Partial Least Squares Structural Equation Modeling IIIb

Evaluation of structural models

Morten Berg Jensen

Department of Economics and Business

February 29, 2024

Outline

Outcome

This lecture will help you to understand

- The necessary steps for an assessment of the structural model
  - Assessment of collinearity issues and model relationships
  - Assessment of explanatory and predictive power
  - (Model comparisons)

Evaluation of the structural model

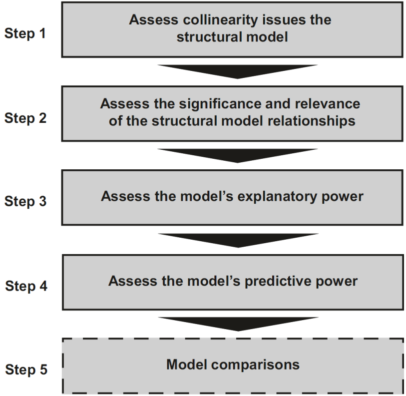

Figure: Fig 6.1

Collinearity assessment

- Examine each part of the model separately. E.g. asses Y_1, Y_2 and Y_3 together, and then Y_2, Y_3 and Y_4
- Tolerance < 0.2 (VIF > 5) indicates potential collinearity problems
- If multicollinearity is an issue $\rightarrow$ consider eliminating constructs or merging predictors into a single construct

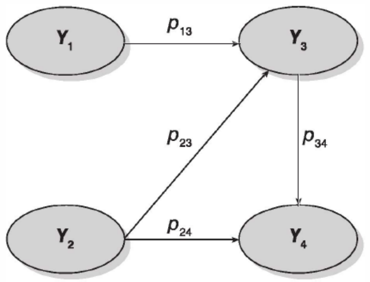

Figure: Collinearity in the structural model. (Source: Hair et al. 2022)

Structural model path coefficients

- Path coefficients have values approximately between -1 and +1
- Bootstrapping is used to obtain standard errors, *t*-values, *p*-values and confidence intervals
- Identify significant relations
- Asses if a small but significant effect is relevant
- Look at direct, indirect and total effects

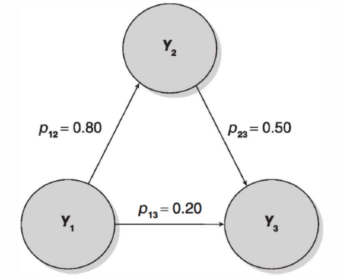

Figure: Direct, indirect, and total effects. (Source: Hair et al. 2022)

Explanatory power – R^2 value

- R^2: The amount of variance in the endogenous variable, that can be explained by the exogenous variables linked to it. It is calculated as the squared correlation between the predicted value and actual value of the construct
- Represent in-sample predictive power
- Whether R^2 is considered “high” depends on the field of research
- If we want to compare models with different number of exogenous variables or data sets with different sample sizes $\rightarrow$ use adjusted R^2, R^2_{adj}

$$R^2_{adj} = 1 - (1 - R^2)\frac{n-1}{n-k-1}$$

where k is the number of exogenous latent variables and n is the sample size

Explanatory power – Effect Size f^2

- Effect size f^2: Used to evaluate if an exogenous construct has a substantive impact on an endogenous construct

$$f^2 = \frac{R^2_{included} - R^2_{excluded}}{1 - R^2_{included}}$$

  where $R^2_{included}$ and $R^2_{excluded}$ are the R^2 values of the endogenous latent variable when a selected exogenous variable is included or excluded from the model
- Guidelines for assessment: $f^2 < 0.02$: no effect, $0.02 \leq f^2 < 0.15$: small effect, $0.15 \leq f^2 < 0.35$: medium effect, $0.35 \leq f^2$: large effect
- Compare the rank order of the effect sizes to the rank order of the path coefficients

Predictive power – cross-validation I

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Holdout 1	Training	Training	Training	Training
Training	Holdout 2	Training	Training	Training
Training	Training	Holdout 3	Training	Training
Training	Training	Training	Holdout 4	Training
Training	Training	Training	Training	Holdout 5

Figure: Fig 6.2

- For each fold: Estimate the model on the training data and predict on the holdout data
- Number of folds: Training sample size should meet minimum sample size requirement
- Default is to use 10 folds and use 10 repetitions of cross-validation (to stabilize results)

Predictive power – cross-validation II

- The prediction error for indicator k for observation i:

$$\hat{e}_{i,k} = x_{i,k} - \hat{x}_{i,k}$$

  where $\hat{x}_{i,k}$ is the out-of-sample prediction of $x_{i,k}$
- Root mean square error for indicator k:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{e}_{i,k})^2}$$

  this is the generally recommended statistic
- For highly skewed distributions, the mean absolute error might be preferable:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{e}_{i,k}|$$

Predictive power - linear regression benchmark I

- Generally, the smaller the value of *RMSE* or *MAE* the higher the predictive power
- The linear regression model benchmark is suggested as a benchmark for evaluating the absolute level of *RMSE* or *MAE*
- This involves regressing each of the indicators of the dependent construct on the indicators of the exogenous latent variables
- The resulting models are used for predicting the pertinent indicator

