ISSUES AND PROCEDURES IN ADOPTING STRUCTURAL EQUATION MODELING TECHNIQUE

Siu Loon HOE[1]

PhD, Adjunct Research Fellow, UWA Business School,
The University of Western Australia, Australia

E-mail: hoesl@graduate.uwa.edu.au

Abstract: *When applying structural equation modeling (SEM) technique for analytical procedures, various issues are involved. These issues may concern sample size, overall fit indices and approach. Initiates of SEM may find it somewhat daunting in resolving these technical issues. The purpose of this paper is to highlight key issues in adopting SEM technique and various approaches available. This paper provides a discussion on the sample size, fit indices, standardized paths, unidimensionality test and various approaches in relation to SEM. It is hoped that having reviewed the paper, new researchers can devote more time to data analysis instead of procedural issues involved.*

Key words: Structural equation modeling; fit indices; unidimensionality; eigenvalue

Introduction

The modern positivist paradigm for conducting scientific research rests on developing sound theoretical frameworks followed by rigorous testing of these theories. One often adopted technique is structural equation modeling (SEM). SEM is a powerful statistical technique that combines measurement model or confirmatory factor analysis (CFA) and structural model into a simultaneous statistical test.

SEM is particularly valuable in inferential data analysis and hypothesis testing where the pattern of inter-relationships among the study constructs are specified a priori and grounded in established theory. It has the flexibility to model relationships among multiple predictor and criterion variables, and statistically tests a priori theoretical assumptions against empirical data through CFA (Chin, 1998). In most cases, the method is applied to test 'causal' relationships among variables.

In applying SEM technique for analytical procedures, many issues are involved. These issues may concern various overall fit indices and selection of the appropriate approach (Lei & Wu, 2007). Initiates of SEM may find it somewhat daunting in resolving these issues. The purpose of this paper is to highlight key issues in adopting SEM technique and various approaches so that researchers can devote more time to data analysis instead of dealing with procedural issues. This paper provides a discussion on the sample size, fit indices, standardized paths, unidimensionality test and various approaches in relation to SEM. Wherever appropriate, numerical examples were provided to illustrate the issues and procedures highlighted.

Structural equation modeling and sample size

Typically, a hypothesized model is tested with a linear equation system through SEM. This method of study investigates the extent to which variations in one variable corresponded to variations in one or more variables based on correlation co-efficient. SEM is usually used because it permits the measurement of several variables and their inter-relationships simultaneously. It is more versatile than other multivariate techniques because it allows for simultaneous, multiple dependent relationships between variables.

The hypothesized causal relationships can be tested among the theoretical constructs using software programs such as EQS (Bentler, 2002) to estimate and evaluate the structural portion of the model. The raw data for the variables are input into the software to generate the iterations, goodness-of-fit indices and standardized paths. The various variables are usually summated scales where the attributes measuring a common underlying construct are summed and divided by the number of items.

McQuitty (2004) suggested that it is important to determine the minimum sample size required in order to achieve a desired level of statistical power with a given model prior to data collection. Schreiber et al (2006) mentioned that although sample size needed is affected by the normality of the data and estimation method that researchers use, the generally agreed-on value is 10 participants for every free parameter estimated. Although there is little consensus on the recommended sample size for SEM (Sivo et al, 2006), Garver and Mentzer (1999), and Hoelter (1983) proposed a 'critical sample size' of 200. In other words, as a rule of thumb, any number above 200 is understood to provide sufficient statistical power for data analysis.

Fit indices

There are several indicators of goodness-of-fit and most SEM scholars recommend evaluating the models by observing more than one of these indicators (Bentler & Wu, 2002; Hair et al. 1998). Marsh, Balla and McDonald (1988) proposed that the criteria for ideal fit indices are relative independence of sample size, accuracy and consistency to assess different models, and ease of interpretation aided by a well defined pre-set range. Based on this stated criteria, Garver and Mentzer (1999) recommended the nonnormed fit index (NNFI); the comparative fit index (CFI), and the root mean squared approximation of error (RMSEA). Therefore, the commonly applied fit indices are NNFI and CFI (>0.90 indicates good fit), RMSEA (<0.08 indicates acceptable fit), and commonly used χ^2 statistic (χ^2/ d.f. ratio of 3 or less).

