Structural Equation Modeling (SEM) with SmartPLS

Case Study

A Company wants to measure the effect of customer satisfaction on customer loyalty through SEM. To do that, the survey was collected and model was established based on theory with following latent variables and indicators. Each statement (indicator) was measure on a 7-point scale (1 =fully disagree to 7 = fully agree) and received 344 valid responses from the respondents.

<table>
<thead>
<tr>
<th>Competence (COMP)</th>
<th>Indicator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>comp_1</td>
<td></td>
<td>[The company] is a top competitor in its market.</td>
</tr>
<tr>
<td>comp_2</td>
<td></td>
<td>As far as I know, [the company] is recognized worldwide.</td>
</tr>
<tr>
<td>comp_3</td>
<td></td>
<td>I believe that [the company] performs at a premium level.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Likeability (LIKE)</th>
<th>Indicator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>like_1</td>
<td></td>
<td>[The company] is a company that I can better identify with than other companies.</td>
</tr>
<tr>
<td>like_2</td>
<td></td>
<td>[The company] is a company that I would regret more not having if it no longer existed than I would other companies.</td>
</tr>
<tr>
<td>like_3</td>
<td></td>
<td>I regard [the company] as a likeable company.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Customer Satisfaction (CUSA)</th>
<th>Indicator</th>
<th>Question</th>
</tr>
</thead>
</table>
| cusa                        |           | How satisfied are you with [company]?

<table>
<thead>
<tr>
<th>Customer Loyalty (CUSL)</th>
<th>Indicator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cusl_1</td>
<td></td>
<td>I would recommend [company] to friends and relatives.</td>
</tr>
<tr>
<td>cusl_2</td>
<td></td>
<td>I would choose [company] as my mobile phone services provider.</td>
</tr>
<tr>
<td>cusl_3</td>
<td></td>
<td>I will remain a customer of [company] in the future.</td>
</tr>
</tbody>
</table>

Steps to Perform Structural Equation Modeling (SEM)

1. **Specify the measurement model**: As latent variables are not directly observed, they are formed from one or more indicators/statements. There are two types of measurement models: Formative and Reflective. In the present case customer loyalty, the CUSL is reflective as the arrow direction is toward the indicators/questions.

2. **Specify the structural model**: Based on theory, we should choose latent variables and specify the model.
Applying the traditional multivariate techniques, we can estimate the two endogenous variables (CUS and CUSL) in single analysis. Formulate the hypothesis:

\[ H_1: \text{Customer satisfaction has a positive effect on customer loyalty} \]

Check the path coefficient and significance value to confirm the hypothesis

Make four latent variables from 10 indicators/statements using Factor Analysis: COMP, LIKE, CUSA and CUSL. Calculate coefficients and variance explained like Multiple Regression: dependent variable (CUSL) and independent variables (COMP, LIKE and CUSA). Perform another regression for dependent variable (CUSA) and independent variables (COMP and LIKE).

In the process we are using indicators and little cumbersome to explain in different regression analysis. In SEM we use the latent variables to explain the path coefficients and run the model simultaneously.

Create a project in SmartPLS

- Click on the SmartPLS icon.
- Click on “New Project” on the left-hand upper side of the screen.
- The screen will ask for the name of the new project. I have named it “Corporate Reputation Project”. Click OK.
- The project would appear on the left hand pane.
- Double click on the first option under the new project name “Corporate Reputation project”.
- Import the sample data file Corporate Reputation Data.csv. Data files need to be with .CSV extension. Kindly note while importing data, file to be renamed with no special character Corporate_Reputation_Data.
- Click at **Corporate Reputation Project**. Indicators available and an space for drawing model can be seen.
- Draw the latent constructs and its indicators. Drag and drop from the list of indicators and rename the latent variable as desired.

![Diagram 1](image1.png)

- Similarly draw the other latent variables (**Comp, Like, Cusa, Cusl**). When you draw more than one constructs, the color of the constructs changes to Red until all the constructs are connected.

![Diagram 2](image2.png)

- Use **Connect** tab in the top of the pane and connect the independent (exogenous) variables with dependent (endogenous) variables (Please see your structural model for reference).

