2011

An evaluation of PLS based complex models: the roles of power analysis, predictive relevance and GoF index

Shahriar Akter
University of New South Wales, sakter@uow.edu.au

John D'Ambra
University of New South Wales

Pradeep Ray
University of New South Wales

Publication Details

Akter, S., D'Ambra, J. & Ray, P. (2011). An evaluation of PLS based complex models: the roles of power analysis, predictive relevance and GoF index. Proceedings of the 17th Americas Conference on Information Systems (AMCIS2011) (pp. 1-7). Detroit, USA: Association for Information Systems.

An evaluation of PLS based complex models: the roles of power analysis, predictive relevance and GoF index

Abstract

Structural equation modeling (SEM) is an important tool to estimate a network of causal relationships linking two or more complex concepts. The PLS approach to SEM, also known as component based SEM, is becoming more prominent for estimating large complex models due to its soft modeling assumptions. This 'soft modeling' refers to the greater flexibility of PLS technique in developing and validating the complex models. However, to establish rigor in such complex modeling, this study highlights the critical roles of power analysis, predictive relevance and GoF index. The findings of the study show that power analysis is essential to establish conjectures based on IT artifacts, predictive relevance is vital to measure how well observed values are reproduced by the model and finally, GoF index is crucial for assessing the global validity of a complex model.

Disciplines

Business | Social and Behavioral Sciences

Publication Details

Akter, S., D'Ambra, J. & Ray, P. (2011). An evaluation of PLS based complex models: the roles of power analysis, predictive relevance and GoF index. Proceedings of the 17th Americas Conference on Information Systems (AMCIS2011) (pp. 1-7). Detroit, USA: Association for Information Systems.

AN EVALUATION OF PLS BASED COMPLEX MODELS: THE ROLES OF POWER ANALYSIS, PREDICTIVE RELEVANCE AND GOF INDEX

Shahriar Akter
Australian School of Business
The University of New South Wales
s.akter@unsw.edu.au

John D'Ambra
Australian School of Business
The University of New South Wales
j.dambra@unsw.edu.au

Pradeep Ray
Australian School of Business
The University of New South Wales
p.ray@unsw.edu.au

ABSTRACT

Structural equation modeling (SEM) is an important tool to estimate a network of causal relationships linking two or more complex concepts. The PLS approach to SEM, also known as component based SEM, is becoming more prominent for estimating large complex models due to its soft modeling assumptions. This 'soft modeling' refers to the greater flexibility of PLS technique in developing and validating the complex models. However, to establish rigor in such complex modeling, this study highlights the critical roles of power analysis, predictive relevance and GoF index. The findings of the study show that *power analysis* is essential to establish conjectures based on IT artifacts, *predictive relevance* is vital to measure how well observed values are reproduced by the model and finally, *GoF index* is crucial for assessing the global validity of a complex model.

Keywords

Complex models, PLS path modeling, power analysis, predictive relevance, GoF index.

INTRODUCTION

Partial Least Squares (PLS) is the dominant approach to establish rigor in complex models as it can avert the limitations of covariance based Structural Equation Modeling (CBSEM) with regard to distributional properties, measurement level, sample size, model identification and factor indeterminacy (Chin 1998a, 2010; Wetzels et al. 2009; Akter et al. 2010, 2011). We define complex models as the larger model with many latent variables and indicators, such as, a model with 10 or more constructs and 50 or more items (Chin 2010). In this particular case, using CBSEM causes difficulties to handle such larger models "due to the algorithmic nature requiring inverting of matrices" (Chin 2010, P. 661). As such, PLS path modeling is widely adopted to estimate large complex models with small sample size because it can remove the uncertainty of improper solutions. Although it is suitable for complex models; however, to establish rigor in such modeling this study investigates the implications of power analysis, predictive relevance and GoF index.