The NNFI, also known as the Tucker-Lewis index, compares a proposed model's fit to a nested baseline or null model. Additionally, NNFI measures parsimony by assessing the degrees of freedom from the proposed model to the degrees of freedom of the null model. NNFI also seems resilient against variations in sample size and, thus, is highly recommended. An acceptable threshold for this index is 0.90 or greater. Bentler (1990) developed the CFI as a noncentrality parameter-based index to overcome the limitation of sample size effects. This index ranges from 0 to 1, with 0.90 or greater representing an acceptable fit. RMSEA is an extremely informative criterion in evaluating model fit. The RMSEA index measures the discrepancy between the observed and estimated covariance matrices per degree of freedom (Steiger, 1990). It measures the discrepancy in terms of the population and not the sample. Thus, the value of this fit index is expected to better approximate or estimate the population and not be affected by sample size. Again, values run on a continuum from 0 to 1. Values less than 0.05 indicate good fit, values up to 0.08 reasonable fit and ones between 0.08 and 0.10 indicate mediocre fit.

Chi-square (χ^2) is the most common method of evaluating goodness-of-fit. A low χ^2 value, indicating nonsignificance, would point to a good fit. This is because chi-square test is used to assess actual and predicted matrices. Thus, non-significance means that there is no considerable difference between the actual and predicted matrices (Hair et al., 1998). Therefore, low χ^2 values, which result in significance levels greater than 0.05 or 0.01, indicate that actual and predicted inputs are not statistically different. The significance levels of 0.1 or 0.2 should be exceeded before nonsignificance is confirmed (Fornell, 1983).

In terms of a model's goodness-of-fit, p-values indicate whether the model is significantly different than the null model. In statistics, the null is usually '0'. This, however, is not necessarily so in SEM. The null hypothesis is the hypothesized model in which the parameters were set up for the hypothesized model, indicating whether a path should exist or not between variables. A high ρ-value, or a value larger than zero, would mean that the null hypothesis is rejected leading to a high probability that it would be wrong in doing so (MacLean & Gray, 1998). Thus, a high 'ρ' is good as it indicates that the observed model is not significantly different from what was expected. Conversely, a low ρ-value, or one close to zero, implies a 'bad model' because the null hypothesis is rejected with a low probability of being wrong in reaching that conclusion.

There is a limitation to the chi-square test. The χ^2 is highly sensitive to sample size especially if the observations are greater than 200. An alternate evaluation of the χ^2 statistic is to examine the ratio of χ^2 to the degrees of freedom (d.f.) for the model (Joreskog & Sorbom, 1993). A small χ^2 value relative to its degree of freedom is indicative of good fit. Kline (1998) suggested that a χ^2/ d.f. ratio of 3 or less is a reasonably good indicator of model fit.

As an example, fit indices were generated for a hypothesized model using SEM technique and presented in Table 1.

Table 1. Examples of fit indices of a hypothesized model

Model	d.f.	χ^2	NNFI	CFI	RMSEA
Hypothesized Model	18	416.69	-0.043	0.330	0.321

As indicated in these results, the goodness of fit measures for the hypothesized model came nowhere near the minimum requirements for the benchmark fit indices. The χ^2 value is 416.69 based on 18 d.f. and probability value (ρ) for χ^2 statistic is less than 0.001. The NNFI = -0.043, CFI = 0.330 and RMSEA = 0.321 which indicate a 'bad fit' for the hypothesized model (Bentler, 1990). Since the hypothesized model did not have a 'good fit', it was rejected.