![Diagram 3](image3.png)
Now, to run this model, go to “calculate” tab on the top right of the pane and click on “consistent PLS algorithm”. Kindly note that Consistent PLS algorithm performs a correction of reflective constructs’ correlations to make results consistent with a factor-model. Consistent PLS is used when all constructs are reflective. In case of mix of reflective and formative regular PLS is recommended.

3. PLS Path Model Estimation: While running the PLS path model, one should pay attention to path weighting method (path weighting is the recommended as it provides the highest R² value for endogenous latent variables). Following will be the result after the calculation. Let’s assess the results one by one in the next steps.

<table>
<thead>
<tr>
<th></th>
<th>R Square</th>
<th>R Square Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUSL</td>
<td>0.504</td>
<td>0.500</td>
</tr>
<tr>
<td>CUSA</td>
<td>0.024</td>
<td>0.019</td>
</tr>
</tbody>
</table>
4. **Assess the results of Measurement Models:** The very first thing we need to look at the outer loadings. For example, COMP has three indicators which have loadings of 0.667, 0.785 and 0.751 (>0.6). There is another section of Construct Reliability & Validity which assess the quality of each latent variable.

![Internal Consistency Reliability and Convergent Validity](image)

**Internal Consistency Reliability**
- Composite Reliability (CR) > 0.708 – 0.60 - 0.70 is acceptable.
- Cronbach's alpha (α> 0.7 or 0.6)

**Convergent Validity**
- Average Variance Extracted (AVE) > 0.5
- Discriminant Validity
- Fornell-Larcker criterion
- Cross Loadings

---

**Convergent validity** is the extent to which a measure correlates positively with other measures (indicators) of the same construct. To establish **convergent validity**, researchers consider the **outer loadings of the indicators**, as well as the **average variance extracted (AVE)**.

**Indicator reliability** denotes the proportion of indicator variance that is explained by the latent variable. However, reflective indicators should be eliminated from measurement models if their loadings within the PLS model are smaller than 0.4 (Hulland 1999, p. 198).

**In our case, we decide to remove cusl_1 (outer loadings <0.4) and hence new PLS model was generated.**

---

Instructor: Dr. Prabhat Mittal M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
URL: [http://people.du.ac.in/~pmittal/]
Discriminant validity is the extent to which a construct is truly distinct from other constructs by empirical standards.

- **Cross-Loadings**: An indicator’s outer loadings on a construct should be higher than all its cross loadings with other constructs.

- **Fornell-Larcker criterion**: The square root of the AVE of each construct should be higher than its highest correlation with any other construct (Fornell and Larcker, 1981).

  The AVE values are obtained by squaring each outer loading, obtaining the sum of the three squared outer loadings, and then calculating the average value. For example, with respect to construct COMP, 0.667, 0.785, and 0.751 squared are 0.445, 0.616, and 0.564. The average value (AVE) is 0.542. Square-root of AVE=0.736 (diagonal value)

  Henseler, Ringle and Sarstedt (2015) show by means of a simulation study that these approaches do not reliably detect the lack of Discriminant validity in common research situations. These authors therefore propose an alternative approach to assess Discriminant validity: the Heterotrait-monotrait ratio of correlations (HTMT). If the HTMT value is below 0.90, Discriminant validity has been established between two reflective constructs.

5. **Assessing Results of the Structural Model**: Once we know that the indicators in the latent variables are reliable, we should assess the results of structural model. Run Bootstrapping procedure to check the statistical significance test of path coefficients, checking T-statistics which should be greater than 1.96 (5% significance level). Cautious note: Sometimes Consistent PLS results with n/a:


   - R-square: amount of variance in the endogenous constructs explained by all of the exogenous constructs linked to it
   - Effect size f-square: The change in the $R^2$ value when a specified exogenous construct is omitted from the model can be used to evaluate whether the omitted construct has a
     - $-0.02$ → small
     - $-0.15$ → medium
     - $-0.35$ → large effects (Cohen, 1988)
   - Blindfolding $Q^2 > 0$ for a certain reflective endogenous latent variable indicate the path model’s predictive relevance for this particular construct. This procedure does not apply for formative endogenous constructs.