Partial Least Squares Structural Equation Modeling Ib

Model estimation

Morten Berg Jensen

Department of Economics and Business

February 22, 2024

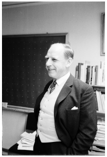
Outline

Outcome

This lecture will help you to understand

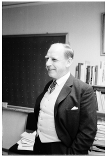
- The formalisation of the relationships among latent variables as well as between latent and manifest variables
- The iterative PLS-SEM algorithm
- How to interpret the model and its statistical properties

Herman Wold (1908 - 1992) I



- Swedish Professor of statistics at Uppsala University
- Considered the “grandfather” of the family of partial least squares methods
- Also known for Cramer-Wold theorem, Wold decomposition, macroeconomics and contributions to utility theory
- His work on causality has been characterized as being decades ahead of time (see e.g. Judea Pearl’s book “Causality” (2nd edition))
- Member of Nobel Economic Science Prize Committe from 1968 - 1980

Herman Wold (1908 - 1992) II



- The development of the PLS algorithm in 1977 was founded on two other developments he made
    - Fixed-Point algorithm: Iterative ordinary least squares algorithm to estimate coefficients of a system of simultaneous equations
    - NIPALS (Non-linear Iterative Partial Least Squares) algorithm: Which among other things can calculate principal components and canonical correlations

Structural equation modeling (SEM)

- "Structural": Each inner equation should represent a hypothesized causal relationships between a set of (latent) variables, and the form of each equation conveys the assumptions that the analyst has asserted
- Two main ways of estimating SEM models

1. Based on the covariance matrix of the variables: Developed primarily by Karl Jöreskog, a doctoral student of Herman Wold
2. Using the PLS-SEM algorithm (which we will soon visit) developed by Herman Wold

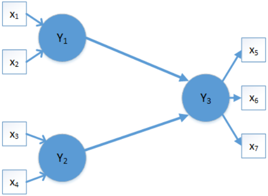
Inner relations I

Let N be the sample size, K the number of indicators/observed variables and H be the number of latent variables

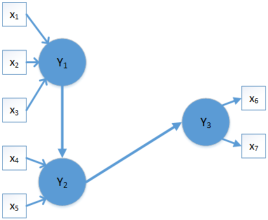
$$Y = Y \cdot \underbrace{(P + t(P))}_{C} + \nu \tag{1}$$

- Y is a $N \times H$ matrix of latent variables
- P is a $H \times H$ matrix of inner weights/path coefficients
  - Is upper diagonal
  - Restrictions on P determines the inner model
- ν is a $N \times H$ matrix of inner errors

Inner relations II



$$P = \begin{bmatrix} 0 & 0 & p_{13} \\ 0 & 0 & p_{23} \\ 0 & 0 & 0 \end{bmatrix}$$



$$P = \begin{bmatrix} 0 & p_{12} & 0 \\ 0 & 0 & p_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

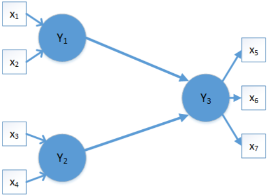
Figure: Different constraints on the inner weights

Outer relations - correlation weights (mode A) I

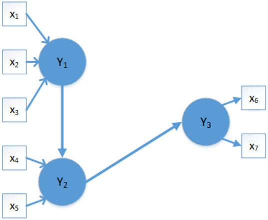
$$X = Y \cdot L + \varepsilon \tag{2}$$

- X is a $N \times K$ matrix of standardized indicators
- Y is a $N \times H$ matrix of latent variables
- L is a $H \times K$ matrix of correlation weights/loadings
- ε is a $N \times K$ matrix of outer errors

Outer relations - correlation weights (mode A) II



$$L = \begin{bmatrix} l_{11} & l_{12} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & l_{23} & l_{24} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & l_{35} & l_{36} & l_{37} \end{bmatrix}$$



$$L = \begin{bmatrix} l_{11} & l_{12} & l_{13} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & l_{24} & l_{25} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & l_{36} & l_{37} \end{bmatrix}$$

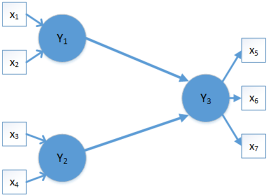
Figure: Different constraints on the correlation weights

Outer relations - regression weights (mode B) I

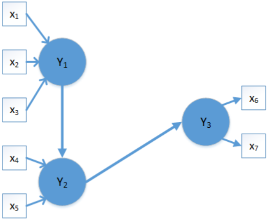
(3) $$Y = X \cdot \Omega + \delta$$

- Y is a $N \times H$ matrix of latent variables
- X is a $N \times K$ matrix of standardized indicators
- Ω is a $K \times H$ matrix of regression weights/weights
- δ is a $N \times K$ matrix of validity errors

