Panel:
Using Structural Equation Modeling (SEM)
Using Partial Least Squares (SmartPLS)

Presenters:
Dr. Faizan Ali, Assistant Professor
Dr. Cihan Cobanoglu, McKibbon Endowed Chair Professor
University of South Florida Sarasota-Manatee

Housekeeping

Panel:
Using Structural Equation Modeling (SEM)
Using Partial Least Squares (SmartPLS)

Presenters:
Dr. Faizan Ali, Assistant Professor
Dr. Cihan Cobanoglu, McKibbon Endowed Chair Professor
University of South Florida Sarasota-Manatee
Raise Hand

Webinar will be recorded and sent to all attendees
Sponsors

http://www.iibaconference.org
http://conference.anahei.org

ANAHEI Journals

International Interdisciplinary Business-Economics Advancement Journal
(IIBA Journal)
ISSN 2375-9615

http://www.anahei.org
Webinar Outline

• Introduction to SEM
• Usage of SEM
• Various methodological issues
• Formative vs. reflective constructs
• Modelling using PLS
• Evaluation of measurement model
• Evaluation of structural model
**Statistical Methods**

- With first-generation statistical methods, the general assumption is that the data are *error free*.
  - Multiple regression
  - Logistic regression
  - Analysis of variance
  - Cluster analysis
  - Exploratory factor analysis
  - Multidimensional scaling

- With second-generation statistical methods, the measurement model stage attempts to identify the *error component* of the data.
  - SEM
    - CB - SEM
    - PLS - SEM

**Measurement error**

- Measurement error is the difference between *true value of variable* and *value obtained by using scale*.

- Types of measurement error
  - *Random error can affect the reliability of construct*
  - *Systematic error can affect the validity of construct* (Hair et al. 2014)

- Source of error
  - Poorly worded questions in survey
  - Incorrect application of statistical methods
  - Misunderstanding of scaling approach
Structural equation modeling (SEM)

• An advanced statistical tool used to assess complex models with many relationships, perform confirmatory factor analysis, and incorporate both unobserved and observed variables.

• It combines characteristics of factor analysis and multiple regressions to simultaneously examine both direct and indirect effects of independent and dependent variables.

• Largely used in various academic disciplines over the last decade

• There are two approaches to estimate the relationships in a structural equation model (SEM):
  • Covariance-based SEM (CB-SEM)
  • Variance-Based - VB-SEM/ PLS-SEM.

Research Article


Patrícia Oom do Valle¹ and Guy Assaker²
The use of partial least squares path modeling in international marketing

Author(s): Jörg Henseler, Christian M. Ringle, Rudolf R. Sinkovics
Citation: Jörg Henseler, Christian M. Ringle, Rudolf R. Sinkovics (2009). The use of partial least squares path modeling in international marketing, in Rudolf R. Sinkovics, Pervez N. Ghaure (ed.) New Challenges to International Marketing (Advances in International Marketing, Volume 20) Emerald Group Publishing Limited, pp.277 - 319

On the Use, Usefulness, and Ease of Use of Structural Equation Modeling in MIS Research: A Note of Caution

Wynne W. Chin and Peter A. Todd
MIS Quarterly
Vol. 19, No. 2 (Jun., 1995), pp. 237-246

Published by: Management Information Systems Research Center, University of Minnesota
DOI: 10.2307/249690
Stable URL: http://www.jstor.org/stable/249690
Page Count: 10
SEM Software / Applications

<table>
<thead>
<tr>
<th>CB-SEM</th>
<th>PLS-SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMOS</td>
<td>SmartPLS</td>
</tr>
<tr>
<td>LISREL</td>
<td>PLS-Graph</td>
</tr>
<tr>
<td>MPLUS</td>
<td>PLS-GUI</td>
</tr>
<tr>
<td>EQS</td>
<td>SPADPLS</td>
</tr>
<tr>
<td>SAS</td>
<td>LVPLS</td>
</tr>
<tr>
<td>R</td>
<td>WarpPLS</td>
</tr>
<tr>
<td>SEPATH</td>
<td>PLS-PM</td>
</tr>
<tr>
<td>CALIS</td>
<td>semPLS</td>
</tr>
<tr>
<td>LISCOMP</td>
<td>Visual PLS</td>
</tr>
<tr>
<td>Lavaan</td>
<td>PLSPath</td>
</tr>
<tr>
<td>COSAN</td>
<td>XLSTAT</td>
</tr>
</tbody>
</table>
Usage of SEM in Hospitality Research

- 22.3% articles published in four top hospitality journals from 2008 to 2010 used SEM (Line & Runyan, 2014).
- 10.2% articles published in four top hospitality journals from 2000 to 2009 used SEM (Yoo et al., 2011).
- 7.51% articles published in the Cornell Hospitality Quarterly from 2008 to 2011 used SEM (Law et al., 2012).