The PLS approach to Structural Equation Models, also known as PLS Path Modeling (PLS-PM) is regarded as a component based SEM (Chin et al. 2003; Tenenhaus 2008) to model multiple causes and multiple indicators of a single latent variable, and to latent path models (Wold 1975, 1982, 1985). In contrast to ML (maximum likelihood) approaches, PLS path modeling is able to estimate large number of latent variables and indicators with a small sample size (Jöreskog and Yang 1996; Kelava et al. 2008; Klein and Moosbrugger 2000; Moosbrugger et al. 2009). Indeed, PLS path modeling is a favorable technique for estimating complex models because it can ensure more theoretical parsimony and less model complexity (Chin 2010; Edwards 2001; Law et al. 1998; MacKenzie et al. 2005; Wetzels et al. 2009). As such, it is used to develop large, complex, hierarchical models through the repeated use of manifest variables (Guinot et al. 2001; Lohmöller 1989; Noonan and Wold 1982; Tenenhaus et al. 2005; Wold 1982). For instance, if a second-order construct consists of five underlying first order constructs, each with four manifest variables, the second-order construct can be estimated using all 20 (5*4) manifest variables of the underlying first-order constructs. According to Wetzels et al. (2009), *"This approach also allows us to derive the (indirect) effects of lower-order constructs, or dimensions, on outcomes of the higher-order construct."* Furthermore, PLS path modeling provides more accurate estimates of mediating and moderating effects by accounting for the measurement error that attenuates the estimated relationships and improves the validation of theories (Chin et al. 2003; Helm et al. 2010; Henseler & Fassott 2010). Overall, PLS path modeling provides robust solutions, especially when the objective is prediction, the model is relatively complex, the sample size is small, and the phenomenon under study is new or changing (Chin & Newsted 1999).

Though PLS path modeling gains its prominence for estimating complex models, however, such models require rigorous assessment using power analysis, predictive relevance and GoF index to provide evidence supporting the research model. First, power analysis (1-β) is necessary because "the stability of the estimates can be affected contingent on the sample size" (Chin 1998, p. 305). In this regard, Dijkstra (1983) and Schneeweiss (1993) discuss the magnitude of standard errors for PLS based complex models resulting from not using enough observations (consistency) and indicators for each latent variable (consistency at large). Second, predictive relevance (Q^2) is critical to assess the predictive validity of a complex model (Stone 1974; Geisser 1975; Fornell and Cha 1994; Chin 1998a). It refers to "a synthesis of cross validation and function fitting with the perspective that the prediction of observables is of much greater relevance than the estimation of what are often artificial construct – parameters" (cf. Chin 2010, p. 679; Geisser 1975, p. 320). Finally, GoF (Goodness of Fit) index is crucial to assess the global validity of a PLS based complex model (Tenenhaus et al. 2005). It is defined as the geometric mean of the average communality and average R^2 for all endogenous constructs. Overall, the importance of these techniques have been evidenced in numerous studies, however, there is a paucity of research in the IS domain which have synthesized their applications under one umbrella. Thus, the objective of this study is to establish rigor in PLS path modeling by investigating the implications of power analysis, predictive relevance and GoF index in large complex models.

The remainder of the paper is organized as follows: The next section explores the roles of power analysis, predictive relevance and GoF index in large complex models, then we discuss the overall findings of the study and finally, we conclude the paper with limitations and future research directions.

POWER ANALYSIS

Power (1-β) can be defined as "the probability of rejecting H_0, when H_1 is true" (Larsen and Marx 1981). In other words, power is the probability of obtaining a statistically significant result (H_1), that is, successfully rejecting the H_0 (Cohen 1988). In positivist IS research, the importance of power analysis lies in establishing facts under study by successfully rejecting H_0, accepting H_1 and making decisions on IT artifacts. However, power is less understood and less explored in IS domain (Goodhue et al. 2007; Baroudi and Orlikowski 1989; Sawyer and Ball 1981; Mazen et al.1987).

In developing and testing complex models using PLS path modeling, power analysis is important to validate the implications of sample sizes. Though it is generally assumed that "sample size is less important in the overall model" (Falk & Miller 1992, p. 93), however, adequate sample size is important to improve overall estimates and reduce standard errors (Hui & Wold 1982; Marcoulides and Saunders 2006). Specifically, if small sample sizes (N= 20) were used in large complex models, it would not detect low valued structural path coefficients (β = 0.20) until large sample sizes (N>150) are used (Chin & Newsted 1999, p. 333). Besides, in case of moderately non-normal data, "a markedly larger sample size is needed despite the inclusion of highly reliable indicators in the model" (Marcoulides and Saunders 2006, p. vi). These findings are consistent with Joreskog and Wold's observations (1982. P. 266) which highlight that, "PLS estimates are asymptotically correct in the joint sense of consistency (large number of cases) and consistency at large (large number of indicators for each latent variable)." Thus, adequate sample size is necessary to achieve power in PLS based estimates in order to ensure rigor in complex modeling.

