Testing Nonlinear Effects in PLS Path Models: A Simulation-based Comparison of Available Approaches

Jörg Henseler.
Nijmegen School of Management, Radboud University, Nijmegen, The Netherlands.
j.henseler@fm.ru.nl
Bradley Wilson.
RMIT University, School of Applied Communication, Melbourne, Australia.
brad.wilson@rmit.edu.au
Theo Dijkstra
University of Groningen. Econometrics & Operations Research. Groningen, Netherlands
theo.dijkstra@sns.nl

Introduction

Gefen, Straub, and Boudreau (2000, p. 6) point out, "SEM has become *de rigueur* in validating instruments and testing linkages between constructs." Along with the development of scientific disciplines, the complexity of hypothesized relationships has steadily increased (Cortina, 1993). Researchers direct their interest more and more from investigating linear effects between constructs towards investigating nonlinear effects. Covariance-based structural equation modeling (CBSEM) has a growing body of research dedicated to the modeling of nonlinear relationships (c.f. Schumacker and Marcoulides, 1998). In constrast, estimating nonlinear effects by means of PLS path modeling has not received any attention for the last two decades. In our contribution, we discuss four approaches to modeling nonlinear effects with PLS, compare their performance by means of a Monte Carlo simulation, and develop guidelines for PLS researchers.

A Brief Review of Nonlinear Modeling Approaches

Herman Wold himself regarded PLS path modeling as readily equipped to estimate nonlinear effects between latent variables (Wold, 1982). The four approaches that have been suggested so far include: (1) Wold's original approach, (2) the product indicator approach (Chin, Marcolin, and Newsted, 1996, 2003), and (3) a two-stage approach (Chin, Marcolin, and Newsted, 2003; Henseler and Fassott, 2007). Moreover, (4) an additional orthogonal approach suggested by Little, Bovaird, and Widaman (2006), is adapted to nonlinear PLS path modeling.

Wold's Original Approach. The first attempt to model nonlinear structural relationships in PLS path models was made by Herman Wold (1982). Although he initially considered only a model with a quadratic term, his approach is generalizable to other nonlinear relations between latent variables. The premise behind Wold's original approach is to take the nonlinearity in the structural model into account within the PLS algorithm. More precisely, during the algorithm, proxy variables for the nonlinear terms are calculated and used to determine the latent variable scores. The PLS algorithm delivers estimates for the latent variable scores by means of an iterative process. This approach requires a modification of the PLS algorithm which until now has not been implemented in any of the leading PLS software[1].

The Product-Indicator Approach. The first approaches using CBSEM to study nonlinear effects were marked by Busemeyer and Jones (1983) and Kenny and Judd (1984). They proposed to build product terms between the indicators of the latent independent variable and the indicators of the latent moderator variable. These product terms serve as indicators of the interaction term. Chin, Marcolin, and Newsted (1996, 2003) were the first to transfer this approach to PLS path modeling. This product indicator approach can easily be adapted to other nonlinear effects, in particular, polynomial effects. In the latter case, polynomial terms can be created by product indicators as well. Whilst the product indicator approach has been considerably difficult to implement in CBSEM context and alternatives have been suggested (Jöreskog and Wang, 1996), it was found to be easily implementable in PLS path modeling. Yet, it remains unclear whether indicators should only be multiplied with themselves in order to form polynomial terms or if indicators should also be multiplied with every other indicator of the respective latent variable.

The Two-Stage Approach. The idea of estimating structural equation models with nonlinear effects between latent variables in two steps, in which the first stage is dedicated to estimation of the measurement model, is not new (c.f. Anderson and Gerbing, 1988). Researchers could use the estimated latent variable scores as the basis for calculating nonlinear terms, e.g. create polynomial terms. Relying on the estimated measurement models, the second stage then delivers parameter estimates of the linear and nonlinear effects by multiple linear regression or by a single indicator PLS path model. The idea of applying a two-stage approach to PLS path modeling in the context of nonlinear effects was initially expressed by Chin, Marcolin, and Newsted (2003). Chin, Marcolin, and Newsted (2003) and Henseler and Fassott (2007) recommend the usage of the two-stage approach to formative exogeneous variables.

