Available online at www.elixirpublishers.com (Elixir International Journal)

Social Science

Elixir Soc. Sci. 49 (2012) 9930-9934

ISSN: 2229-712X

Statistics for social science: structural equation modeling approach

Ahmed Audu Maiyaki
Department of Business Administration & Enrepreneurship, Bayero University, Kano, Nigeria.

ARTICLE INFO

Article history:
Received: 17 May 2012;
Received in revised form:
23 July 2012;
Accepted: 7 August 2012;

Keywords
Statistics,
Social science,
Structural Equation Modelling.

ABSTRACT

The paper attempts at introducing Structural Equation Modelling (SEM) as an advanced analytical technique in the area of social and behavioural research. Although other multivariate analytical techniques are equally being employed by the social and behavioural scientist, SEM provides a more holistic analysis and handles complex models. A literature review was conducted and consequently, the procedure, merits and demerits of SEM were revealed. It was concluded that the SEM is superior to other multivariate analytical techniques such as Analysis of Variance, Multivariate Analysis of Variance, Factor Analysis and Regressions.

Introduction

Statistical methods being employed in the social science research are diverse just as the information and the type of data involved are varied. This is partly due to the fact the social science cover a wide range of disciplines such as Economics, Management, Political Science, Sociology, Communication, Psychology, Education and Anthropology among others (O'Rourke, Hatcher & Stepanski, 2005). Additionally, the type of data and information being gathered in social research could be quantitative or qualitative. Under quantitative, the nature of data could be of nominal scale, ordinal, interval and ratio (Sekaran & Bougie, 2010). Thus, the nature of data available for social research usually determines the type of statistics to be adopted; whether descriptive or inferential/ parametric or non-parametric. To this end, the power for statistical analysis has revolutionized the ways in which behavioral and social scientists plan, conduct, and evaluate their research (Davey & Savla, 2010).

Despite this diversity in the area of research and the nature of data used, social science research still has some common characteristics. For example, regardless of the discipline under investigation, most of the research in social science basically constitutes collection of data from the field and analysing the data to derive meaning out of it (O'Rourke et al., 2005). Similarly, most social scientists use a common language when conducting and reporting their research findings. For instance, researchers in both psychology and management speak of "testing null hypotheses" and "obtaining statistically significant *p* values."

Due to the advancement of technology, researchers do not have to analyse data manually. A number of computer programmes and softwares have been developed for data analysis.

Furthermore, a primary objective of multivariate analysis is to increase explanatory power and statistical efficiency in research. Although, techniques such as factor analysis, analysis of variance, multivariate analysis of variance and multiple regressions among others are used to address a number of theoretical and practical questions; however, the techniques are inadequate in resolving multiple relationships at a time (Hair, Black, Babin & Anderson, 2010). In other words, a technique among the aforementioned can not be used to test entire theory and at the same time considering all possible information. Structural Equation Modelling (SEM) is one of the advanced statistical techniques that assess a series of multiple dependent relationships simultaneously. The present trend is tend to be increasingly growing popularity of using more sophisticated analytical technique like SEM in the services marketing related research (Hooper, Coughlan & Mullen, 2008; Nel, Heerden, Chan, Ghazisaeedi, Halvorson & Steyn, 2011).

To this end, this chapter aims at introducing SEM as an advanced statistical tool in the area of social and behavioural sciences.

Literature review: structural equation modelling (SEM)

According to Tabachnick and fidell (2007) Structural Equation Modelling (SEM) is a collection of statistical techniques that allow a set of relationships between one or more Independent Variables (IVs), either continuous or discrete, and one or more Dependent Variables (DVs), either continuous or discrete, to be examined. Additionally, in SEM both IVs and DVs can be analysed as factors or measured variables; just as the technique evaluates whether the model provides a reasonable fit to the data and the contribution of each of the IVs to the DVs. SEM examines the structure of interrelationships expressed in a series of equations similar to those of multiple regressions. Put differently, it provides estimates for a series of separate but interdependent, multiple regression equations simultaneously by specifying the structural model (Hair, Bush & Ortinau, 2006).