The power dynamics depend on three parameters: the significance level (α) of the test, the sample size (N) of the study and the effect size (ES) of the population (Cohen 1988). In order to assess the adequacy of sample size of large complex models, the power analysis should be conducted on the portion of the model with the largest number of predictors (Chin & Newsted 1999). Though early researchers used to rely on power charts (see, e.g., Scheffé, 1959) and power tables (e.g., Cohen, 1988), now these are supplemented by PC based efficient, precise, and easy-to-use power analysis programs (Goldstein 1989) such as, G*Power 3.1.2 (Faul et al. 2009). The general convention is that the power of a statistical test should be at least 0.80 (Cohen 1988, p. 56). Thus, high power (> 0.80) indicates that there is high degree of probability of producing significant results when the relationship is truly significant. It also proves that the study has adequate confidence on the hypothesized relationships in the research model. Using G*Power 3.1.2 (Faul et al. 2009), Table 1 & Figure 1 estimate the power of a large complex model with 15 constructs, 0.05 significance level (α) and the effect size (ES) of 0.15. It shows that the power of the overall model increases with the increased number of sample sizes. Thus, it is evident that an adequate sample size leads to an adequate power, which is ultimately necessary for establishing IT artifacts in a large complex model.

Sample size	Power
20	0.19
40	0.48
60	0.66
80	0.79
100	0.87
120	0.92
140	0.96
160	0.97

Table 1: sample size and power

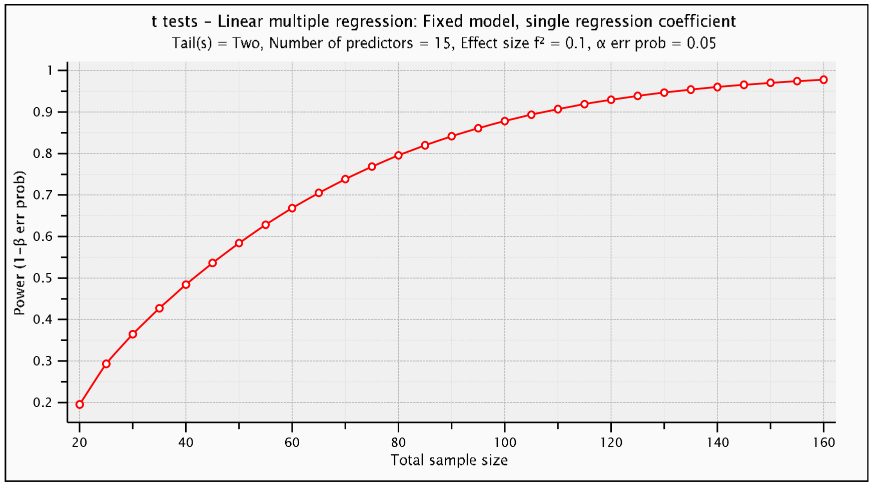

Figure 1: Statistical power in a complex model

PREDICTIVE RELEVANCE

In addition to the size of R^2, the predictive sample reuse technique (Q^2) can effectively be used as a criterion for predictive relevance (Stone 1974; Geisser 1975; Fornell and Cha 1994; Chin 2010). Based on blindfolding procedure, Q^2 evaluates the predictive validity of a large complex model using PLS. While estimating parameters for a model under blindfolding procedure, this technique omits data for a given block of indicators and then predicts the omitted part based on the calculated parameters. Thus, Q^2 shows how well the data collected empirically can be reconstructed with the help of model and the PLS parameters (Fornell & Cha 1994).The predictive measure for the block is based on the following parameters:

$$Q^2 = 1 - \frac{\sum_D E_D}{\sum_D O_D}$$

Where,

E = The sum of squares of prediction error
O = The sum of squares error using the mean for prediction
D = Omission distance

Q^2 can be obtained using two different types of prediction techniques, that is, *cross validated communality* and *cross validated redundancy*. The first one is obtained by predicting data points using latent variable score, whereas the latter one is obtained by predicting the questionable blocks using the latent variables used for prediction. Chin (2010) suggests using the latter to estimate the predictive relevance of a large complex model.