Outer relations - regression weights (mode B) II



$$\Omega = \begin{bmatrix} \omega_{11} & 0 & 0 \\ \omega_{21} & 0 & 0 \\ 0 & \omega_{32} & 0 \\ 0 & \omega_{42} & 0 \\ 0 & 0 & \omega_{53} \\ 0 & 0 & \omega_{63} \\ 0 & 0 & \omega_{73} \end{bmatrix}$$



$$\Omega = \begin{bmatrix} \omega_{11} & 0 & 0 \\ \omega_{21} & 0 & 0 \\ \omega_{31} & 0 & 0 \\ 0 & \omega_{42} & 0 \\ 0 & \omega_{52} & 0 \\ 0 & 0 & \omega_{63} \\ 0 & 0 & \omega_{73} \end{bmatrix}$$

Figure: Different constraints on the regression weights

Assumptions

- The relations in the model are linear (as given in the previous slides)
- We do not need to assume a specific distribution of our variables
- For identification we assume
    - $Var(Y_j) = 1$ for $j = 1, ..., H$
    - $E(Y_j) = 0$ for $j = 1, ..., H$

The PLS-SEM algorithm

1. Initialization: Standardize all indicators, and create an initial (poor) approximation of the LV scores
2. Inner approximation: Estimate inner weights (using one of three weighting schemes) and use these weights to approximate the LV's
3. Outer approximation: Estimate temporary correlation and/or regression weights (depending on type of measurement model) to approximate LV's as weighted sums of their indicators
4. Convergence: Iterate step 2 and 3 until there is convergence of temporary correlation and regression weights
5. Final estimates: Estimate correlation weights, regression weights and inner weights using standardized LV's

- See also small R script for illustration

PLS-SEM algorithm - step 1: initialization

- Standardize all indicators, X
- Choose initial weights by setting $\omega_{ij} = 1$ for all i, j and let $W^0 = \Omega$
- Make approximate LV's, $\hat{Y}^{*0} = X \cdot W^0$

PLS-SEM algorithm - step 2: inner approximation

- We can chose between three different estimates of the inner weighting schemes, $\hat{P}^{i-1} = f(\hat{Y}^{*(i-1)}), i = 1, \ldots$
  - Factorial: The inner weights are the correlation between adjacent approximate LV's
  - Centroid: The inner weights are the sign of the correlation between adjacent approximate LV's
  - Path: The inner weights are the regression coefficients in an OLS regression with the endogenous approximate LV as dependent and its exogenous approximate LV' as independent variables (preferred)
- Update approximate LV's,

$$\hat{Y}^{i} = \hat{Y}^{*(i-1)} \cdot (\hat{P}^{i-1} + t(\hat{P}^{i-1})) = \hat{Y}^{*(i-1)} \cdot \hat{C}^{i-1}, i = 1, \ldots$$

- Standardize approximate LV's

PLS-SEM algorithm - step 3: outer approximation

- Update approximate LV's as a weighted sum

$$\hat{Y}^{*i} = X \cdot \hat{W}^i, i = 1, \ldots$$

  where $\hat{W}^i = g(\hat{Y}^i)$ is a $K \times H$ matrix of temporary weights and $\hat{Y}^*$ is the approximate LV's
- When a LV is measured reflectively, it is the estimated correlation weights, $\hat{L}^i$, that are used as weights in the weighted sum
- When a LV is measured formatively, it is the estimated regression weights, $\hat{\Omega}^i$, that are used as weights in the weighted sum
- Standardize approximate LV's

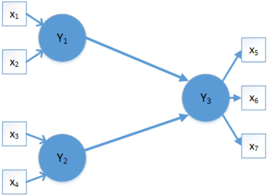
PLS-SEM algorithm - step 4 and 5

- Based on the initialization and step 2 and 3 we get the sequence,

$$\hat{Y}^{*0}, \hat{Y}^{1}, \hat{Y}^{*1}, \hat{Y}^{2}, \hat{Y}^{*2}, \ldots$$

- When $\hat{W}^i$ only change by a small amount from one iteration to the next, the iterations stops. We denote the final weights (after convergence) as $\hat{W}$
- The final LV scores are calculated as $\hat{Y} = X \cdot \hat{W}$
- Using the LV scores, $\hat{Y}$, and the indicators, X, the final parameter estimates, $\hat{P}$ (using the path weighting scheme), $\hat{L}$ and $\hat{\Omega}$ are calculated

Estimation after convergence – inner relations



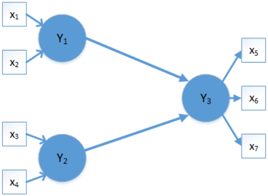
$$P = \begin{bmatrix} 0 & 0 & p_{13} \\ 0 & 0 & p_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