Of the articles, 379 utilized CB-SEM and 45 PLS-SEM.

A critical look at the use of SEM in international business research

Richter, Nicole; Sinkovics, Rudolf R; Ringle, Christian; Schlägel, Christopher

Structural Equation Modeling Algorithm and Its Application in Business Analytics

Shahryar Sorooshian (Universiti Malaysia Pahang, Malaysia)

Source Title: Organizational Productivity and Performance Measurements Using Predictive Modeling and Analytics
Copyright © 2017 Pages: 23
DOI: 10.4018/978-1-5225-0654-6.ch002

Figure 1. Scopus documents

Source: Scopus (2016)

Poll
Methodological Issues Ignored

- Sample Size
- Model Complexity
- Prediction-Based Modelling
- Data Normality
- Formative and Single Item Constructs
- Weak Theoretical Support.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>PLS</th>
<th>CBSEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective</td>
<td>Prediction-oriented</td>
<td>Parameter-oriented</td>
</tr>
<tr>
<td>Approach</td>
<td>Variance-based</td>
<td>Covariance-based</td>
</tr>
<tr>
<td>Assumption</td>
<td>Predictor specification (nonparametric)</td>
<td>Typically multivariate normal distribution and independent observations (parametric)</td>
</tr>
<tr>
<td>Parameter estimates</td>
<td>Consistent as indicators and sample size increase (i.e., consistency of large)</td>
<td>Consistent</td>
</tr>
<tr>
<td>Latent variable scores</td>
<td>Explicitly estimated</td>
<td>Indeterminate</td>
</tr>
<tr>
<td>Epistemic relationship between an LV and its measures</td>
<td>Can be modelled in either formative or reflective mode</td>
<td>Typically only with reflective indicators. However, the formative mode is also supported.</td>
</tr>
<tr>
<td>Implications</td>
<td>Optimal for prediction accuracy</td>
<td>Optimal for parameter accuracy</td>
</tr>
<tr>
<td>Model complexity</td>
<td>Large complexity (e.g., 100 constructs and 1,000 indicators)</td>
<td>Small to moderate complexity (e.g., less than 100 indicators)</td>
</tr>
<tr>
<td>Sample size</td>
<td>Power analysis based on the portion of the model with the largest number of predictors. Minimal recommendations range from 30 to 100 cases.</td>
<td>Ideally based on power analysis of specific model - minimal recommendations range from 200 to 800.</td>
</tr>
<tr>
<td>Type of optimization</td>
<td>Locally iterative</td>
<td>Globally iterative</td>
</tr>
<tr>
<td>Significance tests</td>
<td>Only by means of simulations; restricted validity</td>
<td>Available</td>
</tr>
<tr>
<td>Availability of global Goodness of Fit (GoF) metrics</td>
<td>Are currently being developed and discussed</td>
<td>Established GoF metrics available</td>
</tr>
</tbody>
</table>
PLS-SEM

- Partial Least Squares (PLS) focuses on the prediction of a specific set of hypothesized relationships that maximizes the explained variance in the dependent variables (Hair, Ringle, & Sarstedt, 2011).

SmartPLS
(Ringle et al., 2005)

“SmartPLS 3 is a milestone in latent variable modeling. It combines state of the art methods (e.g., PLS-POS, IPMA, complex bootstrapping routines) with an easy to use and intuitive graphical user interface.”

Joe F. Hair, DBA Founder, Senior Scholar, and Professor, Kennesaw State University, USA.
Presence of PLS-SEM

"PLS-SEM: Indeed a Silver Bullet" by Joseph F. Hair Jr., Christian M ...
digita commons.kennesaw.edu > FacPubs > 2367
by JF Hair Jr. - 2011 - Cited by 2403 - Related articles
The current paper reviews PLS-SEM and its algorithm, and provides an ... that PLS-SEM path modeling, if appropriately applied, is indeed a “silver bullet” for ...