Q^2 is generally estimated using an omission distance of 5-10 under existing PLS software packages. The rule of thumb indicates that a cross validated redundancy $Q^2 > 0.5$ is regarded as a predictive model (Chin 2010). For illustrative purpose, this study estimates cross validated redundancy Q^2 of a large complex model depicted in Fig. 2. In this model, IT continuance is predicted by 10 latent variables, in which 9 are reflective constructs and the other one is formative (control variables). Among all the latent variables, perceived usefulness, perceived quality, perceived trust and social influence were found significant. Using an omission distance of 7, the study obtains a Q^2 of 0.636 which is an indicative of a highly predictive model (see Fig. 2). This finding indicates that prediction of observables or potential observables is of much greater relevance than the estimation of what are often artificial construct parameters (Geisser 1975).

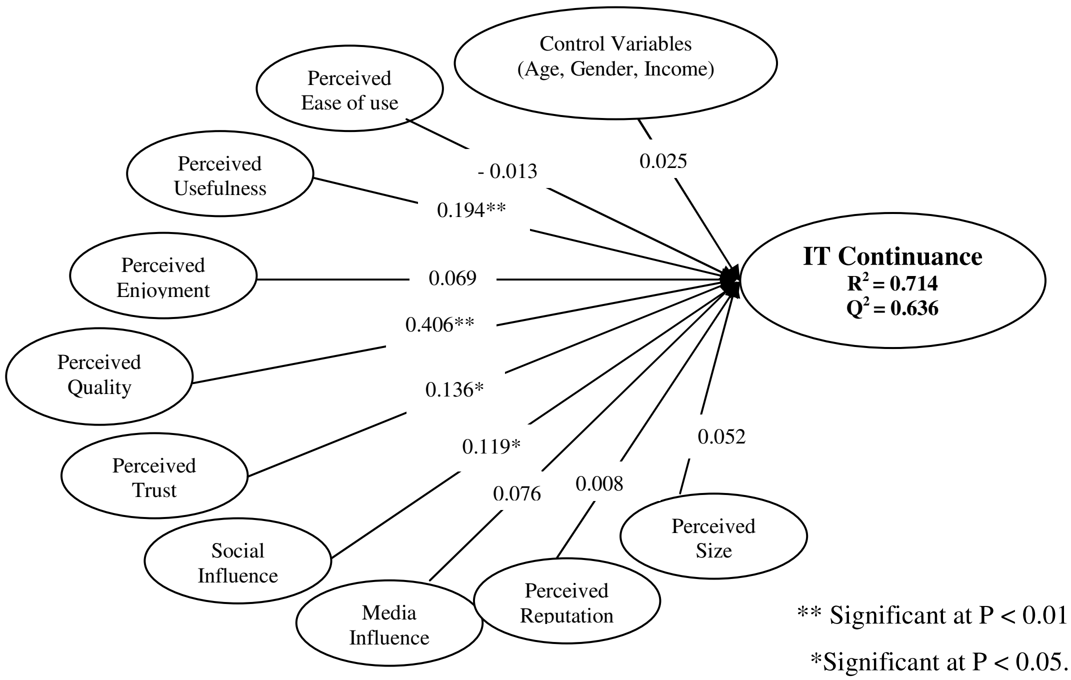

Figure 2: Q^2 in a complex model

GOF INDEX

Goodness of Fit (GoF) index is defined as the geometric mean of the average communality and average R^2 for all endogenous constructs (Tenenhaus et al. 2005). It can be used to determine the overall prediction power of the large complex model by accounting for the performance of both measurement and structural parameters. According to Chin et al. (2010, p. 680), “The intent is to account for the PLS model performance at both the measurement and the structural model with a focus on overall prediction performance of the model”. Though the index is suitable for evaluating reflective indicators, however it can be applied for formative indicators knowing the fact that it would increase the predictiveness of the inner model at the cost of the outer model (Chin 2010). As such, GoF index is applied for both reflective and formative latent variables in a complex case as it provides a measure of overall fit (Vinzi et al. 2010). This index is suggested by Tenenhaus et al. (2005) for assessing the global validity of PLS based complex models. As Tenenhaus et al. note (p. 173), “As a matter of fact, differently from SEM-ML, PLS path modeling does not optimize any global scalar function so that it naturally lacks of an index that can provide the user with a global validation of the model (as it is instead the case with χ^2 and related measures in SEM-ML). The GoF represents an operational solution to this problem as it may be meant as an index for validating the PLS model globally”.