In SEM analysis, variables are categorised into observed and unobserved variables. While observed variables can be directly measured and are represented by items or questions, unobserved or latent variable can not be measured directly. Latent construct/variable is normally measured indirectly through multiple observed variables which are also known as indicators (Hair et al. 2006). Similarly, in SEM parlance, all the independent variables are referred to as exogenous, while the dependent variables are called endogenous variables. The exogenous variables are determined by the factors outside the

Tele:
E-mail addresses: aamaiyakiy@gmail.com,aamaiyaki@yahoo.com

model and thus, do not receive path (one headed arrow) from any other variable. On the other hand, endogenous constructs are determined by factors within the model, this is represented by receiving path from other variables (Hair et al. 2006). The first stage in SEM analysis is model specification via Confirmatory Factor Analysis (CFA) instead of Exploratory Factor Analysis (EFA) which is normally conducted in regression type of analysis (Tabachnick & Fidell, 2007). Hence, the model is estimated, evaluated and probably re-specified with the ultimate aim of testing the model and hypotheses.

Additionally, in SEM there are majorly two types of models, measurement and structural. Firstly, the measurement model has to do with the relationships between the observable variables and the unobservable variables in the model. Thus, measurement model helps in assessing the relationships between constructs.

A measurement model is a typical CFA and frequently considered to be the "null model" (Salim, 2007). In the null model, the covariances for the latent variables are assumed to be zero. CFA tests a measurement model by assessing the validity of individual measures based on the overall model's fit and other evidence of construct validity (Hair et al. 2010). An example of measurement model graphics is provided in figure 1.

The second model is known as the structural model and it comprises the exogenous and endogenous variables and it emphasises on the model fit. This type of model is conceptual representation of the structural relationships between constructs based upon an established theory. An example of the structural model is shown in figure 2. In summary, the aforementioned types of models made up of the two-basic steps in conducting Structural Equation Modeling (SEM).

In CFA the tests can either be that of "first-order" or "Second-order" factors. The first-order CFA is designed to test the multidimensionality of a construct, through the covariances between the observed variables and the latent construct (Byrne, 2010; Hair et al. 2010).

On the other hand Second-order test contains two layers of latent constructs and theoretically it can be extended to infinite multiple layers (Loehlin, 2004). However, researchers more often than not perform Second-order test.

According to Hair et al. (2010) and Kline (2011) a minimum of three first-order (first level) constructs are needed to conduct a single higher-order test and that each first-order variable should have at least 3 items. Similarly, with a reflective second-order all first-order constructs which are now the indicators of the second-order factor are expected to move together or covary just like the observed items indicating the first-order constructs.

Finally, second-order test may be appropriate when there is theoretical basis to expect multiple layers of a particular constructs exist; and all the first-order constructs are expected to affect other nomologically related factors in the same way. It should however be noted that SEM technique of analysis is highly dependent on the theory because it is considered as a confirmatory analysis. Support of theory is needed in specifying relationships in both the measurement and structural models and also in modifying the proposed relationships (Hair et al., 2010).

Figure 1: An Example of Measurement Model

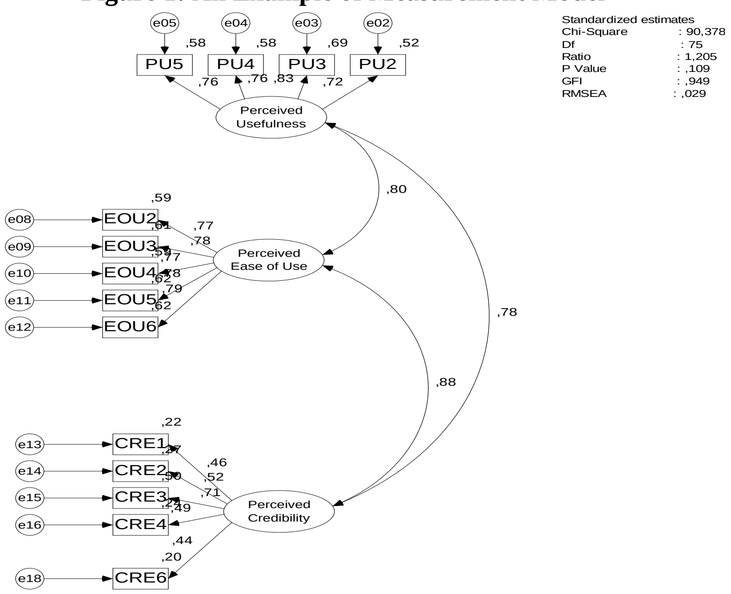
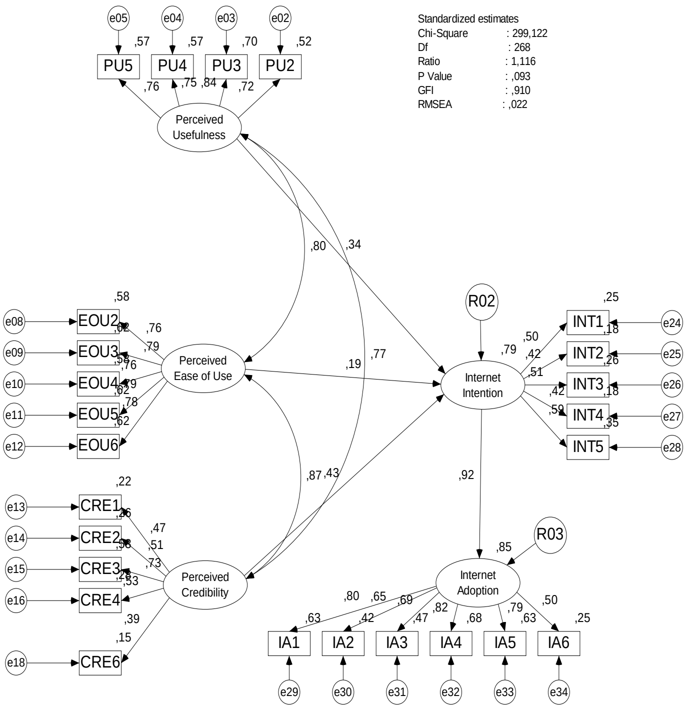


