLS should not be used for explanation (i.e. the estimation of true, unbiased population parameters) but may instead be more useful for prediction.

This study is of particular interest to Information Systems researchers, as this field is the main user of the PLS technique for estimating structural equation models (Rouse and Corbitt, 2008). Many of the major developments in the use of PLS have occurred in the IS context (e.g. Chin et al., 2003; Goodhue et al., 2007; Wetzels et al., 2009). Additionally, a number of recent editorials on PLS in MIS Quarterly highlight the important role that PLS plays in the IS discipline (Marcoulides and Saunders, 2006; Marcoulides, Chin, and Saunders, 2009; Ringle et al., 2012). The last of these (Ringle et al., 2012) reported 65 studies in MIS Quarterly over the period of 2001 to 2011. This is more than three times the combined number of PLS studies in the top three marketing journals (JMR, JM, JAMS) in the same period. Ringle et al. (2012) also suggest that a focus on prediction is one of the main stated reasons for researchers to adopt PLS over other techniques, both in IS as well as the marketing discipline. However, they note that, despite the stated predictive aim of many PLS studies, none report appropriate predictive ability metrics.

In this study, we examine the predictive ability of PLS and compare it to the predictive ability of covariance analysis, which is traditionally associated more with explanation and testing rather than with prediction. Using a simulation study, we examine the predictive ability of PLS and covariance estimates for a range of models under conditions of differing sample sizes, numbers of indicators, and item loadings. To our knowledge, this is the first study to provide a systematic evaluation of predictive ability of different estimation and prediction methods.

The remainder of the paper is structured as follows. We next present a brief description of blindfolding, the primary and recommended technique to evaluate predictive ability. This is followed by a description of the simulation study. We present and discuss the results of that study and conclude the paper with recommendations for researchers and an outlook to future research required in this area.

Evaluating Predictive Ability using Blindfolding

While in traditional regression models the R^2 proportion of explained variance is an indicator of the predictive strength of the model, researchers have recently advocated the use of blindfolding for assessing the predictive strength of structural equation models (Chin, 2010; Ringle et al., 2012). In blindfolding, the researcher omits a number of observations from the data set, estimates the model parameters, and uses the estimated model to predict the omitted observations[1].

Blindfolding can be done on any set of variables. However, the predictive ability of the model typically concerns the indicator variables for the endogenous latent variables. Blindfolding proceeds by considering a block of N cases and K indicators, e.g. of the indicators of the endogenous latent variables. Beginning with the first data point (row 1, column 1) of this block, every k-th observation is omitted where k is the omission distance. To estimate the model, the omitted values are typically replaced with the variable mean, though other imputation techniques may be used. Based on the estimated model, the estimates for the omitted values are compared to the observed values, using the squared difference (E). At the same time, the difference between the variable mean (or otherwise imputed values) and the observed values are also compared using the squared difference (O). Beginning with the second data point, another set of values is omitted and the squared differences are computed. This process is repeated k times. The predictive measure for these variables is then calculated as

$$Q^2 = 1 - \frac{\sum_k E_k}{\sum_k O_k}$$

Based on different procedures for predicting observations from the model, one can distinguish communality-based and redundancy-based blindfolding, with correspondingly differing values for Q^2_{comm} and Q^2_{red} predictive measures. In communality-based blindfolding, the predicted values are based on the estimated composite scores and the factor loadings. For redundancy-based blindfolding, factor loadings are also used but the composite scores themselves are predicted from the structural model using the estimated multiple regression coefficients. This takes into account the unexplained variance in endogenous latent variables. Redundancy-based blindfolding is applicable only to observations of indicators of endogenous latent variables, while communality-based blindfolding can be applied to all observed variables.

There are different recommendations for the blindfolding omission distance k in the literature, though generally between 5 and 10. The blindfolding distance represents an assumption as to how far out of sample the future values are, which are to be predicted. The omission distance indicates how much of the sample will be discarded for parameter estimation: For a given omission distance k, a proportion of $^1/_k$ of the sample values will be discarded. A small omission distance (e.g. $k = 5$) will retain relatively less of the original sample for the parameter estimation than a large omission distance (e.g. $k = 20$). As a result, the distributional characteristics of the estimated sample are more likely to differ from those of the complete sample for small distances than for large distances. As a consequence, the predicted values for small omission distances will be further from the estimating distributional characteristics than for large omission distances. Hence, there is no single "correct" omission distance, but only an assumption by the researcher how far out of sample the model will be asked to predict. Unfortunately, given the small number of predictive studies in information systems (Shmueli and Koppius, 2011), we do not know which omission distance represents a typical prediction situation.

Chin (2010) recommends to use redundancy-based blindfolding to assess the predictive relevance of ones "theoretical/structural model" (p. 680) and suggests that a value of $Q^2 > 0.5$ is indicative of a predictive model.

---

[1] Blindfolding should not be confused with Jackknifing. The latter is a resampling technique with the aim of computing empirical parameter standard errors or confidence intervals, and is closer to the bootstrap than the blindfolding. In contrast to the k-distance blindfolding, where k individual observations are removed without regard for the cases they belong to, and are then predicted from the remaining sample, the "remove k" jackknife removes k entire cases and re-estimates model parameters but does not predict. In contrast to bootstrapping, where a new sample is generated by sampling with replacement from the original sample, the jackknife creates each new sample by simply omitting k cases. This makes the jackknife deterministic, in contrast to the bootstrap.