- We recall that the inner relations are given by $Y = Y \cdot (P + P') + \nu$
- The inner weights, P, are estimated by a range of separate OLS regressions using the estimated LV scores
- For example, the weights p_{13} and p_{23} are estimated from the linear regression

$$\hat{Y}_3 = p_{13}\hat{Y}_1 + p_{23}\hat{Y}_2 + \nu_3$$

- We use $\hat{}$ on the LV, to indicate that we are dealing with the LV approximations from the PLS-SEM algorithm

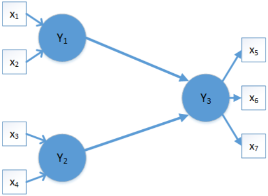
Estimation after convergence – outer relations I



$$L = \begin{bmatrix} l_{11} & l_{12} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & l_{23} & l_{24} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & l_{35} & l_{36} & l_{37} \end{bmatrix}$$

- We recall that the outer relations with respect to the correlation weights, L, are given as $X = Y \cdot L + \varepsilon$
- The individual correlation weights in L are estimated as the OLS regression where the indicator is the dependent variable and its respective LV is the independent variable

Estimation after convergence – outer relations II



$$L = \begin{bmatrix} l_{11} & l_{12} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & l_{23} & l_{24} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & l_{35} & l_{36} & l_{37} \end{bmatrix}$$

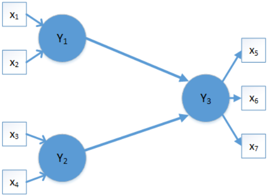
- For example, the correlation weights, l_{11} and l_{35} are estimated from two separate linear regressions:

$$x_1 = l_{11}\hat{Y}_1 + \varepsilon_1$$

$$x_5 = l_{35}\hat{Y}_3 + \varepsilon_5$$

- The regression estimates are the same as the correlation between an indicator and its LV approximation because the variables are standardized

Estimation after convergence – outer relations III



$$\Omega = \begin{bmatrix} \omega_{11} & 0 & 0 \\ \omega_{21} & 0 & 0 \\ 0 & \omega_{32} & 0 \\ 0 & \omega_{42} & 0 \\ 0 & 0 & \omega_{53} \\ 0 & 0 & \omega_{63} \\ 0 & 0 & \omega_{73} \end{bmatrix}$$

- The regression weights are calculated by regressing each LV approximation on its indicators. For example:

$$\hat{Y}_1 = \omega_{11} x_1 + \omega_{21} x_2 + \delta_1$$

$$\hat{Y}_3 = \omega_{53} x_5 + \omega_{63} x_6 + \omega_{73} x_7 + \delta_3$$

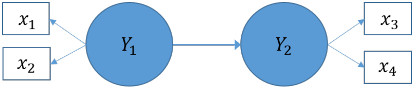
Interpretation

- With theoretical support of our model, the inner weights are often interpreted as causal effects
- A correlation weight can be interpreted as the effect of the LV on the indicator
- When a construct is measured reflectively, the correlation weights are used for interpretation
- A regression weight can be interpreted as the effect of the indicator on the LV
- When a construct is measured formatively, the regression weights are used for interpretation

Model estimation - statistical properties I

- Focus on maximizing the explained variance in the dependent variables. i.e. focus is on prediction
- No distributional assumptions $\rightarrow$ rely on bootstrapping for significance testing
- Assume that all indicator variance is useful and should be used to estimate LV scores $\rightarrow$ measurement error of indicators is transferred to latent variables $\rightarrow$ typically results in overestimation of the parameters in the measurement model and attenuated estimates of the parameters in the structural model (called PLS-SEM bias)

Model estimation - statistical properties II



- In empirical relevant settings PLS-SEM does not estimate correlations between latent variables, $cor[Y_1, Y_2]$, but correlations between conditional expectations of latent variables. For example, in the model given above PLS-SEM will estimate

$$cor[E(Y_1|x_1, x_2), E(Y_2|x_3, x_4)]$$

- Increasing both sample size and number of indicators sufficiently, we will estimate the correlation between the latent variables $\rightarrow$ consistency of PLS-SEM (called consistency at large)

PLS-SEM and principal component analysis (PCA)

- PCA
  - In PCA we extract a set of components from the set of observed variables we want to analyze
  - Each component is orthogonal to the other components
  - Each component is constructed as a weighted sum of all observed variables
- PLS-SEM
  - The set of observed variables we want to analyze is divided in H sets, where each set only measure a specific latent variable
  - The division of the variables is determined based on *prior theory* or *subject matter knowledge*
  - The components are not orthogonal by construction (the purpose of PLS-SEM is to investigate if the latent variables are related to each other)
  - A latent variable is constructed as a weighted sum of the observed variables intended to measure that latent variable (not all observed variables)

Corporate reputation model from Eberl (2010)