Predictive power - linear regression benchmark II

1. High predictive power: All indicators have lower RMSE in PLS-SEM compared to linear regression benchmark
2. Medium predictive power: The majority of indicators have lower RMSE in PLS-SEM compared to linear regression benchmark
3. Low predictive power: A minority of indicators have lower RMSE in PLS-SEM compared to linear regression benchmark
4. Lack of predictive power: No indicators have lower RMSE in PLS-SEM compared to linear regression benchmark

- Use the direct antecedent (DA) approach for models with mediators
- Focus on key construct(s)

Alternative model configurations

- In many instances when a theory is considered in a new context alternative models emerge
- Generally a model with a high R^2 is preferred but at the same time parsimonious models are seen as more likely to generalize to other settings
- The Bayesian information criterion (BIC) seeks to balance those two goals
- The model with the smallest BIC value is preferred
- The relative benefits of the different models can be determined based on Akaike weights (building on BIC values)
- These weights indicate a model's relative likelihood given the data and a set of competing models

Full corporate reputation model

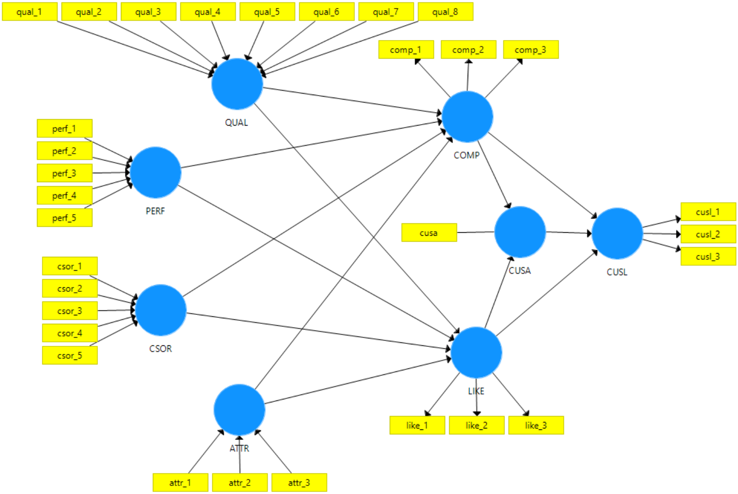

Figure: Full corporate reputation model. (Source: Hair et al. 2022)

Corporate reputation model – structural model evaluation I

- Collinearity
  - No issues, all VIF's are below 5 (and vast majority below 3)
- Path coefficients, direct effects

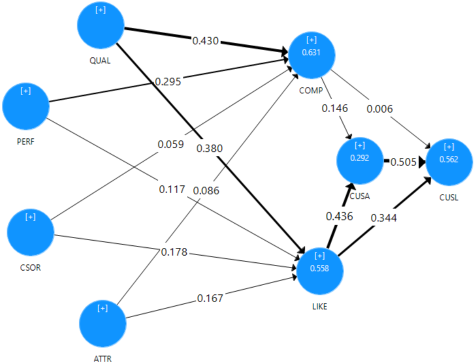

Figure: Path coefficients

Corporate reputation model – structural model evaluation II

- Path coefficients relevance, direct effects
  - QUAL is the strongest driver of COMP and LIKE
  - CSOR and ATTR are more important for LIKE than for COMP
  - LIKE is the most important corporate reputation dimension in terms of creating customer satisfaction and customer loyalty
- Path coefficients relevance, total effects
  - QUAL is the most important of the four formative driver constructs in terms of creating customer satisfaction and customer loyalty

Corporate reputation model – structural model evaluation III

- Path coefficients significance
  - All direct effects are significant on a 5% level except the following four relations: CSOR $\rightarrow$ COMP, ATTR $\rightarrow$ COMP, PERF $\rightarrow$ LIKE and COMP $\rightarrow$ CUSL
- The results indicate that companies should focus on creating likeability instead of competence to increase the customer loyalty

Corporate reputation model – structural model evaluation IV

- Explanatory power – R^2 and R^2_{adj}
  - The model can explain more than 50% of the variance in all endogenous constructs, except CUSA
  - Only minor difference between R^2 and R^2_{adj}
- Explanatory power – f^2
  - Using the guidelines for assessment we have that
    - CUSA has a large effect on CUSL
    - Most other constructs have small effects on their respective endogenous constructs

Corporate reputation model – structural model evaluation V

- Predictive power

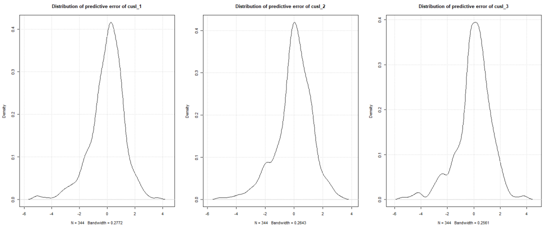

Figure: Prediction errors

  - Only slightly negative skewed prediction errors
  - PLS-SEM predictions is better than linear benchmark predictions on all indicators
- Comparing model 1 to model 2 and 3 gives a very strong weighting to model 1

Exercises

- Answer question 1, 2, 3(a), and 3(b) on page 136 in Hair et al. 2021