Study Design

To compare the predictive ability of models estimated using PLS to those estimated using covariance analysis, we used a simulation study. A simulation study is a controlled experiment in which observations are generated from a model with parameters fixed by and known to the researcher. The PLS and covariance algorithms are then used to estimate the model parameters with the simulated data. The advantage of simulation studies is that different conditions of sample size, numbers of indicators for each latent variable and loadings can be examined. In effect, it constitutes a controlled experiment. As such, it emphasizes internal validity over external validity (i.e. realism).

In our study, we examine the three models shown in Figures 1, 2 and 3, as these models were used in an existing simulation study on PLS (Evermann and Tate, 2010). They represent a range of model complexity comparable to actual models in the IS literature. Ringle et al. (2012) report a minimum of 3 and mode of 7 latent variables per model in the IS literature that uses PLS. Table 1 shows the different conditions for which data was generated and model parameters were estimated. These conditions are also representative of estimation conditions in the literature; Ringle et al. (2012) report a mean of 3.58 indicators per construct and a mean sample size of 238.12.

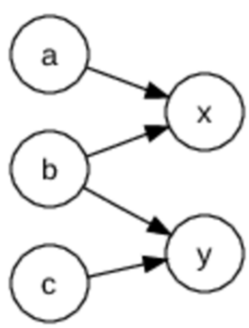

Figure 1: Model 1

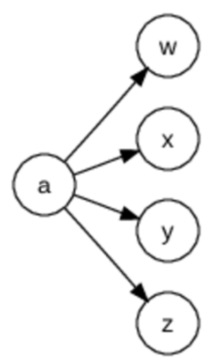

Figure 2: Model 2

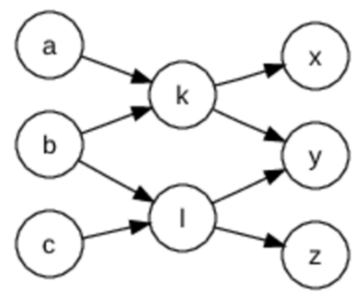

Figure 3: Model 3

Table 1: Experimental conditions

Sample size	100, 250, 750
Number of indicators per latent construct	3, 5, 7
Factor loadings (unstandardized)	Low (0.75), medium (1), high (1.25)

We conduct our simulations under very conservative conditions. For example, our variables are continuous from a multivariate normal distribution, all structural paths are significant, there are no formative indicators in the model, and the estimated model was correctly specified

Data was generated for unstandardized structural regression coefficients of 0.75 and an error variance of 0.1 for all indicator variables. In summary, this yields 3 x 3 x 3 = 27 conditions for each of the three models. For each of these experimental conditions, we estimated 200 samples. For each sample, we estimated the model using both PLS and covariance analysis with ML (Maximum Likelihood) estimation. We implemented both communality- and redundancy-based blindfolding for PLS and covariance analysis, using blindfolding omission distances of $k = 5$ and $k = 20$ for the indicators of all endogenous latent variables. We used mean substitution for estimating the models. Thus, in summary we performed $27 \times 200 \times (20 + 5) = 135000$ PLS and covariance estimations for each of the three models.

For outcome measures, we computed the mean communality-based and redundancy-based Q^2 for each of the 200 samples. We also computed the mean r^2 proportion of explained variance for all endogenous latent constructs. Finally, we wished to compare the model-based prediction methods of SEM and PLS to an atheoretical, data-driven prediction method. While a wide range of such methods exist (Hastie et al., 2009), the EM algorithm is most familiar to researchers as it is frequently used for missing value imputation and is available in many statistical packages. This method of predicting missing values does not rely on a statistical model, but assumes a multivariate-normal distribution of the observed values. It then estimates the means and covariances using the maximum-likelihood method and samples the missing values from the resulting multivariate-normal distribution (Schafer, 1997).

Results and Discussion

The complete results of our simulation are shown in Tables 2-7 in the Appendix of the paper. We have plotted excerpts of our results for model 3 in Figures 4 (mean R^2) and Figure 5 (Q^2). Figure 4 shows the mean R^2 values as a function of the estimation mode (PLS or SEM), prediction mode (Communality or Redundancy), sample size and number of indicators. The figure shows that the mean R^2 is always higher for SEM-estimated models than for PLS estimated models. In contrast to SEM-estimated models, the mean R^2 for PLS-estimated models increases as the number of indicators increases. Figure 5 shows the Q^2 predictive ability metrics, again as a function of the estimation mode (PLS or SEM), prediction mode (Communality or Redundancy), sample size and number of indicators. The figure shows that the communality-based predictive ability from SEM-estimated models is much lower than for PLS-estimated models, and also lower than for redundancy-based prediction. As we did not identify many model-specific phenomena, the equivalent figures for the other two models show similar information and are therefore not included.

Our first observation is that, while there are some commonalities between the two blindfolding omission distances, many of our results differ between the different omission distances. We present and discuss our results accordingly.

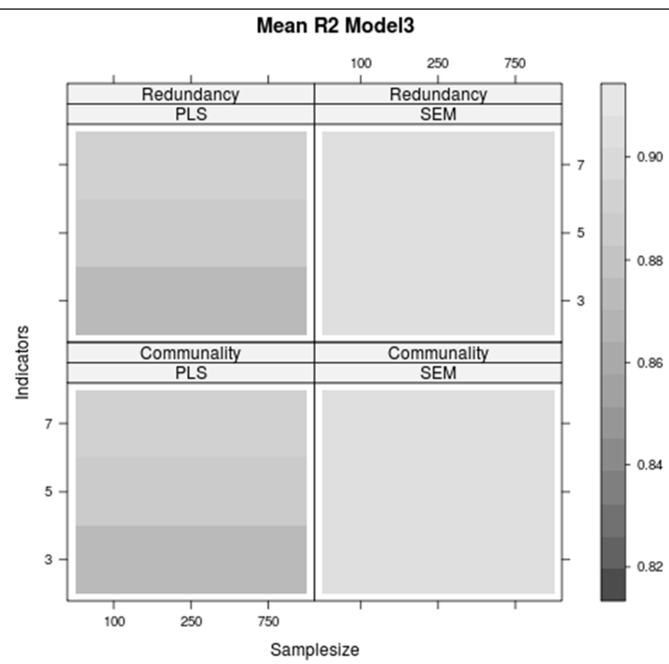

Figure 4: Mean R^2 for model three as a function of estimation mode (PLS/covariance SEM), number of indicators, sample size and (unstandardized) loadings. Omission distance k=5