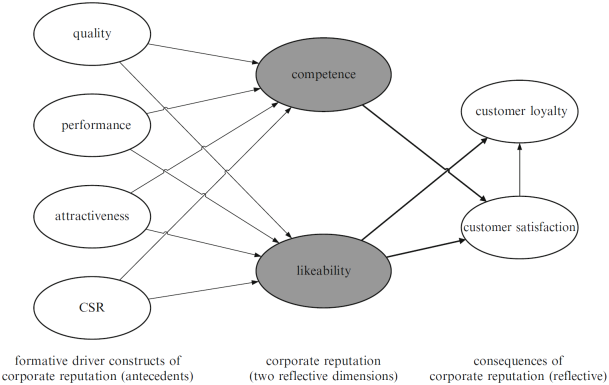


Figure: The extended corporate reputation model. (Source: Eberl 2010, in Handbook of Partial Least Squares: Concepts, Methods and Applications)

Corporate reputation I

“(..) *a stakeholder’s overall evaluation of a company over time. This evaluation is based on the stakeholder’s direct experiences with the company, any other form of communication and symbolism that provides information about the firm’s actions and/or a comparison with the actions of other leading rivals*”
(Source: Gotsi and Wilson 2001, Corporate reputation: seeking a definition)

Corporate reputation II

- The two core dimensions of corporate reputation: Competence (COMP) and likeability (LIKE)
- Cognitive evaluation: Competence. A more rationally based dimension
- Affective evaluation: Likeability. A more emotionally based dimension
- The model aims to explain the effect of corporate reputation on customer satisfaction (CUSA) and customer loyalty (CUSL)

Antecedent to corporate reputation

- Drivers of corporate reputation
  - Quality (QUAL)
  - Performance (PERF)
  - Attractiveness (ATTR)
  - Corporate social responsibility (CSOR)
- We will look further into these constructs in the lecture concerning the evaluation of the formative measurement model

Data sample and scales

- Telephone interviews: asking respondents about four major network providers in German mobile communication market
- 7-point Likert scales, with higher scores indicating higher levels of agreement with these statements

Competence (COMP)	
comp_1	[The company] is a top competitor in its market.
comp_2	As far as I know, [the company] is recognized worldwide.
comp_3	I believe that [the company] performs at a premium level.
Likeability (LIKE)	
like_1	[The company] is a company that I can better identify with than other companies.
like_2	[The company] is a company that I would regret more not having if it no longer existed than I would other companies.
like_3	I regard [the company] as a likeable company.
Customer loyalty (CUSL)	
cusl_1	I would recommend [company] to friends and relatives.
cusl_2	If I had to choose again, I would choose [company] as my mobile phone service provider.
cusl_3	I will remain a customer of [company] in the future.

Figure: Table 3.1

Subset of corporate reputation model

- In this lecture we will focus on the subset of the model where we have reflective indicators

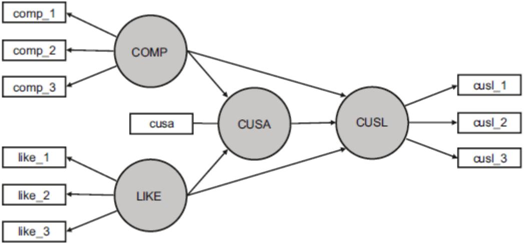


Figure: Figure 3.2

Corporate reputation model – specification and estimation

- Data for the corporate reputation model are readily available – either directly from the `seminr` package or via the books homepage
- Data are basically already cleaned and it is just a matter of following the instructions in Hair et al. 2021
- To specify the measurement model as well as the structural model follow the instructions in Hair et al. 2021
- After having estimated the model you can assess its convergence properties – after four iterations we obtained convergence
- Many of the results from the estimation appear as point estimates – uncertainty can be assessed via bootstrapping

Exercises

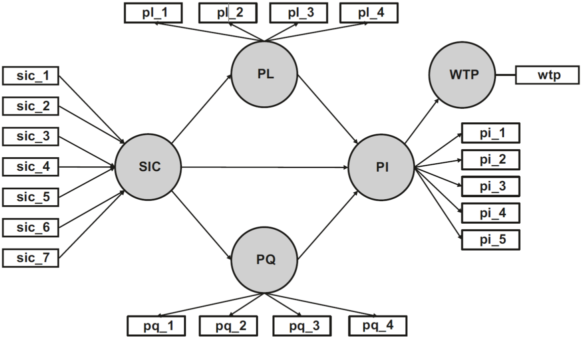


Figure: Fig 3.10

- The influencer model analyzes if consumers are likely to follow social media influencers' purchase recommendations
- See also chapter 3 in Hair et al. 2021 for a description
- Complete exercises 1,2, and 3 on page 71 in Hair et al. 2021