The use of partial least squares path modeling in international marketing
J.Henseler, C.M. Ringle - Advances in international marketing ..., 2009 - emeraldinsight.com
The advent of structural equation modeling (SEM) with latent variables has changed the nature of research in international marketing and management. Researchers acknowledge the possibilities of distinguishing between measurement and structural models and ...
Cited by 2872 Related articles • All 14 versions • Cite • Save • More

An assessment of the use of partial least squares structural equation modeling in marketing research
JF Hair, M Sarstedt, CM Ringle, JA Mena - Journal of the academy of ..., 2012 - Springer
Abstract Most methodological fields undertake regular critical reflections to ensure rigorous research and publication practices, and, consequently, acceptability in their domain. Interestingly, relatively little attention has been paid to assessing the use of partial least ... Cited by 1296 Related articles • All 15 versions • Web of Science 382 • Cite • Save • More

Citations A primer on partial least squares structural equation modeling (PLS-SEM)
JF Hair Jr, GTM Hult, C Ringle, M Sarstedt - 2016 - Sage Publications
Cited by 1955 Related articles • All 6 versions • Cite • Save • More

http://www.journals.elsevier.com/long-range-planning/most-cited-articles
Justification for usage of PLS-SEM

• Primary objective of study is prediction and explanation of target constructs.
• Smaller sample sizes
• Complex models
• No assumptions about the underlying data (Normality assumptions)
• Support reflective and formative measurement models as well as single item construct.
• Weaker theoretical support/Integration of multiple theories.
• Works with ordinal and binary scaled questions.
**Missing Data Treatment and Data Normality**

- In general, missing data occurs when respondents intentionally or unintentionally fail to answer at least one question in a survey.

  - If > 25% missing, throw out the questionnaire.
  - Less than 5% values missing per indicator, use:
    - Use the midpoint of the scale
    - Mean of the respondent
    - Expectation maximization (EM)

- **Construct missing**: If respondent did not answer all items related to one construct, delete it.

- **Suspicious Respond Patterns**: Remove responses

- **Data Normality** – Report Skewness and Kurtosis

---

**PLS-SEM**

- A PLS path model consists of two elements:
  - The structural model displays the relationships (paths) **between the constructs**.
  - The measurement models display the relationships **between the constructs** and the indicator variables (rectangles).
The decision of whether to measure a construct reflectively or formatively is not clear-cut (Hair et al., 2014).
REFLECTIVE MEASUREMENT MODEL

The goal of reflective measurement model assessment is to ensure the reliability and validity of the construct measures and therefore provide support for the suitability of their inclusion in the path model.

- **Reliability** is the extent to which an assessment tool produces stable and consistent results.

- **Validity** refers to the extent to which the construct measures what it is supposed to measure.

REFLECTIVE MEASUREMENT MODEL EVALUATION

- Internal Consistency Reliability
  - Composite Reliability (CR > 0.70 - in exploratory research 0.60 to 0.70 is acceptable).
  - Cronbach’s alpha (α > 0.7 or 0.6)

- Indicator reliability (> 0.708)
  - Squared Loading - the proportion of indicator variance that is explained by the latent variable

- Convergent validity
  - Average Variance Extracted (AVE > 0.5)

- Discriminant validity
  - Fornell-Larcker criterion
  - Cross Loadings
  - HTMT Criteria (New Tool).
Convergent validity

• An established rule of thumb is that a latent variable should explain a substantial part of each indicator's variance, usually at least 50%.

• This means that an indicator's outer loading should be above 0.708 since that number squared (0.7082) equals 0.50.

Discriminant validity

• 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.
  • For example, with respect to construct \( Y_1 \), 0.60, 0.70, and 0.90 squared are 0.36, 0.49, and 0.81. The sum of these three numbers is 1.66 and the average value is therefore 0.55 (i.e., 1.66/3).
### Discriminate Validity

<table>
<thead>
<tr>
<th></th>
<th>COMP</th>
<th>CUSA</th>
<th>CUSL</th>
<th>LIKE</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMP</td>
<td>(\sqrt{AVE})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUSA</td>
<td>0.4356</td>
<td>Single-item Construct</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUSL</td>
<td>0.4496</td>
<td>0.6892</td>
<td>(\sqrt{AVE})</td>
<td></td>
</tr>
<tr>
<td>LIKE</td>
<td>0.6452</td>
<td>0.5284</td>
<td>0.6146</td>
<td>Formative</td>
</tr>
</tbody>
</table>