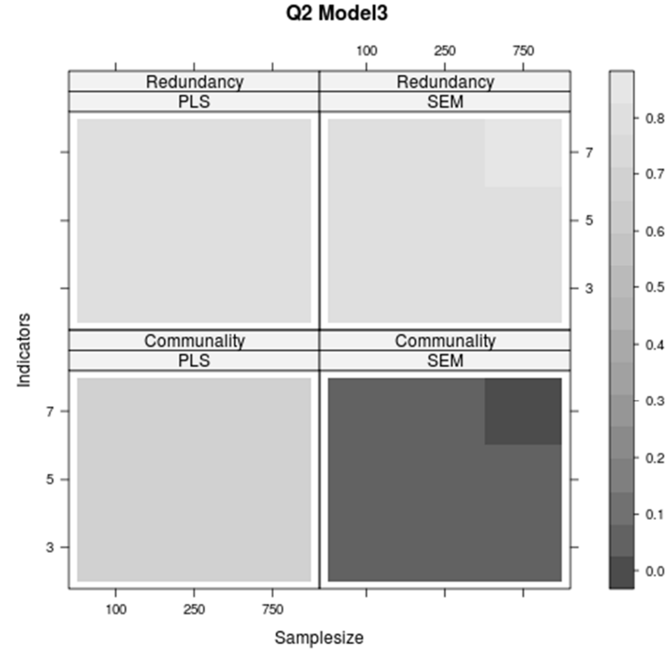

Figure 5: Q^2 for model three as a function of prediction mode (communality/redundancy), estimation mode (PLS/covariance SEM), sample size and (unstandardized) loadings. Omission distance k=5

Results independent of omission distance

A number of observations can be made that are independent of the omission distance. First, for small samples ($n = 100$), communality- and redundancy-based prediction from PLS estimated models generally dominates over the same mode of prediction based on covariance-based SEM estimated models. This holds for all models, loadings and numbers of indicators. In fact, the communality-based Q^2 for covariance-based SEM models is much lower than that for PLS models, while the redundancy-based Q^2 is marginally, but consistently lower by approximately 5%.

Second, the communality-based Q^2 for PLS is always larger than the communality-based Q^2 for the covariance SEM analysis, for all models and all experimental conditions. This reflects the known PLS bias for overestimating measurement loadings (Lohmöller, 1989).

Third, prediction based on EM imputation dominates both communality- and redundancy-based prediction based on PLS or covariance-SEM estimated models for medium and large sample sizes ($n = 250,\ n = 750$) for almost all models and experimental conditions.

Fourth, the redundancy-based Q^2 for both PLS and covariance-based SEM is always above the recommendation of 0.5 for a predictive model (Chin, 2010). It increases with loadings, which is unsurprising, as these determine the residual measurement error in the observed variables and therefore have a direct impact on the predictive ability of the model. It also increases with the number of indicators. Again, this is not surprising as more indicators increase the reliability of the scale and can thus reduce the prediction error. In contrast, there is little to no effect of sample size on redundancy-based Q^2 for either estimation type.

Fifth, the communality-based Q^2 for PLS increases with the sample size and loadings. It is easy to see that the prediction error is reduced as the loadings are increased, because in that case the measurement error is decreased. Moreover, previous research has established that the performance of PLS estimation improves with increasing sample size ("consistency-at-large", Lohmöller, 1989) and the estimated parameter values approach those of covariance estimation in the limit of infinite sample size. The communality-based Q^2 for covariance-SEM estimated models does not show these effects.

Sixth, the mean R^2 for PLS is always less than the mean R^2 for covariance SEM estimation. While most of these differences are small (less than 5%), there are a few large differences for low loadings and a small number of indicators (up to 15%). These results are not unexpected, as it is well-known that PLS de-emphasizes structural estimates, and over-emphasizes measurement loadings (Lohmöller, 1989). In fact, in many applications this is argued to be an advantage of PLS over covariance SEM, as it is argued to de-emphasize the often uncertain theory underlying the statistical model.

Seventh, as expected, the R^2 for the covariance-based SEM estimation is stable across all conditions for all models, whereas the R^2 for the PLS estimation varies with loadings, sample size, and number of indicators. The variation with loadings is a result of the way PLS estimates the weights. In the estimation loops, the composite scores are calculated alternatingly based on the structural model and the measurement model ("inner" and "outer" estimation, cf. Lohmöller, 1989). Hence, the loadings have a strong impact on the estimates in the structural model, thus allowing for a strong influence of the observed variables on the structural coefficients and the resulting R^2 values. The variations with sample size and number of indicators is an example of the consistency-at-large property of PLS estimates, which approach those of covariance-based estimation with increasing sample size and number of indicators (Lohmöller, 1989).