The Orthogonalizing Approach. Little, Bovaird, and Widaman (2006) have recently suggested an orthogonalizing approach for modeling interactions among latent variables. The main objective of their approach was to overcome the problems of multicollinearity that often occur when nonlinear terms and linear terms simultaneously enter into multiple regression as independent variables. Although this approach was applied to CBSEM only, it is easily

---

[1] We considered the following PLS software: LVPLS, PLS-Graph (Chin, 1993-2003), SmartPLS, SPAD-PLS.

transferable to PLS path modeling. Basically, the orthogonalizing approach is an extension of the use of residual centering for moderated multiple regressions as described by Lance (1988). Residual centering is essentially a two-stage OLS procedure in which a nonlinear term is regressed onto its respective linear term. The resulting residuals are then used to represent the nonlinear term. The variance of this new orthogonalized nonlinear term contains the unique variance that fully represents the nonlinear effect, independent of the linear effect (see Little, Bovaird, and Widaman, 2006, for an analogue argumentation for interaction effects). As a consequence of the orthogonality of the nonlinear term, the parameter estimates of the linear effects in a model with nonlinear terms are identical to the parameter estimates of the linear effects in a model without the nonlinear terms. Furthermore, residual centering yields a regression coefficient for the nonlinear term that can directly be interpreted as the effect of the nonlinear term on the dependent variable (c.f. Lance, 1988, on interactions) and thus replace the assessment of the increase in the coefficient of determination due to the inclusion of the nonlinear term. From the fact that PLS calculates the latent variable scores as linear combinations of the respective indicators, it can be derived that a nonlinear term that is created in this manner is orthogonal to its constituting latent variable.

Methodology

We implement the PLS algorithm in R 2.3.1 (R Development Core Team, 2006). For the simulation we define an underlying true model as consisting of one exogenous latent variable (with a beta of 0.5), a quadratic effect (with small, medium, large, or zero effect size), and one endogenous latent variable. The latent variables have six indicators and a reliability (Cronbach's Alpha) of 0.9 each, and are estimated in Mode A. Besides the effect size of the quadratic effect, we vary the sample size (with levels of 20, 50, 100, 200, and 500). We create 500 samples (runs) for each condition. In each condition, all four approaches for the analysis of interaction effects between latent variables using PLS path modeling are used to estimate the model. We select the centroid scheme as inner weighting scheme. Three types of PLS estimation outcomes are measured for each run: (1) path coefficient estimates for the single as well as the interaction effect, (2) bootstrap t-values for all effects, and (3) squared correlations between the predicted latent variable scores of the endogenous variable and its true scores.

Discussion

Based on the results it is possible to give recommendations to researchers who want to analyze nonlinear effects using PLS. We do recognize that the choice of approaches should mainly be based on the researcher's objectives. It remains unclear whether our findings will be applicable for formative measures. Future research needs to guide analysts on how to analyze nonlinear effects with formative indicants. Our study was limited to quadratic terms. Further research could strive for the replication of our in progress recommendations for other commonly used nonlinear functions like e.g. exponential or logarithmic effects. Further extensions can be followed by combining polynomial terms with interaction effects.

Bibliography

Anderson, James C., and David W. Gerbing. 1988. "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach." Psychological Bulletin. 103:411-423.

Chin, W. W., Marcolin, B. L., and Newsted, P. N. (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.

Cortina, J. M. (1993). Interaction, nonlinearity, and multicollinearity: Implications for multiple regression. Journal of Management, 19, 915–922.

Gefen, D., D. W. Straub, and M.-C. Boudreau, (2000) "Structural Equation Modeling and Regression: Guidelines for Research Practice," Communications of the AIS, 4(7) 1-77. http://citeseer.ist.psu.edu/gefen00structural.html.

Henseler, J., and Fassott, G. (2007) "Testing Moderating Effects in PLS Path Models: An Illustration of Available Procedures," in Vinzi, V. E., Chin, W. W., Henseler, J., and Wang, H. (eds.) Handbook Partial Least Squares Path Modeling. Springer, Heidelberg, forthcoming.

Kenny, D.A., Judd, C.M. (1984) Estimating the nonlinear and interactive effects of latent variables. Psychological Bulletin, 96, 201 –210.

Lance, C. E. (1988) "Residual Centering, Exploratory and Confirmatory Moderator Analysis, and Decomposition of Effects in Path Models Containing Interactions," Applied Psychological Measurement (12:2), June, pp. 163–175.

Little, T. D., Bovaird, J. A., and Widaman, K. F. (2006) "On the Merits of Orthogonalizing Powered and Product Terms: Implications for Modeling Interactions Among Latent Variables," Structural Equation Modeling (13:4), October, pp. 497–519.

Lohmöller, J.-B. (1989) Latent Variable Path Modeling with Partial Least Squares. Heidelberg: Physica-Verlag.

Schumacker, Randall E. and George A. Marcoulides (1998), Interactions and Nonlinear Effects in Structural Equation Modeling. Mahwah, NJ: Lawrence Erlbaum Associates.

Wold, H. (1982) 'Soft modeling. The basic design and some extensions', in Jöreskog, K.G., & Wold, H. (eds) Systems under indirect observation. Causality, structure, prediction. Part I, 1–54. North-Holland: Amsterdam.