The GoF index is bounded between 0 and 1. Because of the descriptive nature of GoF index, there is no inference based criteria to assess its statistical significance (Vinzi et al. 2010). However, Wetzels et al. (2009) suggest using 0.50 as the cut off value for communality (Fornel and Larcker 1981) and different effect sizes of R^2 (Cohen 1988) to determine GoF_{small} (0.10), GoF_{medium} (0.25) and GoF_{large} (0.36) (see Table 2). These may serve as baselines for validating the PLS based complex models globally.

For the model depicted in Fig. 2, this study obtains a GoF value of 0.750, which exceeds the cut-off value of 0.36 for large effect sizes of R^2 (Cohen 1988). It indicates that the model has a better prediction power in comparison with the baseline values (GoF criteria). This finding adequately validates the complex PLS model globally. It may be noted that GoF index can be estimated for both PLS path modeling and CBSEM.

GoF	GoF criteria
GoF = $\sqrt{\overline{communality} \times \overline{R}^2}$ Range of GoF value: GoF = (0 < GoF < 1)	Communality = 0.50 (Fornel and Larcker 1981) R^2 effect: Small = 0.02, Medium = 0.13, Large = 0.26 (Cohen 1988) Thus, $GoF_{small} = \sqrt{0.5 * 0.02} = 0.10$ $GoF_{medium} = \sqrt{0.5 * 0.13} = 0.25$ $GoF_{large} = \sqrt{0.5 * 0.26} = 0.36$

Table 2: GoF Index and its criteria

DISCUSSION AND FUTURE RESEARCH DIRECTIONS

In validating complex models, PLS path modeling clearly surpasses CBSEM in any settings (exploratory or confirmatory) because of its flexible or soft modeling assumptions. According to Lohmoller (1989, p. 64), "*It is not the concepts nor the models nor the estimation techniques which are 'soft', only the distributional assumptions."* Thus, scholars (e.g., Fornell & Bookstein,1982; Hulland, 1999) believe that PLS path modeling is more suitable for real world applications and more advantageous to employ when models are complex. As Chin (2010, p. 661) notes, *"It is under this backdrop of high complexity that PLS, regardless of whether applied under a strong substantive and theoretical context or limited/exploratory conditions, comes to the fore relative to CBSEM.*" As such, to establish rigor in evaluating PLS based complex models, this study synthesizes the important roles of statistical power analysis, predictive relevance and GoF index.

The study focuses on power analysis as it evaluates the robustness of a large complex model by assessing its ability to reject a false null hypothesis (H_0). Specifically, the power of a statistical test is defined as rejecting the false null hypothesis (H_0) when the alternative hypothesis (H_1) is true (Cohen 1988). It also assesses the probability of finding significant associations among the latent variables in the structural model when the relationships are there ((Baroudi and Orlikowski 1989; Sawyer and Ball 1981; Mazen et al. 1987).

The study discusses predictive sample reuse technique (or, Q^2) as it is able to assess the predictive validity of a large complex model using blindfolding procedure (Stone 1974; Geisser 1975; Fornell and Cha 1994; Chin 1998a). It shows how well the data collected empirically can be reconstructed with the help of model and the PLS parameters (Fornell & Cha 1994).

Finally, the study presents GoF index as it assesses the global validity of PLS based complex model. Specifically, GoF index provides a measure of overall model fit by using the geometric mean of average communality and average R square. The study provides a set of guidelines for estimating GoF criteria to validate complex models.

Overall, PLS path modeling enjoys specific advantages over CBSEM in complex modeling; however, it needs to address some critical challenges to reap its full benefits as the ultimate technique for complex modeling. For example, first, PLS path modeling should have the flexibility of imposing constraints on model coefficients (weights, loadings, path coefficients) in order to specify any information or conjectures available a priori in estimating model parameters (Vinzi et al. 2010). Second, it should allow specific treatment of categorical variables, outliers, non linearity and mutual causality both in measurement and structural models. We believe these challenges represent fascinating areas for future research to establish PLS path modeling as a principal paradigm for complex modeling.