Results differing by omission distance

The main difference between the two blindfolding omission distances examined here is in the communality-based Q^2 values. We first discuss the communality-based Q^2 for PLS-based estimation. For the small omission distance ($k = 5$), the communality-based Q^2 is always smaller than the redundancy-based Q^2 for PLS estimated models. In fact, the communality-based Q^2 for model two is well below the recommendation of 0.5 for a predictive model (Chin, 2010). This may be a result of the fact that this model has a large number of endogenous latent variables compared to the number of exogenous latent variables (Figure 2) whereas the other two models are more balanced. Consequently, a larger proportion of the sample is missing during blindfolding. However, for the larger omission distance ($k = 20$) the communality-based Q^2 for PLS estimated models dominates the redundancy-based Q^2 for medium and large samples, even for our second model.

The communality-based Q^2 values for covariance-based SEM estimated models are very low for the small omission distance ($k = 5$) to the point that they may effectively be considered as zero for all models. In contrast, for the larger omission distance ($k = 20$), the covariance-SEM based communality Q^2 is well above the recommended value of 0.5 (Chin, 2010) when sample sizes are medium ($n = 250$) or large ($n = 750$). They remain effectively zero for small samples. However, even in the medium and large sample conditions, the covariance-SEM based communality Q^2 is lower than the PLS based communality Q^2 for all conditions. As indicated above, this is not entirely surprising, as PLS is argued to overemphasize the measurement model loadings compared to covariance-SEM based estimation, and this bias leads to increased communality-based predictive ability.

For the smaller omission distance ($k = 5$), the redundancy-based prediction from PLS estimated models dominates the redundancy-based prediction from covariance-SEM estimated models, except for large sample sizes ($n = 750$) for model three, where the two perform similarly. However, for the large omission distance ($k = 20$), this is the case only for small samples. In contrast, for medium ($n = 250$) or large ($n = 750$) samples, the redundancy-based Q^2 for covariance-SEM estimated models dominates that for PLS estimated models.

Recommendations and Conclusion

PLS is frequently used because it is argued to be appropriate for prediction, rather than model testing. A recent high-profile editorial in MIS Quarterly (Ringle et al., 2012) calls for increased reporting of predictive ability of PLS models, as 15% of PLS studies claim that prediction is an important reason for choosing PLS. Moreover, Shmueli and Koppius (2011) have argued for increased emphasis on prediction in IS research. These calls motivated the present study. To our knowledge, this is the first study to provide a systematic evaluation of predictive ability of different estimation and prediction methods.

Our results generally support the claim that PLS is good choice for estimating models for use in prediction. Our results also support the specific recommendation by Chin (2010), who suggests that the redundancy-based Q^2 is the appropriate metric for assessing the predictive ability of the structural model.

We were surprised by the good performance of the EM imputation algorithm to recover the blindfolded values. This leads us to suggest that, when data are multivariate normal and the emphasis is purely on prediction, the observations for which dependent values are to be predicted should simply be treated as missing values and EM imputation should be performed; no statistical model is required for this. However, this situation is unlikely to occur in practice, where most data are not multivariate normal so that the performance of the EM imputation will suffer. While we have no data on how much the predictive performance of PLS or covariance-based SEM estimated models will suffer for non-normal data, it would be surprising if EM imputation outperformed prediction based on statistical models for realistic data sets. However, a wide range of other atheoretical, data-driven predictive techniques exist, with a long history of study and use in predictive modeling, data mining, etc. (Hastie et al., 2009). We thus caution researchers that, despite our results, PLS path modeling may not be the best predictive technique for any given data set.

Based on our results, we make the following recommendations when structural equation models are estimated for predictive purposes:

- For small sample sizes, we always recommend redundancy-based prediction from PLS estimated models. For small samples, this prediction mode dominates covariance-SEM based prediction for all experimental conditions and also dominates communality-based prediction for both PLS and covariance-based SEM estimated models.
- For medium and large samples, when prediction is to be made for values relatively close to those in the estimation sample, our results suggest that redundancy-based prediction from covariance-SEM estimated models is in many situations the superior prediction method. If covariance-SEM model estimation is not possible, e.g. because of non-linear structural relationships or under-identification of the model, then *communality-based* prediction from PLS estimated should be used.
- For medium and large samples, when prediction is to be made for values less close to the estimation sample, our results suggest that redundancy-based prediction from PLS estimated models is the superior prediction method. Given that the PLS literature recommends blindfolding omission distances closer to $k = 5$ than to $k = 20$, this should become a standard recommendation.
- When predictive ability is interpreted as the ability to explain variance in the endogenous latent variables, rather than the ability to predict individual observations, covariance-based SEM estimation should be used.

However, while these are specific recommendations based on our results, we recommend that, in line with the notion that prediction is possible without explanation (Gregor, 2006) and that prediction allows more flexible and data-driven approaches than model testing (Shmueli and Koppius, 2011), researchers should use both methods of estimation (PLS and covariance-based SEM) and both methods of prediction (communality-based and redundancy-based) to explore the best way to predict individual scores from the specific model. If prediction is indeed the main aim of the study, the fact that a model shows lack of fit by traditional metrics, such as the PLS goodness-of-fit index or the various fit indices for covariance-based estimation, is irrelevant. Moreover, to pursue predictive validity, post-hoc model modifications should be explored because the bias of significance tests is not a concern in this case. Thus, while PLS may typically be the preferred option for prediction, researchers should explore both estimation methods and both prediction methods.

We note a distinct difference between small and large blindfolding omission distances. As we indicated earlier, there is not correct value for the omission distance. Instead, the blindfolding omission distance is a measure for how far "out of sample" the to be predicted values are. A small omission distance omits a larger proportion of the sample for parameter estimation than a large omission distance. Consequently, a small omission distance suggests that the to be predicted values are further from the sample values in terms of their distributional characteristics. The "632 bootstrap procedure" (Efron and Tibshirani, 1997; Hastie et al., 2009) for cross-validation has been shown to be superior to other cross-validation methods. We recommend that PLS researchers investigate the performance of this method and PLS users adopt this cross-validation method for future work.

