Treating Heterogeneity in PLS Path Modeling Using Latent Class Moderating Effects

Armin Monecke and Friedrich Leisch

Institut für Statistik
Ludwig-Maximilians-Universität München

3rd Workshop on Psychometric Computing, 24.02.2011, Tübingen, Germany

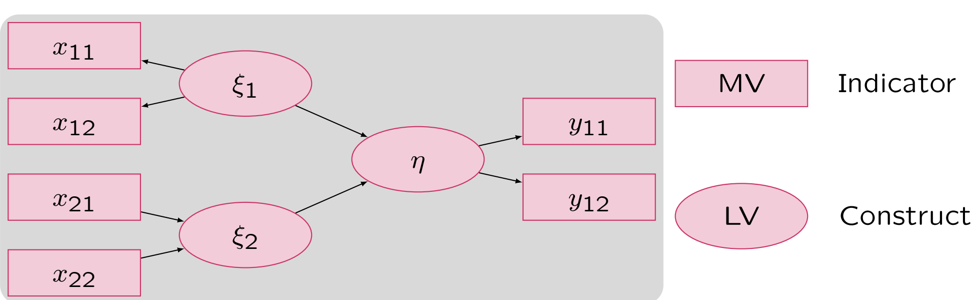
Overview

Outline:

PLS Path Modeling: Kick-Start

PLS Path Modeling consits of basically two steps:

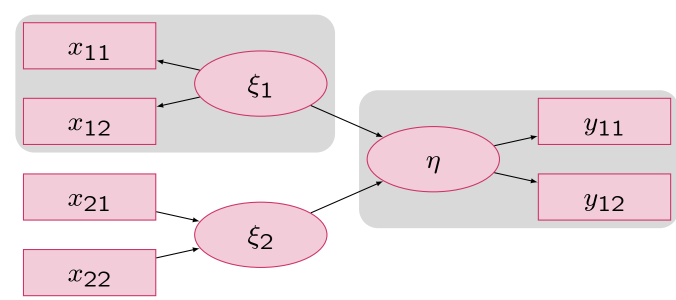
1. Determine factor scores by an iterative procedure, based on the hypothetical model.
2. Use the factor scores to estimate the path coefficients.



PLS Path Modeling: Kick-Start

The relations between MVs and LVs are refered to as measurement or outer model.

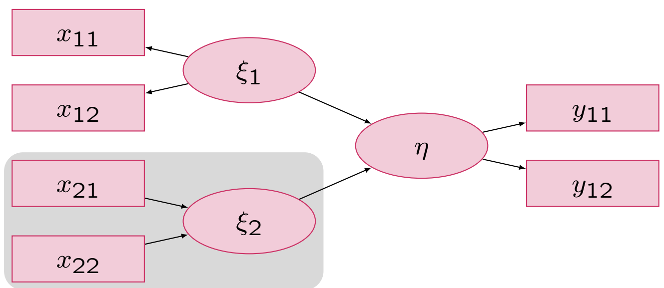
- Reflective measurement for exogenous latent variable ξ_1 and endogenous variable η:



PLS Path Modeling: Kick-Start

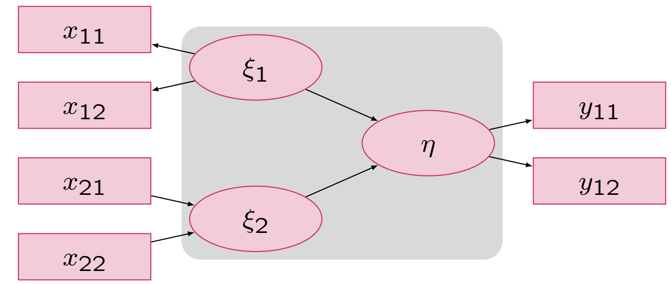
The relations between MVs and LVs are refered to as measurement or outer model.

- Formative measurement for exogenous latent variable ξ_2:



PLS Path Modeling: Kick-Start

Relations between LVs are called structural or inner model.



- Exogenous variables are LVs without predecessors.
- All other LVs are endogenous.

Heterogeneity in PLS Path Models

The assumption that the data is collected from a single homogenous population is often unrealistic or may turn to be false. The presence of heterogeinity not accounted for may lead to biased or erroneaous results.
Within the context of PLS path modelling there are four prominent approaches to deal with heterogeneity:

- Pathmox (Sanchez & Aluja 2007): observed.
- REBUS-PLS (Esposito Vinzi et al. 2008): unobserved.
- FIMIX-PLS (Ringle et al. 2010): unobserved.
- PLS Typological Path Modeling (Squillacciotti 2010): unobserved.