Figure 2: A typical Structural Model



Goodness of Fit Indices

Assessment of model validity is another important exercise in SEM that needs to be explained. A model is valid when it fits the empirical data and this is done with the aid of Goodness-of-Fit (GOF) indices. There are a number of GOF indices which researchers have developed over the years to reflect various ways in which the model can be able to represent the data. GOF indices are basically categorised into three classes namely: absolute fit indices, incremental fit indices and parsimonious fit indices (see table 1). Absolute fit indices are a direct measure of how well a model fit the empirical data. Therefore, they provide the most basic assessment of how well a theory fit an empirical data (Hair et al. 2010). However, they do not compare the GOF of different models rather each model is assessed independently. Chi-square statistics is the most fundamental absolute fit index, but it is highly sensitive to both the sample size and the number of observed variables to the extent that even minor misspecification could lead to poor model fit with large sample and vice-versa (Davey & Savla, 2010; Tabachnick & Fidell, 2007; Hair et al. 2010). Similarly, Hair et al. (2010) added that mere achieving p-value via non-significant in the value of Chi-square (X^2) does no always guarantee good model fit; because simple models with small sample size have a tendency towards a non-significant X^2 even if they do not meet other validity criteria.

A model is considered to have a good fit if the X^2 statistics is not significant through P-value (i.e. P>0.05). This however, is difficult to achieve with large sample size and large number of observed variables. Thus, Hair et al. (2010) observed that as N increases so does the X^2 value even if the differences between matrices are identical. They added that, the X^2 statistics is also likely to be greater if the number of observed variables increases. Due to the aforementioned limitation associated with X^2, other measures of model fit have been developed in order to rectify the bias against large sample and model complexity (Byrne, 2010). To this end, Hair et al. (2010) concluded that the X^2 statistical test or resulting p-value is less meaningful as the sample size increases or the number of observed variables becomes large (see table 4.4).

Another absolute fit index is Goodnes-of-fit Index (GFI) and it was an early attempt to provide a fit statistic that was less sensitive to sample size. However, the index is still affected by sample size due to the effect of N on sampling distributions. The values of GFI range between 0 and 1 with higher values indicating better fit and the reverse is the case. Tanaka and Huba (1989) observe that GFI is analogues to R^2 in regression analysis. Like R^2 this index can also be adjusted for the number of parameters estimated in the model through AGFI.