While this study aimed to investigate the predictive abilities of PLS and covariance-based SEM estimation for different modes of prediction (communality- and redundancy-based), our simulations have been conducted under very conservative conditions. For example, our variables were continuous from a multivariate normal distribution, all structural paths are significant, there were no formative indicators in the model, and the model was correctly specified. In practice, it is unlikely that all, or even many, of these assumptions are met to the extent they were for this study. Thus, future extensions of this work should investigate the predictive performance of PLS and SEM for discrete data (e.g. from Likert scales) and varying degrees of skewness and kurtosis of the data.

Further, it is impossible to even know whether a model is correctly specified in practice, so one should in general assume that the estimated model is not the true generating model. While establishing the correctness of the model is not a priority from the perspective of prediction that we have assumed here, it is known that parameter estimates are biased for misspecified models. These biased estimates will affect the predictive ability of the model, but it is unclear what the direction and magnitude of these effects will be for the different modes of estimation and prediction. Future research needs to examine predictive ability under a wider range conditions.

In general, while this brief, initial study has shown that PLS may be a more appropriate choice than covariance SEM when the goal is prediction, this must be qualified. When the goal is prediction, the underlying model is typically not important, and thus the predictive ability of PLS should best be compared to other, often atheoretical, prediction techniques, such as canonical regression, kernelized PLS regression etc. (Hastie et al., 2009). Generally, any PLS path model will impose constraints on the estimation and prediction that are not present when using e.g. a simple, direct PLS or canonical regression between independent and dependent variables. McDonald (1996) write with respect to PLS prediction that "a path models is … generally subobtimally predictive" and that "if the object of the analysis were to predict the response variables, … we cannot do better than to use a multivariate regression … or the corresponding canonical variate analysis." (pg. 266) In contrast, when the correctness of the model is important, SEM should be preferred, as PLS cannot test the correctness of the model (Evermann and Tate, 2010).

To conclude, this study contributes the first systematic analysis of predictive ability of different structural equation estimation methods and different prediction methods. The resulting recommendations are based on strong empirical evidence under a range of different conditions.

References

Chin, W.W., Marcolin, B.L. and Newsted, P.R. 2003. "A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study." *Information Systems Research* (14:2), 189-217.

Chin, W.W. 2010. "How to write up and report PLS analyses." Esposito Vinzi, E. et al. (eds.) *Handbook of Partial Least Squares*. Berlin, Germany: Springer-Verlag, 655-690.

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

Efron, B. and Tibshirani, R. 1997. "Improvements on crossvalidation: The .632+ bootstrap." *Journal of the American Statistical Association* (92:438), 548-560.

Evermann, J. and Tate, M. 2010. "Testing models or fitting models? Identifying model misspecifications in PLS." *Proceedings of the Proceedings of the 31st International Conference on Information Systems (ICIS),* St. Louis, MS.

Goodhue, D., Lewis, W., and Thompson, R. 2007. "Statistical power in analyzing interaction effects: Questioning the advantage of PLS with product indicators." *Information Systems Research* (18:2), 211-227.

Gregor, S. 2006. "The nature of theory in information systems." *MIS Quarterly* (30:3), 611-642.
Hastie, T., Tibshirani, R. and Friedman, J. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Verlag, Berlin.
Hair, J.F, Ringle C.M. and Sarstedt, M. 2011. "PLS-SEM: Indeed a silver bullet." *Journal of Marketing Theory and Practice* (19:2), 139-151.
Marcoulides, G.A. and Saunders, C. 2006. "PLS: A silver bullet?" *MIS Quarterly* (30:2), iii-ix.
Marcoulides, G.A., Chin, W.W. and Saunders, C. 2009. "A critical look at partial least squares modeling." *MIS Quarterly* (33:1), 171-175.
McDonald, R.P. 1996. "Path analysis with composite variables." *Multivariate Behavioral Research* (31:2), 239-270.
Lohmöller, J.B. 1989. *Latent Variable Path Modeling with Partial Least Squares*. Heidelberg, Germany: Physica-Verlag.
Reinartz, W., Haenlein, M., and Henseler, J. 2009 "An empirical comparison of the efficacy of covariance-based and variance-based SEM." *International Journal of Research in Marketing* (26*)*, 332-344.
Ringle, C.M., Sarstedt, M., and Straub, D.W. 2012. "A critical look at the use of PLS-SEM in MIS Quarterly." *MIS Quarterly* (36:1), iii-xiv.
Rouse, A.C. and Corbitt, B. 2008. "There's SEM and "SEM": A critique of the use of PLS regression in information systems research." *Proceedings of the 19th Australasian Conference on Information Systems,* 3-5 December, Christchurch, New Zealand.
Schafer, J.F. 1997. *Analysis of Incomplete Multivariate Data*. Boca Raton, FL: CRC Press/Chapman & Hall.
Shmueli, G. and Koppius, O.R. 2011. "Predictive analytics in information systems research." *MIS Quarterly* (35:3), 553-572.
Wetzels, M., Odekerken-Schröder, G., and van Oppen, C. 2009. "Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration." *MIS Quarterly* (33:1), 177-195.

Appendix

The appendix contains the complete results of the simulation study in tabular form. Each row in the table represents one of the 27 experimental conditions, The columns Q2C represent the communality-based Q^2 (for PLS and covariance-based SEM estimation) and the columns Q2R represent the redundancy-based Q^2 (for PLS and covariance-based SEM estimation). The columns R^2 represent the mean R^2 of the endogenous latents (for PLS and covariance-based SEM estimation) and the column EM Imputed represents the Q^2 when the blindfolded values are treated as missing values and imputed by an Expectation-Maximization algorithm (Schafer, 1997).