Heterogeneity in PLS Path Models

All four approaches have in common to:

1. fit the global model,
2. find homogenous groups,
3. fit local models for each group.

Thus, the number of path coefficients increases multiplicatively with the number of groups while local sample sizes become smaller, especially, when the resulting grouping is unbalanced.

Why not using the grouping indicator as a moderator variable in the global model?

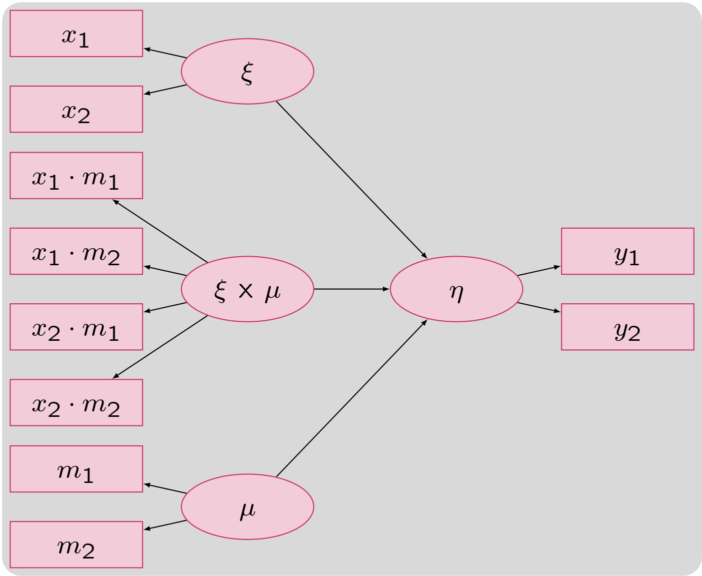
Moderating Effects in PLS Path Models

Approaches for metric moderarator variables:

- Product Indicator (Chin et al. 2003)
- Two-Stage (Henseler & Fassot 2010)
- Hybrid (Wold, 1982)
- Orthogonalizing (Little et al. 2006)

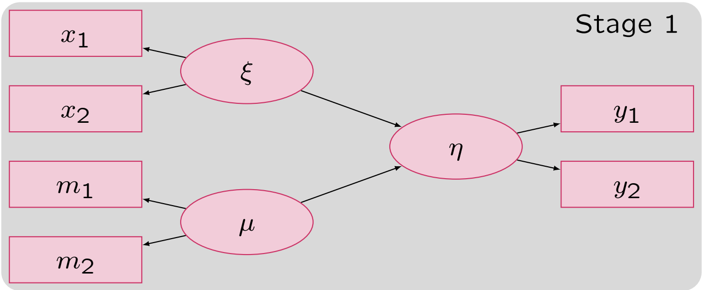
Note: Moderating effects in PLS path modelling remain an open topic, especially, when nominal indicators are involved.

Product Indicator Approach



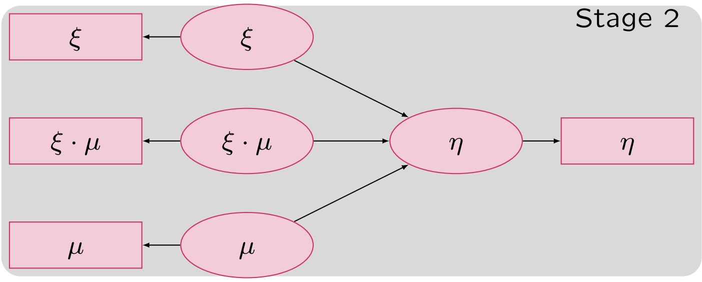
Two-Stage Approach

Stage 1: Run the PLS path model with only the main effects and save the estimated factor scores for further analysis.



Two-Stage Approach

Stage 2: Calculate the interaction term $\xi \cdot \mu$ as the element-wise product of the latent variable scores of the exogenous variable ξ and the moderator variable μ. Then the path coefficients are obtained by a multiple regression on the exogenous variable η with the interaction term and the latent variable scores of ξ and μ as independent variables.



Other Approaches

Hybrid Approach (Wold 1982): Initially designed for estimation of models with nonlinearities in the structural model, e.g. with a quadratic term and later generalized for other nonlinearities – in particular interaction effects.