Standardized paths and test for unidimensionality

Besides the 'goodness-of-fit' indices, SEM may also be used to look at paths among the variables. The causal paths can be evaluated in terms of statistical significance and strength using standardized path coefficient that range between -1 and +1. Based on α of 0.05, the test statistic generated from the EQS output should be greater than $\pm$ 1.96 to indicate that the null hypothesis can be rejected. The rejection of the null hypothesis means that the structural coefficient is not zero (Bentler, 2002; Byrne, 1994). After reviewing the statistical significance of the standardized paths, the strength of relationships among the variables can then be reviewed. According to Chin (1998), standardized paths should be at least 0.20 and ideally above 0.30 in order to be considered meaningful for discussion.

As another example, the standardized paths of a hypothesized model were computed and shown in Table 2.

Table 2. Examples of standardized paths of a hypothesized model

Hypothesis	Causal Path	Standardized Path Coefficient
H1	Informal knowledge acquisition → Market knowledge use	0.209***
H2	Informal knowledge dissemination → Market knowledge use	0.210**
H3	Shared vision → Informal knowledge acquisition	Nonsignificant
H4	Shared vision → Informal knowledge dissemination	0.452****
H5	Interpersonal trust → Informal knowledge acquisition	0.150*
H6	Interpersonal trust → Informal knowledge dissemination	Nonsignificant

Note: * $\rho < 0.05$
** $\rho < 0.01$
*** $\rho < 0.001$
**** $\rho < 0.0001$

In this hypothesized model, four of the paths were statistically significant. Comparing these results with the hypotheses, the standardized path coefficient of 0.209 seems to indicate that informal knowledge acquisition is positively associated with market knowledge use (H1). Also, the standardized path coefficient of 0.210 suggests that informal knowledge dissemination is also positively associated with market knowledge use (H2).

The path between shared vision and informal knowledge acquisition was not statistically significant indicating that shared vision is not positively associated with informal knowledge acquisition (H3). The standardized path coefficient between shared vision and informal knowledge acquisition was 0.452. This seems to suggest that shared vision is positively associated with informal knowledge dissemination (H4).

Although the results point to a significant association between interpersonal trust and informal knowledge acquisition (H5) with standardized path coefficient of 0.150, this path adds minimal value to the understanding of the relationship between interpersonal trust and informal knowledge acquisition. The reason is because the standardized path coefficient failed to meet the minimum benchmark for path strength. Chin (1998) has proposed that standardized paths should be at least 0.20 and ideally above 0.30 in order to be considered meaningful. The path between interpersonal trust and informal knowledge dissemination was nonsignificant suggesting that interpersonal trust is not positively associated with informal knowledge dissemination (H6).

Once the overall model fit has been evaluated, the variables can be assessed for unidimensionality. In accordance to accepted practice (Anderson, 1987; Churchill 1979; Gerbing & Anderson, 1988), the property of scales for unidimensionality was assessed. Unidimensionality is referred to as the existence of one construct underlying a set of items. Germain, Droge and Daugherty (1994) suggested the use of principal components analysis to test for unidimensionality. Based on this suggestion, each variable should be separately subject to principal components analyses to determine the eigenvalue. As a rule, eigenvalues that are greater than 1 provide support for the unidimensionality of these scales.

As an illustration, the eight variables in a particular study were separately subject to principal components analyses and the eigenvalues presented in Table 3.

Table 3. Examples of eigenvalues of measures

Measure	Component	Initial Eigenvalues Total	% of Variance	Cumulative %
Market Knowledge Use	1	4.048	57.824	57.824
	2	1.040	114.852	72.676
	3	0.726	10.365	83.041
	4	0.425	6.071	89.112
	5	0.327	4.664	93.776
	6	0.247	3.528	97.304
	7	0.189	2.696	100.00
Structural Knowledge Acquisition	1	1.925	64.155	64.155
	2	0.651	21.702	85.858
	3	0.424	14.142	100.00
Structural Knowledge Dissemination	1	3.094	61.888	61.888
	2	0.660	13.193	75.082
	3	0.461	9.214	84.296
	4	0.427	8.547	92.836
	5	0.358	7.164	100.00
Informal Knowledge Acquisition	1	2.620	52.396	52.396
	2	0.910	18.191	70.587
	3	0.627	12.549	83.136
	4	0.507	10.139	93.275
	5	0.336	6.725	100.00
Informal Knowledge Dissemination	1	1.786	59.539	59.539
	2	0.731	24.362	83.901
	3	0.483	16.099	100.00
Shared Vision	1	3.147	78.679	78.679
	2	0.317	7.936	86.615
	3	0.283	7.081	93.697
	4	0.252	78.679	100.00
Interpersonal Trust	1	3.272	81.804	81.804
	2	0.398	9.961	91.765
	3	0.198	4.959	96.725
	4	0.131	3.275	100.00
Perceived Importance of Market Knowledge	1	4.049	80.974	80.974
	2	0.473	9.462	90.436
	3	0.205	4.097	94.534
	4	0.156	3.122	97.656
	5	0.117	2.344	100.00