Table 2: Predictive ability results for Model 1 (largest Q^2 value highlighted), k=5

N	I	L	Q2C (PLS)	Q2R (PLS)	Q2C (SEM)	Q2R (SEM)	R2 (PLS)	R2 (SEM)	EM Imputed
100	3	0.75	0.550	**0.723**	0.142	0.677	0.831	0.920	0.715
100	3	1	0.575	0.784	0.105	0.742	0.864	0.918	**0.787**
100	3	1.25	0.595	0.822	0.078	0.784	0.885	0.920	**0.830**
100	5	0.75	0.557	**0.737**	0.107	0.702	0.863	0.920	0.710
100	5	1	0.585	**0.798**	0.073	0.765	0.886	0.919	0.782
100	5	1.25	0.596	**0.826**	0.054	0.795	0.898	0.918	0.821
100	7	0.75	0.561	**0.747**	0.083	0.716	0.879	0.920	0.693
100	7	1	0.587	**0.802**	0.059	0.771	0.895	0.918	0.775
100	7	1.25	0.601	**0.832**	0.042	0.802	0.904	0.919	0.809
250	3	0.75	0.550	0.718	0.145	0.686	0.829	0.920	**0.733**
250	3	1	0.579	0.785	0.102	0.757	0.865	0.919	**0.800**
250	3	1.25	0.597	0.822	0.074	0.797	0.885	0.920	**0.839**
250	5	0.75	0.558	0.735	0.102	0.712	0.861	0.918	**0.743**
250	5	1	0.587	0.797	0.069	0.777	0.886	0.918	**0.811**
250	5	1.25	0.601	0.829	0.050	0.810	0.899	0.919	**0.841**
250	7	0.75	0.564	0.744	0.080	0.726	0.878	0.919	**0.745**
250	7	1	0.589	0.803	0.054	0.785	0.896	0.919	**0.813**
250	7	1.25	0.601	0.830	0.038	0.813	0.902	0.917	**0.842**
750	3	0.75	0.549	0.715	0.145	0.687	0.827	0.918	**0.743**
750	3	1	0.581	0.786	0.102	0.762	0.865	0.918	**0.809**
750	3	1.25	0.598	0.821	0.074	0.802	0.884	0.918	**0.845**
750	5	0.75	0.559	0.735	0.102	0.717	0.862	0.919	**0.758**
750	5	1	0.588	0.797	0.068	0.781	0.886	0.918	**0.821**
750	5	1.25	0.602	0.829	0.047	0.816	0.898	0.919	**0.851**
750	7	0.75	0.563	0.742	0.079	0.731	0.877	0.918	**0.764**
750	7	1	0.590	0.802	0.051	0.790	0.895	0.919	**0.824**
750	7	1.25	0.602	0.831	0.035	0.820	0.903	0.918	**0.853**

Table 3: Predictive ability results for Model 2 (largest Q^2 value highlighted), k=5									
N	I	L	Q2C (PLS)	Q2R (PLS)	Q2C (SEM)	Q2R (SEM)	R2 (PLS)	R2 (SEM)	EM Imputed
100	3	0.75	0.298	0.592	0.085	0.551	0.731	0.844	**0.602**
100	3	1	0.318	0.697	0.065	0.647	0.783	0.849	**0.688**
100	3	1.25	0.327	**0.747**	0.051	0.696	0.806	0.849	0.740
100	5	0.75	0.302	**0.618**	0.063	0.584	0.779	0.849	0.608
100	5	1	0.319	**0.709**	0.048	0.666	0.808	0.849	0.688
100	5	1.25	0.327	**0.755**	0.037	0.709	0.823	0.850	0.734
100	7	0.75	0.300	**0.621**	0.051	0.595	0.795	0.846	0.601
100	7	1	0.319	**0.708**	0.037	0.668	0.816	0.845	0.690
100	7	1.25	0.326	**0.757**	0.029	0.714	0.830	0.849	0.729
250	3	0.75	0.291	0.591	0.084	0.564	0.734	0.848	**0.607**
250	3	1	0.309	0.690	0.062	0.653	0.781	0.848	**0.696**
250	3	1.25	0.319	**0.741**	0.046	0.703	0.803	0.847	**0.741**
250	5	0.75	0.292	0.611	0.058	0.591	0.778	0.848	**0.616**
250	5	1	0.310	**0.703**	0.041	0.673	0.808	0.849	0.701
250	5	1.25	0.319	**0.751**	0.031	0.718	0.823	0.849	0.744
250	7	0.75	0.294	**0.620**	0.046	0.605	0.798	0.849	0.618
250	7	1	0.311	**0.709**	0.032	0.683	0.820	0.849	0.703
250	7	1.25	0.319	**0.752**	0.024	0.722	0.829	0.848	0.747
750	3	0.75	0.288	0.593	0.082	0.570	0.737	0.850	**0.612**
750	3	1	0.306	0.692	0.059	0.660	0.783	0.850	**0.700**
750	3	1.25	0.315	0.742	0.043	0.708	0.805	0.849	**0.746**
750	5	0.75	0.289	0.608	0.057	0.594	0.777	0.848	**0.624**
750	5	1	0.307	0.702	0.039	0.678	0.808	0.849	**0.706**
750	5	1.25	0.316	**0.749**	0.028	0.722	0.822	0.849	**0.749**
750	7	0.75	0.290	0.616	0.043	0.607	0.796	0.848	**0.630**
750	7	1	0.307	0.707	0.029	0.686	0.820	0.850	**0.712**
750	7	1.25	0.315	0.752	0.021	0.726	0.829	0.848	**0.753**