CONCLUSION

PLS path modeling gains its prominence for estimating complex multivariable relationships among observed and latent variables. It is a component based approach to structural equation modeling that estimates both measurement models and structural model. It is well known as a soft modeling approach which is more oriented to optimizing predictions. It is suitable for complex modeling because *"PLS models do not break down as do MLE models especially when sample sizes are 100 or less, there are 2 indicators per construct and independent variable correlations are low"* (Hulland 2010, p. 322). As such, PLS path modeling is more robust for real world applications and more advantageous to employ when models are complex. However, to establish rigor in such modeling, this study necessitates taking into account the critical roles of power analysis, predictive relevance and GoF index.

REFERENCES

1. Akter, S., D'Ambra, J. & Ray, P. (2011). Trustworthiness in mHealth Information Services: An Assessment of a Hierarchical Model with Mediating and Moderating Effects using Partial Least Squares (PLS), *Journal of the American Society for Information Science and Technology, 62(1), 100-116*
2. Akter, S., D'Ambra, J. & Ray, P. (2010). Service Quality of mHealth platform: Development and Validation of a Hierarchical Model using PLS, *Electronic Markets,* 20(3), 209-227.
3. Baroudi, J., and Orlikowski, W. (1989). The Problem of Statistical Power in MIS Research, *MIS Quarterly,* 13, 1, 87-106.
4. Chin, Wynne W. (1998a). Commentary: Issues and Opinion on Structural Equation Modeling, *MIS Quarterly,* 22 (March), vii–xvi.
5. Chin, Wynne W. (1998b). The Partial Least Squares Approach for Structural Equation Modeling, in *Modern Methods for Business Research,* George A. Marcoulides, ed., Mahwah, NJ: Lawrence Erlbaum, 295–336.
6. Chin, W. W., and Newsted, P. R. (1999) Structural Equation Modeling analysis with Small Samples Using Partial Least Squares, in R. Hoyle (ed.), S*tatistical Strategies for Small Sample Research*, Thousand Oaks, CA: Sage Publications, pp. 307-341.
7. Chin, W. W., Marcolin, B., and Newsted. P. (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.
8. Chin, W. W. (2010). How to write up and report PLS analyses, in *Handbook of Partial Least Squares: Concepts, Methods and Application.* Esposito Vinzi, V.; Chin, W.W.; Henseler, J.; Wang, H. (Eds.), Springer, Germany, 645-689.
9. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences,* Hillside, NJ: Lawrence Erlbaum.
10. Dijkstra, T. Some Comments on Maximum Likelihood and Partial Least Squares Methods, *Journal of Econometrics,* 22, 1983, 67-90.
11. Edwards, J. R. (2001). Multidimensional constructs in organizational behavior research: an integrative analytical framework. *Organizational Research Methods*, 4, 2, 144–192.
12. Faul, F., Erdfelder, E., Buchner, A., & Lang, G. A. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, 41, 4, 1149– 1160.
13. Fornell, C., and Bookstein, F. L. (1982). Two Structural Equations Models: LISREL and PLS Applied to Consumer Exit-Voice Theory, *Journal of Marketing Research,* 19, 4, 440-452.
14. Fornell, C., and Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error, *Journal of Marketing Research,* 18 (1), 39-50.
15. Fornell, C., & Cha, J. (1993). Partial Least Squares (PLS), Unpublished working paper. Ann Arbor: University of Michigan Business School.
16. Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of the American Statistical Association*, 70, 350, 320-328.
17. Goldstein, R. (1989). Power and sample size via MS/PC-DOS computers. *American Statistician*, 43, 253-262.