Except for market knowledge use, only the first eigenvalue was greater than 1 for all the rest of the scales. This provided support for the unidimensionality of these scales. For market knowledge use, two eigenvalues were greater than 1 but the second eigenvalue was only 1.04. Since second eigenvalue is close to 1 and this is a measure that has been used extensively in previous research, it is reasonable to accept the unidimensionality of this scale.

Various approaches

There are several approaches to weighing individual scale items in SEM. These approaches include total aggregation, total disaggregation, partial disaggregation, and partial aggregation (Bagozzi & Heatherton, 1994).

Under the total aggregation approach, all the items are summed into a single indicator or latent variable. The items are arbitrarily given the same weight as is traditionally done in non-SEM research. On the other hand, under the total disaggregation approach, each individual item is taken separately as an individual indicator. Under this approach, SEM weighs the individual items to optimize contribution to the latent variables. Bagozzi and Foxall (1996) mentioned that one of the main drawbacks of the total aggregation approach for CFA is that information is lost and the distinctiveness of the components is obscured. A disadvantage of the total disaggregation approach is it is very sensitive to measurement error which makes it more difficult to obtain satisfactory model fits. The partial aggregation approach retains the idea of a single underlying factor where dimensions of the construct are organized hierarchically as indicators of the factor. Similar to the total aggregation approach, the main drawback is that the unique dimension of the construct may be obscured.

To overcome such limitations, Bagozzi and Heatherton (1994) recommended that a partial disaggregation approach be used. Partial disaggregation is a practical SEM application that allows the use of a large number of indicators to represent a latent variable (Garver & Mentzer, 1999). It is an intermediary level of analysis between the total aggregation and total disaggregation approach. Unlike a total aggregation approach, the partial disaggregation approach helps to reduce the number of parameters to be estimated and to capitalize on the increase in reliability resulting when items on sub-scales are summed. Each dimension can be measured with two indicators wherein each indicator is itself the sum of multiple items.

Bagozzi and Heatherton (1994) suggested between five to seven items that can be randomly divided into two components for partial disaggregation. If there are more than nine items, then there could be three or more components. Based on the partial disaggregation procedures, there is no need to divide the aggregated total by the number of items so that each latent variable is an average score. Thus, the items for each of the latent variables can be randomly grouped into two components each based on odd and even sequence and the raw data input into the measurement model.

Conclusion

This paper has provided a short discussion on the sample size, fit indices, standardized paths, unidimensionality test and various approaches in relation to SEM. Some examples were provided to illustrate the issues highlighted. Through a better understanding of the issues and procedures in adopting the structural equation technique, researchers will be able devote more time to data analysis rather than resolving procedural issues.