Table 4: Predictive ability results for Model 3 (largest Q^2 value highlighted), k=5

N	I	L	Q2C (PLS)	Q2R (PLS)	Q2C (SEM)	Q2R (SEM)	R2 (PLS)	R2 (SEM)	EM Imputed
100	3	0.75	0.621	**0.727**	0.109	0.709	0.820	0.908	0.715
100	3	1	0.650	**0.783**	0.076	0.766	0.854	0.906	0.782
100	3	1.25	0.667	0.812	0.056	0.800	0.874	0.907	**0.823**
100	5	0.75	0.625	**0.732**	0.077	0.723	0.854	0.907	0.676
100	5	1	0.653	**0.787**	0.054	0.778	0.874	0.905	0.761
100	5	1.25	0.668	**0.816**	0.039	0.806	0.887	0.907	0.799
100	7	0.75	0.630	**0.739**	0.061	0.735	0.870	0.908	0.614
100	7	1	0.655	**0.792**	0.042	0.784	0.886	0.908	0.708
100	7	1.25	0.669	**0.818**	0.034	0.812	0.893	0.908	0.758
250	3	0.75	0.621	0.724	0.107	0.721	0.820	0.907	**0.750**
250	3	1	0.651	0.781	0.074	0.780	0.856	0.907	**0.812**
250	3	1.25	0.666	0.811	0.052	0.811	0.874	0.907	**0.844**
250	5	0.75	0.624	0.729	0.073	0.734	0.852	0.906	**0.745**
250	5	1	0.653	0.786	0.049	0.789	0.876	0.907	**0.809**
250	5	1.25	0.669	0.816	0.035	0.820	0.888	0.908	**0.839**
250	7	0.75	0.626	0.732	0.057	**0.741**	0.867	0.906	0.738
250	7	1	0.654	0.788	0.037	0.795	0.884	0.907	**0.803**
250	7	1.25	0.669	0.816	0.027	0.821	0.892	0.907	**0.836**
750	3	0.75	0.620	0.722	0.107	0.724	0.820	0.908	**0.762**
750	3	1	0.651	0.781	0.072	0.785	0.857	0.908	**0.821**
750	3	1.25	0.666	0.810	0.051	0.817	0.875	0.907	**0.852**
750	5	0.75	0.625	0.729	0.072	0.740	0.853	0.907	**0.767**
750	5	1	0.654	0.786	0.046	0.796	0.877	0.908	**0.824**
750	5	1.25	0.668	0.813	0.033	0.823	0.887	0.907	**0.854**
750	7	0.75	0.627	0.731	0.054	0.746	0.868	0.907	**0.768**
750	7	1	0.654	0.786	0.035	0.799	0.884	0.907	**0.826**
750	7	1.25	0.668	0.814	0.024	0.826	0.892	0.907	**0.854**

Table 5: Predictive ability results for Model 1 (largest Q^2 value highlighted), k=20

N	I	L	Q2C (PLS)	Q2R (PLS)	Q2C (SEM)	Q2R (SEM)	R2 (SEM)	R2 (PLS)	EM Imputed
100	3	0.75	0.561	0.722	0.239	0.687	0.921	0.831	**0.725**
100	3	1	0.595	0.791	0.186	0.772	0.921	0.865	**0.796**
100	3	1.25	0.608	0.822	0.142	0.815	0.918	0.883	**0.834**
100	5	0.75	0.569	**0.738**	0.186	0.727	0.923	0.863	0.719
100	5	1	0.596	**0.799**	0.135	0.796	0.918	0.886	0.790
100	5	1.25	0.614	0.831	0.100	**0.832**	0.918	0.897	0.828
100	7	0.75	0.577	**0.749**	0.151	0.745	0.922	0.878	0.713
100	7	1	0.601	**0.806**	0.104	0.805	0.917	0.894	0.782
100	7	1.25	0.613	0.833	0.076	**0.840**	0.919	0.903	0.821
250	3	0.75	0.754	0.719	0.642	0.745	0.919	0.827	**0.799**
250	3	1	0.799	0.789	0.636	0.809	0.918	0.866	**0.873**
250	3	1.25	0.820	0.821	0.596	0.843	0.919	0.884	**0.913**
250	5	0.75	0.772	0.736	0.669	0.768	0.920	0.861	**0.812**
250	5	1	0.815	0.798	0.679	0.825	0.919	0.886	**0.885**
250	5	1.25	0.836	0.829	0.654	0.853	0.918	0.896	**0.922**
250	7	0.75	0.780	0.745	0.674	0.775	0.918	0.879	See note
250	7	1	0.820	0.802	0.691	0.829	0.918	0.895	
250	7	1.25	0.842	0.833	0.678	0.857	0.918	0.904	
750	3	0.75	0.753	0.716	0.643	0.749	0.918	0.828	**0.807**
750	3	1	0.800	0.788	0.636	0.812	0.918	0.865	**0.879**
750	3	1.25	0.822	0.821	0.600	0.846	0.918	0.884	**0.918**
750	5	0.75	0.773	0.735	0.671	0.771	0.918	0.862	**0.824**
750	5	1	0.816	0.797	0.682	0.828	0.919	0.886	**0.892**
750	5	1.25	0.837	0.829	0.658	0.857	0.918	0.898	**0.927**
750	7	0.75	0.781	0.744	0.676	0.779	0.918	0.878	See note
750	7	1	0.822	0.803	0.694	0.833	0.918	0.896	
750	7	1.25	0.842	0.832	0.682	0.861	0.918	0.903	

Note: The EM imputation did not converge within 20000 iterations.