### Discriminant validity

**Loadings and Cross-loadings**

<table>
<thead>
<tr>
<th></th>
<th>ATT</th>
<th>COMP</th>
<th>CSOR</th>
<th>CUSA</th>
<th>CUSL</th>
<th>LIKE</th>
</tr>
</thead>
<tbody>
<tr>
<td>attr_1</td>
<td>0.758</td>
<td>0.496</td>
<td>0.405</td>
<td>0.064</td>
<td>0.062</td>
<td>0.430</td>
</tr>
<tr>
<td>attr_2</td>
<td>0.504</td>
<td>0.269</td>
<td>0.331</td>
<td>0.032</td>
<td>-0.003</td>
<td>0.346</td>
</tr>
<tr>
<td>attr_3</td>
<td>0.889</td>
<td>0.533</td>
<td>0.538</td>
<td>0.054</td>
<td>0.073</td>
<td>0.554</td>
</tr>
<tr>
<td>comp_1</td>
<td>0.546</td>
<td>0.801</td>
<td>0.601</td>
<td>0.082</td>
<td>0.103</td>
<td>0.607</td>
</tr>
<tr>
<td>comp_2</td>
<td>0.451</td>
<td>0.834</td>
<td>0.423</td>
<td>0.113</td>
<td>0.105</td>
<td>0.461</td>
</tr>
<tr>
<td>comp_3</td>
<td>0.519</td>
<td>0.858</td>
<td>0.472</td>
<td>0.100</td>
<td>0.108</td>
<td>0.498</td>
</tr>
<tr>
<td>csor_1</td>
<td>0.486</td>
<td>0.460</td>
<td>0.773</td>
<td>0.070</td>
<td>0.061</td>
<td>0.471</td>
</tr>
<tr>
<td>csor_2</td>
<td>0.333</td>
<td>0.292</td>
<td>0.572</td>
<td>0.105</td>
<td>0.145</td>
<td>0.394</td>
</tr>
<tr>
<td>csor_3</td>
<td>0.481</td>
<td>0.492</td>
<td>0.838</td>
<td>0.069</td>
<td>0.039</td>
<td>0.518</td>
</tr>
<tr>
<td>csor_4</td>
<td>0.422</td>
<td>0.313</td>
<td>0.018</td>
<td>0.008</td>
<td>0.038</td>
<td>0.429</td>
</tr>
<tr>
<td>csor_5</td>
<td>0.484</td>
<td>0.507</td>
<td>0.847</td>
<td>0.145</td>
<td>0.096</td>
<td>0.513</td>
</tr>
<tr>
<td>cusa</td>
<td>0.075</td>
<td>0.118</td>
<td>0.119</td>
<td>1.000</td>
<td>0.614</td>
<td>0.043</td>
</tr>
<tr>
<td>cusl_2</td>
<td>0.057</td>
<td>0.081</td>
<td>0.065</td>
<td>0.504</td>
<td>0.874</td>
<td>0.074</td>
</tr>
<tr>
<td>cusl_3</td>
<td>0.073</td>
<td>0.141</td>
<td>0.082</td>
<td>0.586</td>
<td>0.099</td>
<td>0.070</td>
</tr>
<tr>
<td>like_1</td>
<td>0.577</td>
<td>0.585</td>
<td>0.574</td>
<td>0.056</td>
<td>0.082</td>
<td>0.880</td>
</tr>
<tr>
<td>like_2</td>
<td>0.515</td>
<td>0.509</td>
<td>0.510</td>
<td>0.038</td>
<td>0.060</td>
<td>0.867</td>
</tr>
<tr>
<td>like_3</td>
<td>0.489</td>
<td>0.540</td>
<td>0.512</td>
<td>0.015</td>
<td>0.066</td>
<td>0.845</td>
</tr>
</tbody>
</table>

Bold values are loadings for item which are above for the recommended value of 0.5
FORMATIVE MEASUREMENT MODEL EVALUATION

• The statistical evaluation criteria for reflective measurement scales cannot be directly transferred to formative measurement models where indicators are likely to represent the construct's independent causes and thus do not necessarily correlate highly.

Formative Measurement Model
➢ Assess Collinearity Among Indicators (VIF < 5)
➢ Assess the Significance and relevance of outer weights (T-Value > 1.645).

The estimated values of outer weights in formative measurement models are frequently smaller than the of reflective indicators.
Interpretation:

✓ When an indicator's weight is significant, there is empirical support to retain the indicator.