- Combination of product indicator approach and two-stage approach.
- The product term $\xi \cdot \mu$ is calculated in each iteration of the PLS-algorithm (compare: two-stage approach).
- The product term $\xi \cdot \mu$ is updated in each iteration (compare: product indicator approach).

Other Approaches

Orthogonalizing Approach (Little et al. 2006): Extents the product indicator approach. But instead of the products p_{ij},

$$p_{ij} = x_i \cdot m_j \ \forall\, (i,j),\ i = 1,\ldots,I,\ j = 1,\ldots,J.$$

the residuals ϵ_{ij} (resulting from a linear regression on the main effects' manifest variables) are used.

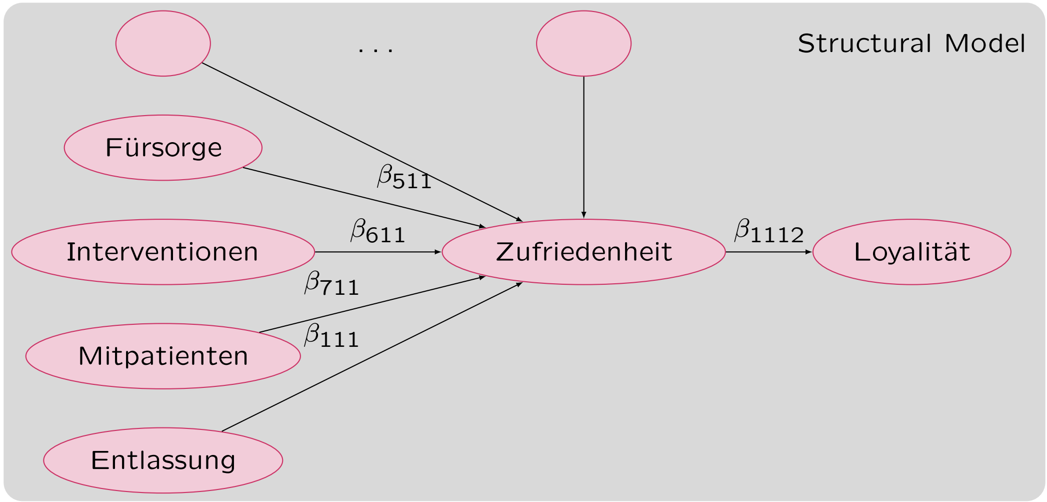
$$p_{ij} = b_{0,ij} + b_{1,ij}x_1 + \cdots + b_{I,ij}x_I + b_{I+1,ij}m_1 + \cdots + b_{I+J,ij}m_J + \epsilon_{ij}$$

Latent Class Moderating Effects

To control the number of parameters we propose the following strategy:

1. Fit the global model and use exploratory model diagnostics to identify set of parameters with large variability, e.g. parallel coordinate plots of bootstapped path coefficients.
2. Model the overdispersion by a convex combination of several coefficient values, e.g. a finite mixture model for this subset of path coefficients. The grouping indicators are determined by a model based clustering on the respective inner models.
3. Refit the global model using the grouping indicators as moderator variables for the respective path coefficients.

Example: Psychosomatic Daycare Center



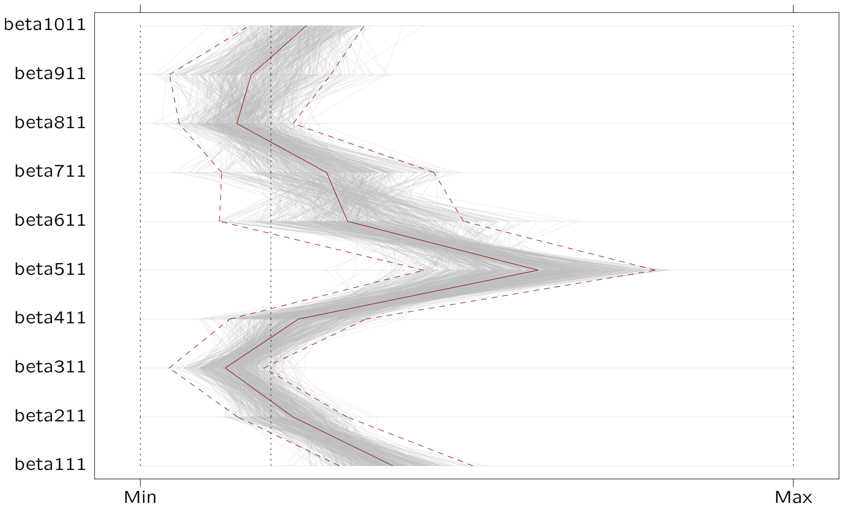
Example: Psychosomatic Daycare Center

The facility surveyed a customer satisfaction study (N=178).