Table 6: Predictive ability results for Model 2 (largest Q^2 value highlighted), k=20									
N	I	L	Q2C (PLS)	Q2R (PLS)	Q2C (SEM)	Q2R (SEM)	R2 (SEM)	R2 (PLS)	EM Imputed
100	3	0.75	0.313	**0.602**	0.160	0.552	0.847	0.735	0.594
100	3	1	0.331	**0.698**	0.131	0.654	0.846	0.781	0.684
100	3	1.25	0.341	**0.749**	0.103	0.715	0.848	0.804	0.733
100	5	0.75	0.313	**0.616**	0.121	0.591	0.846	0.777	0.605
100	5	1	0.331	**0.704**	0.096	0.687	0.849	0.807	0.685
100	5	1.25	0.342	**0.754**	0.073	0.737	0.850	0.821	0.737
100	7	0.75	0.314	**0.619**	0.104	0.613	0.849	0.796	0.607
100	7	1	0.332	**0.707**	0.076	0.697	0.847	0.818	0.687
100	7	1.25	0.341	**0.753**	0.057	0.745	0.850	0.827	0.734
250	3	0.75	0.643	0.595	0.554	0.617	0.848	0.735	**0.684**
250	3	1	0.697	0.692	0.589	0.699	0.848	0.782	**0.790**
250	3	1.25	0.726	0.743	0.583	0.746	0.849	0.803	**0.850**
250	5	0.75	0.647	0.607	0.580	0.635	0.848	0.777	**0.700**
250	5	1	0.703	0.703	0.628	0.713	0.849	0.807	**0.808**
250	5	1.25	0.731	0.750	0.639	0.754	0.848	0.822	**0.867**
250	7	0.75	0.651	0.617	0.588	0.643	0.849	0.797	See note
250	7	1	0.706	0.709	0.641	0.720	0.851	0.819	
250	7	1.25	0.733	0.754	0.659	0.758	0.848	0.829	
750	3	0.75	0.637	0.589	0.558	0.621	0.849	0.735	**0.691**
750	3	1	0.695	0.690	0.592	0.703	0.849	0.782	**0.794**
750	3	1.25	0.723	0.741	0.586	0.748	0.849	0.805	**0.854**
750	5	0.75	0.648	0.611	0.582	0.638	0.849	0.778	**0.714**
750	5	1	0.701	0.701	0.630	0.715	0.848	0.808	**0.815**
750	5	1.25	0.730	0.750	0.643	0.759	0.850	0.822	**0.873**
750	7	0.75	0.650	0.617	0.588	0.644	0.848	0.798	See note
750	7	1	0.704	0.706	0.643	0.721	0.849	0.819	
750	7	1.25	0.731	0.752	0.663	0.761	0.848	0.830	

Note: The EM imputation did not converge within 20000 iterations.

Table 7: Predictive ability results for Model 1 (largest Q^2 value highlighted), k=20

N	I	L	Q2C (PLS)	Q2R (PLS)	Q2C (SEM)	Q2R (SEM)	R2 (SEM)	R2 (PLS)	EM Imputed
100	3	0.75	0.630	0.726	0.212	0.728	0.907	0.820	**0.730**
100	3	1	0.662	0.785	0.153	0.794	0.907	0.856	**0.832**
100	3	1.25	0.677	0.814	0.116	**0.836**	0.909	0.875	0.793
100	5	0.75	0.636	0.736	0.155	**0.749**	0.908	0.852	0.701
100	5	1	0.665	0.790	0.109	**0.814**	0.909	0.876	0.812
100	5	1.25	0.678	0.817	0.079	**0.842**	0.907	0.888	0.772
100	7	0.75	0.637	0.738	0.124	**0.759**	0.905	0.870	0.659
100	7	1	0.664	0.791	0.086	**0.817**	0.906	0.885	0.784
100	7	1.25	0.679	0.817	0.063	**0.848**	0.906	0.893	0.744
250	3	0.75	0.764	0.724	0.651	0.763	0.906	0.820	**0.802**
250	3	1	0.808	0.782	0.632	0.821	0.908	0.857	**0.912**
250	3	1.25	0.830	0.813	0.585	0.850	0.908	0.874	**0.873**
250	5	0.75	0.769	0.731	0.666	0.777	0.907	0.854	See note
250	5	1	0.812	0.787	0.668	0.830	0.907	0.876	
250	5	1.25	0.833	0.814	0.639	0.857	0.907	0.888	
250	7	0.75	0.771	0.732	0.670	0.782	0.908	0.868	
250	7	1	0.814	0.789	0.680	0.832	0.906	0.885	
250	7	1.25	0.834	0.815	0.658	0.858	0.906	0.892	
750	3	0.75	0.763	0.721	0.655	0.768	0.907	0.820	**0.813**
750	3	1	0.807	0.781	0.638	0.825	0.907	0.857	**0.918**
750	3	1.25	0.829	0.810	0.591	0.854	0.907	0.875	**0.880**
750	5	0.75	0.769	0.729	0.670	0.781	0.907	0.854	See note
750	5	1	0.811	0.785	0.671	0.833	0.907	0.877	
750	5	1.25	0.832	0.813	0.641	0.860	0.907	0.887	
750	7	0.75	0.770	0.731	0.672	0.785	0.908	0.868	
750	7	1	0.813	0.788	0.684	0.837	0.907	0.885	
750	7	1.25	0.834	0.815	0.663	0.862	0.906	0.893	

Note: The EM imputation did not converge within 20000 iterations.