✓ When an indicator's weight is not significant but the corresponding item loading is relatively high (> 0.50), the indicator should generally be retained.

✓ If both the outer weight and outer loading are nonsignificant, there is no empirical support to retain the indicator and it should be removed from the model.

PLS-SEM Structural Model Evaluation

PLS-SEM relies on a nonparametric bootstrap procedure to test coefficients for their significance.

• In bootstrapping, a large number of subsamples (i.e., bootstrap samples) are drawn from the original sample with replacement (random from the sampling population).

➢ If the measurement characteristics of constructs are acceptable, continue with the assessment of the structural model results. Path estimates should be statistically significant and meaningful.

➢ Moreover, endogenous constructs in the structural model should have high levels of explained variance — $R^2$ (coefficients of determination).
Coefficient of Determination (R²)

- The coefficient represents the exogenous latent variables' combined effects on the endogenous latent variable.

- It also represents the amount of variance in the endogenous constructs explained by all of the exogenous constructs linked to it.

- The R² value ranges from 0 to 1.

- In scholarly research as a rough rule of thumb:
  - 0.75 is substantial
  - 0.50 is moderate
  - 0.25 is weak
Additional Statistics to Report

- Effect Size ($f^2$)
- Predictive Relevance ($Q^2$)
- Goodness of Fit (GoF) Statistic

Mediator

- A mediating effect is created when a third variable or construct intervenes between two other related constructs.
- The role of the mediator variable then is to clarify or explain the relationship between the two original constructs.
- Indirect effects are those relationships that involve a sequence of relationships with at least one intervening construct involved.
Mediator

• Baron & Kenny (1986) has formulated the steps and conditions to ascertain whether full or partial mediating effects are present in a model.

\[ X \rightarrow M \rightarrow Y \]

\[ P_{12} \]

\[ P_{13} \]

\[ P_{23} \]

Reputation\hspace{1cm}Satisfaction\hspace{1cm}Loyalty

Mediation

• When testing mediative effects, researchers should rather follow Preacher and Hayes (2004,2008) and bootstrap the sampling distribution of the indirect effect, which works for simple and multiple mediator models.

• Run Bootstrapping, and then Indirect Effects + Confidence Interval Bias Corrected.
“My advisor tells me I should use the Baron and Kenny strategy for assessing mediation. But my reading of the literature tells me this isn’t recommended these days. What should I do?”

You have counted on your advisor for guidance and support. Now return the favour. All but the most stubborn of advisors are open to new ideas, and many are too busy or just don’t care enough to stay informed on recent developments. Give him or her a copy of the relevant literature or a copy of my book and make your case. Try my Beyond Baron and Kenny paper for a start (Communication Monographs, 2009, Vol 76, p. 408-420).

"In my mediation analysis examining the direct and indirect effects of X on Y through M, the path from X to M is not statistically significant. Does this mean there is no way that M could mediate the relationship between X and Y. According to Baron and Kenny (1986), it cannot. Should I bother estimating the indirect effect in this case?"

These days, we don't rely on statistical significance criteria for the individual paths in a mediation model in order to assess whether M functions as a mediator. The pattern of significance or non significance for individual paths in a mediation model is not pertinent to whether the indirect effect is significant. You absolutely should estimate the indirect effect. See Hayes (2009) for a brief discussion [PDF], or Chapter 6 of Hayes (2013).
"How can I tell whether I can claim full or partial mediation from the output of one of your mediation macros?"

These are based on the relative size and significance of the total and direct effect. "Full" or "Complete" and "Partial" mediation are outdated, 20th century concepts that have no place in 21st century mediation analysis. I recommend you avoid the use of these terms, and don't attempt to interpret your analysis based on the relative size and significance of the total and direct effects. For a discussion, see section 6.1 in Hayes (2013).
**Higher-Order Models - Hierarchical Component Models (HCM)**

- **Higher-order models** or **HCM** most often involve testing **second-order structures** that contain two layers of components.
- Complex phenomena.

![Diagram of Hierarchical Component Models](image)

![Diagram of Reflective and Formative Models](image)
Thank you!

Dr. Faizan Ali, Assistant Professor  
College of Hospitality and Tourism Leadership  
University of South Florida Sarasota-Manatee  
faizanali@usf.edu

Dr. Cihan Cobanoglu, McKibbon Endowed Chair Professor  
College of Hospitality and Tourism Leadership  
University of South Florida Sarasota-Manatee  
cihan@usf.edu

http://conference.anahei.org/webinar