References

1. Anderson, J. C. **An approach for confirmatory measurement and structural equation modeling of organizational properties,** Management Science, 33, 4, 1987, pp. 525-541
2. Bagozzi, R.P. and Heatherton, T.F. **A general approach to representing multifaceted personality constructs: Application to state self-esteem,** Structural Equation Modeling, 1, 1, 1994, pp. 35-67
3. Bagozzi, R. P. and Foxall, G. R. **Construct validation of a measure of adaptive-innovative cognitive styles in consumption,** International Journal of Research in Marketing, 13, 1996, pp. 201-213
4. Bentler, P. M. **Comparative fit indexes in structural models,** Psychological Bulletin, 107, 2, 1990, pp. 238-246
5. Bentler, P. M. **EQS Structural Equations Program Manual,** Encino, CA, Mulitvariate Software, 2002
6. Bentler, P. M. and Wu, E. J. C. **EQS 6 for Windows User's Guide,** Encino, CA: Mulitvariate Software, 2002
7. Byrne, B. M. **Structural Equation Modeling With EQS And EQS/Windows,** Newbury Park, Sage, 1994
8. Chin, W. W. **Issues and opinion on structural equation modeling,** MIS Quarterly, 22, 1, 1998, pp. 7-16
9. Churchill, G. Jr. **A paradigm for developing better measures of marketing constructs,** Journal of Marketing Research, 16, 1, 1979, pp. 64-73
10. Fornell, C. **Issues in the application of covariance structure analysis: A comment,** Journal of Consumer Research, 9, 4, 1983, pp. 443-448
11. Garver, M. S. and Mentzer, J.T. **Logistics research methods: Employing structural equation modeling to test for construct validity,** Journal of Business Logistics, 20, 1, 1999, pp. 33-57
12. Gerbing, D. W. and Anderson, J. C. **An updated paradigm for scale development incorporating unidimensionality and its assessment,** Journal of Marketing Research, 25, 1, 1988, pp. 186-192
13. Germain, R., Droge, C. and Daugherty, P. J. **The effect of just-in-time selling on organizational structure: An empirical investigation,** Journal of Marketing Research, 31, 4, 1994, pp. 471-483
14. Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. and Black, W. C. **Multivariate Data Analysis** (5th ed.), New Jersey, Prentice-Hall, 1998
15. Hoelter, D. R. **The analysis of covariance structures: Goodness-of-fit indices,** Sociological Methods and Research, 11, 1983, pp. 325–344
16. Joreskog, K. G. and Sorbom, D. **LISREL6 - Computer Program,** Moorseville, IN, Scientific Software, 1986
17. Kline, R. B. **Principles And Practice Of Structural Equation Modeling,** New York, Guilford Press, 1998
18. Lei, P. W. and Wu, Q. **An NCME Instructional Module on Introduction to Structural Equation Modeling: Issues and Practical Considerations,** Educational Measurement, Issues and Practice, 26, 3, 2007, pp. 33-44
19. MacLean, S. and Gray, K. **Structural equation modeling in market research,** Journal of the Australian Market Research Society, 1998, Retrieved September 26, 2001, from World Wide Web: http://www.smallwaters.com/whitepapers/marketing
20. Marsh, H. W., Balla, J. R. and McDonald, R. P. **Goodness-of-fit in confirmatory factor analysis: The effect of sample size,** Psychological Bulletin, 103, 3, 1988, pp. 391-410

21. McQuitty, S. **Statistical power and structural equation models in business research,** Journal of Business Research, 57, 2, 2004, pp. 175-183
22. Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A. and King, J. **Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review,** The Journal of Educational Research, 99, 6, 2006, pp. 323-337
23. Sivo, S. A., Fan, X. T., Witta, E .L. and Willse, J. T. **The Search for 'Optimal' Cutoff Properties: Fit Index Criteria in Structural Equation Modeling,** The Journal of Experimental Education, 74, 3, 2006, pp. 267-289
24. Steiger, J. H. **Structural model evaluation and modification: an interval estimation approach,** Multivariate Behavioral Research, 25, 1990, pp. 173-180

---

[1] Siu Loon Hoe is an Adjunct Research Fellow at UWA Business School, The University of Western Australia. His research interests are focused on organizational learning and boundary spanning. Siu Loon holds a BSc from National University of Singapore, and an MBA (International Business) and PhD (Management) from The University of Western Australia. He is currently Head of Strategic Planning for a Pan-Asian consumer group. In recognition of his academic and professional achievements, Siu Loon was included in International Who's Who of Professionals 2001/02.