18. Guinot, C., Latreille, J., and Tenenhaus, M. (2001). PLS Path Modeling and Multiple Table Analysis: Application to the cosmetic Habits of Women in Ile-de-France, *Chemometrics and Intelligent Laboratory Systems* (58:2), pp. 247-259.
19. Hulland, J. Ryan, M.J. and Rayner, R.K. (2010). Modeling customer satisfaction: a comparative performance evaluation of covariance structure analysis versus partial least squares, in *Handbook of Partial Least Squares: Concepts, Methods and Application.* Esposito Vinzi, V.; Chin, W.W.; Henseler, J.; Wang, H. (Eds.), Springer, Germany, 307-325.
20. Jöreskog, K.G., Yang, F.: Nonlinear structural equation models: the Kenny–Judd model with interaction effects. In: Marcoulides, G.A., Schumacker, R.E. (eds.) Advanced Structural Equation Modeling: Issues and Techniques, pp. 57–88. Lawrence Erlbaum, Mahwah (1996)
21. Kelava, A., Moosbrugger, H., Dimitruk, P., Schermelleh-Engel, K. (2008). Multicollinearity and missing constraints: a comparison of three approaches for the analysis of latent nonlinear effects. Methodology: European Journal of Research Methods for the Behavioral and Social Sciences , 4, 51–66
22. Klein, A.G., Moosbrugger, H. (2000) Maximum likelihood estimation of latent interaction effects with the LMS method. Psychometrika, 65, 457–474.
23. Law, K. S., Wong C., and Mobley, W. H. (1998) Toward a Taxonomy of Multidimensional Constructs, *Academy of Management Review,* 23, 4, 741-755.
24. Lohmöller, J.-B. (1989). *Latent Variable Path Modeling with Partial Least Squares*, Heidelberg: Physica-Verlag.
25. MacKenzie, S. B., Podsakoff, P. M., and Jarvis, C. B. (2005) The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions, *Journal of Applied Psychology,* 90, 4, 710-730.
26. Marcoulides, G. A., C. Saunders. (2006). PLS: A silver bullet? *MIS Quarterly.* 30(2) iii–ix.
27. Mazen, A., M. Magid, M. Hemmasi, M. F. Lewis. (1987). Statistical power in contemporary management research. *Academy of Management Journal,* 30, 2, 369–380.
28. Moosbrugger, H., Schermelleh-Engel, K., Kelava, A., Klein, A.G.: Testing multiple nonlinear effects in structural equation modeling: a comparison of alternative estimation approaches. In: Teo, T., Khine, M.S. (eds.) Structural Equation Modeling in Educational Research: Concepts and Applications, 103–136. Sense, Rotterdam (2009)
29. Noonan, R., and Wold, H. (1982). PLS Path Modeling with Indirectly Observed Variables: A Comparison of Alternative Estimates for the Latent Variable, in *Systems Under Indirect Observation, Part II*, K.G. Jöreskog and H. Wold (eds.), Amsterdam: North Holland, 75-94.
30. Sawyer, A. G., A. D. Ball. 1981. Statistical power and effect size in marketing research. *Journal of Marketing Research* 18, 3, 275–290.
31. Schneeweiss, H. (1993). Consistency at Large in Models with Latent Variables, in *Statistical Modelling and Latent Variables,* K. Haagen, D. J.Bartholomew, and M. Deistler (eds.), Elsevier, Amsterdam.
32. Scheffé, H. (1959). *The analysis of variance*. New York: Wiley.
33. Stone, M. (1974). Cross validatory choice and assessment of statistical predictions, *Journal of the Royal Statistical Society*, 36, 2, 111-147
34. Tenenhaus, M., Vinzi, V. E. Chatelin, Y-M., and Lauro, C. 2005. "PLS Path Modeling," *Computational Statistics and Data Analysis,* 48, 1, 159-205.
35. Vinzi, V.E., Trinchera, L & Amato, S. (2010). PLS path modeling: From foundations to recent developments and open issues for model assessment and improvement , in *Handbook of Partial Least Squares: Concepts, Methods and Application.* Esposito Vinzi, V.; Chin, W.W.; Henseler, J.; Wang, H. (Eds.), Springer, Germany, 47-82.
36. Wold, H. (1982). Soft Modeling: The Basic Design and Some Extensions, in *Systems Under Indirect Observation: Causality, Structure, Prediction* (Volume 2), K. G. Jöreskog and H. Wold (eds.), Amsterdam: North Holland, 1-54.
37. Wold, H. (1985). Partial Least Squares," in *Encyclopedia of Statistical Sciences* (Volume 6), S. Kotz and N. L. Johnson (eds.), New York: Wiley, pp. 581-591.
38. Wold, H., 1975. Soft modelling by latent variables: the non-linear iterative partial least squares (NIPALS) approach. In: Gani, J. (Ed.), Perspectives in Probability and Statistics: Papers, in Honour of M.S. Bartlett on the Occasion of his Sixty-8fth Birthday. Applied Probability Trust, Academic, London, 117–142.