- Only two significant (α=10%) path coefficients affecting 'Gesamtzufriedenheit' (satisfaction) in the gloabal model:
  Fürsorge (care) $\rightarrow$ Geasmtzufriedenheit (0.51)
  Entlassung (release) $\rightarrow$ Geasmtzufriedenheit (0.23).

- A priori, it was asssumed that the patients perception of 'Interventionen' (psychological interventions) would be the main driver for 'Gesamtzufriedenheit' (satisfaction).

Parallel Coordinates: Bootstrap Results



Components of Finite Mixture Model

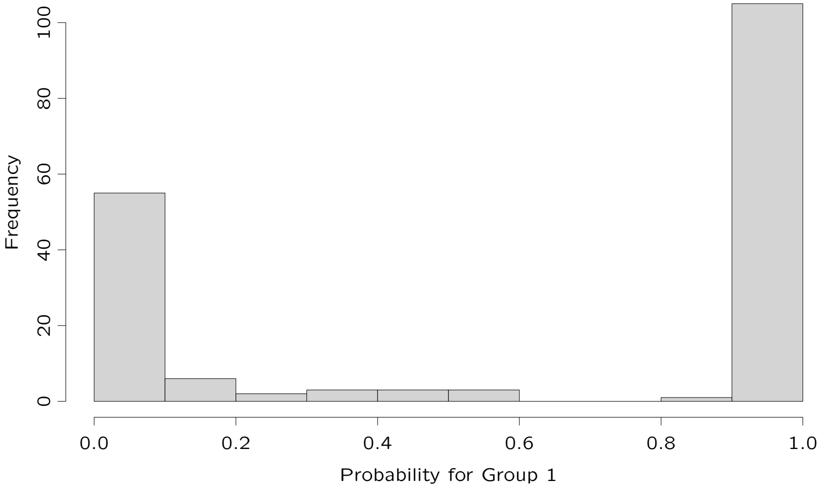
Component 1:

N=109	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.04	0.06	-0.61	0.54
Fuersorge	0.49	0.08	6.21	0.00
Interventionen	0.24	0.10	2.37	0.02
Mitpatienten	0.14	0.08	1.65	0.10
Interventionen:Mitpatienten	0.04	0.04	1.01	0.31

Component 2:

N=69	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.07	0.00	27.22	0.00
Fuersorge	0.97	0.00	324.07	0.00
Interventionen	-0.01	0.00	-1.99	0.05
Mitpatienten	0.00	0.00	1.07	0.29
Interventionen:Mitpatienten	0.01	0.00	4.21	0.00

Histogram: Posterior Group Probabilities



Product Terms

To construct the moderating effect we used a combination of

- the **two-stage** and
- the **orthogonalizing** approach.

$$p_{11} = \text{GroupProbability} \cdot \text{Fürsorge}$$
$$p_{12} = \text{GroupProbability} \cdot \text{Interventionen}$$

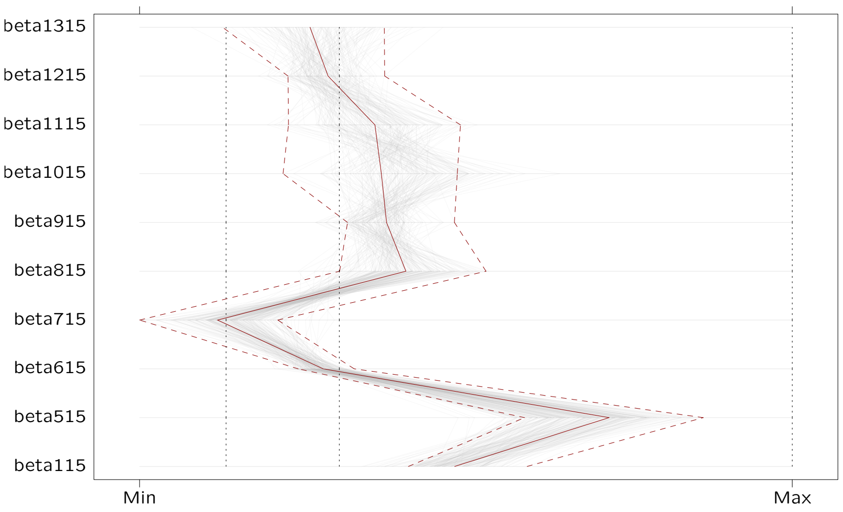
Resulting in the residuals of:

$$p_{11} = b_1 \cdot \text{GroupProbability} + b_2 \cdot \text{Fürsorge} + \epsilon_{11}$$
$$p_{12} = b_1 \cdot \text{GroupProbability} + b_2 \cdot \text{Interventionen} + \epsilon_{12}$$

Path Coefficients (after & before refitting)

Path	Estimate(refit)	Estimate
Entlassung -> Gesamtzufriedenheit	0.25	0.23
Fuersorge -> Gesamtzufriedenheit	0.60	0.51
Group -> Gesamtzufriedenheit	-0.04	–
Group*F -> Gesamtzufriedenheit	-0.27	–
Group*I -> Gesamtzufriedenheit	0.15	–
Group*M -> Gesamtzufriedenheit	0.10	–
Interventionen -> Gesamtzufriedenheit	0.09	0.15
Mitpatienten -> Gesamtzufriedenheit	0.08	0.11
Privatleben -> Gesamtzufriedenheit	-0.02	-0.07
Service -> Gesamtzufriedenheit	-0.06	-0.04

Parallel Coordinates: Bootstrap Results



Summary & Open Questions

+ The proposed strategy enables the researcher to control the number of parameters.
+ Good interpretabilty of the moderation effects.
– The selection of relevant sets of variables for the model based clustering requires a lot of expertise from the researcher.

- Can the Bootstrap Confidence Intervalls still be trusted?
- How to deal with more than two latent groups?

References

A Comparison of Approaches for the Analysis of Interaction Effects Between Latent Variables Using Partial Least Squares Path Modeling,

Henseler, J. and Chin, W. W. (2010). Structural Equation Modeling, 17:82.

Testing Moderating effects in PLS Path Models: An illustration of available procedures,

Henseler, J. and Fassot, G. (2010). In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), Handbook of Partial Least Squares.

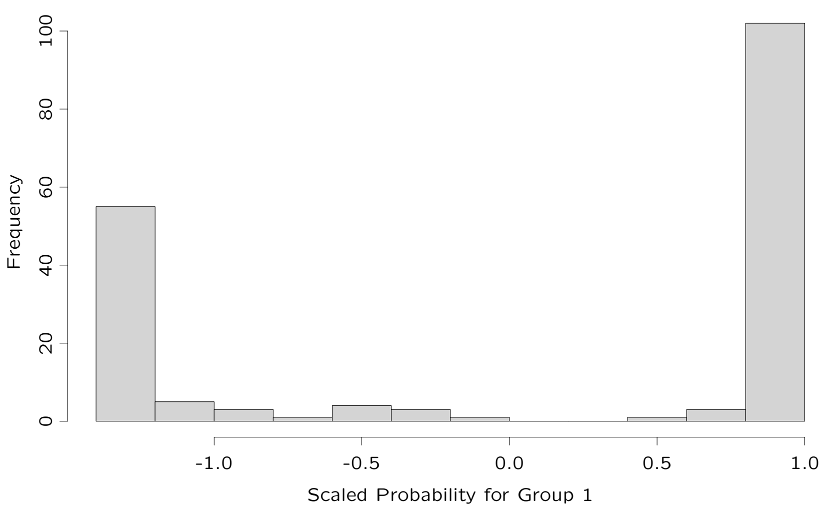
On the merits of orthogonolizing powered and product terms: Implications for modeling interactions among latent variables,

Little, T. D., Bovaird, J.A., & Widman, K. F. (2006). Structural Equation Modeling, 13, 497-519.

Fitting Finite Mixtures of Generalized Linear Regressions in R,

Grün, B. and Leisch, F. (2007). Computational Statistics & Data Analysis 51:11, p.5247–5252.

Histogram: Scaled PGPs



References

semPLS: Structural Equation Modeling Using Partial Least Squares,

Monecke, A. and Leisch, F. (2010). R package version 0.8-7. 2010. http://cran.r-project.org/web/packages/semPLS/.