Root Mean Square Error of Approximation (RMSEA) is a population based index and therefore relatively less sensitive to sample size. This index is equally one of the most widely used measures that attempts to correct the limitation of X^2 GOF with regards to large sample and large observed variables. Studies on several measures reveal that RMSEA is more appropriate to employ in confirmatory or competing models when sample size is large (Hair et al., 2010). The large sample in this sense is considered to be more than 500 respondents. The lower the value of RMSEA the better the model fit. Literature shows that the 0.08 and 0.05 or less as the cut off. Remarkably, RMSEA along with other indices such as RMR, SRMR are in the category of indices known as badness-of-fit measures where high values indicate poor fit while low values indicate good fit (Hair et al. 2010). Normed Chi-square (X^2/df) is yet another absolute GOF index which represents a ratio of X^2 on degree of freedom. Hence, it is more stable for it overcomes some of the limitations of X^2 highlighted earlier on (Hair et al., 2010). This is due to the fact that it measures Chi-square per the degree of freedom and also handles the problem of model complexity. An acceptable value ranges between 1to 3 and value below is an evidence of over fit (Hair et al., 2010; Muhamad, 2008).

The second category of indices are incremental or comparative fit indices because they compare between different models, basically between baseline (null) model and specified model. SEM provides a number of incremental fit indices as standard output. Among the fit indices in this category CFI and TLI are the most widely used (Hair et al. 2010). Comparative Fit Index (CFI) is an improved version of NFI and its values range between 0 and 1 with higher values indicating better fit and vice-versa. This index is relatively insensitive to model complexity. Generally, values above 0.9 are indication of acceptable fit. In the same vein, Tucker Lewis Index (TLI) takes into account of model complexity to some extent and has values range between 0 and 1. However, TLI compares the normed X^2 values for null model and the specified model. Models with higher values indicate better fit than those with lower ones. Similarly, Normed Fit Index (NFI) is among the family of incremental fit indices and it signifies the ratio of the difference in the X^2 value for the fitted model and the null model. NFI values ranges between 0 and 1 with the value approaching 0.95 indicating good fit, while a perfectly model fit give the value of 1. However, several shortcomings have been attributed to this index. For instance, with complex model and large sample size, the value tends to be artificially inflated (Davey & Savla, 2010; Hair et al., 2010). Consequently, this index is less used compared to the first two (CFI & TLI) in the category.

The last category of model fit indices is related to parsimony. According to Hair et al. (2010) a parsimonious model is important in order that a model might fit the data compared to a highly specified complex model. Put differently, this index shows which among competing models is best given its complexity vis-a-vis its fit value. Hence, the indices are meaningful when comparing between different models with varying degree of complexity. When used however, PNFI is the most frequently applied among the parsimonious fit index. According to Hair et al. (2010) a parsimony fit index is improved either by a better fit or by a simpler model which has fewer estimated parameters/paths. Other indices in this group include Akaike Information Criteria (AIC) and Consistent Akaike Information Criteria (CAIC). Small value of normed X^2 indicates that the model contains too many parameters (Muhamad, 2008). On the other hand, with respect to AIC and CAIC low values indicate better fit model.

Table 1: ***A Summary of Goodness of Fit Indices (GOF) and Acceptable Benchmark***

GOF Measure	Acceptable fit level	comments
Absolute Fit Measures Chi-square (X^2) statistics	P-value, p ≥ 0.05	Non-significant X^2 indicates that the empirical data is identical with the model
Goodness of Fit Index (GFI)	Value should be > 0.9	0 = poor fit, 1 = perfect fit
Root Mean Square Error of Approximation (RMSEA)	Value should be < 0.08	0 = perfect fit, 0.1 = poor fit. RMSEA ≤ 0.08 shows reasonable error of approximation
Root Mean Square Residual (RMSR)	Value should be < 0.08	Value closer to 0 shows better fit
Normed Chi-square (X^2/DF)	Ratio between 1 and 2 or between 3 and 5	Wheaton et al. (1977) suggested the ratio of approximately 5 or less as reasonable. Carmines & Mclver (1981) suggested the ratio in the range of 1 to 3 as acceptable
Incremental Fit Measures Adjusted Goodness of Fit Index (AGFI)	Value should be > 0.9	Value adjusted for df, ≥ 0.9 = good model fit
Normed Fit Index (NFI)	Value should be > 0.9	0 = poor fit, 1 = perfect fit
Tunker-Lewis	Value should be >	Lower value indicates poor fit while value that approaches 1

Indext (TLI)	0.9	shows better fit. It is not normed and hence, can fall below 0
Comparative Fit Index (CFI)	Value should be > 0.9	CFI value above 0.9 is associated with a model that fits well
Incremental Fit Index (IFI)	Value should be > 0.9	IFI value above 0.9 shows good fit
Parsimonious Fit Index Measures		
Akaike Information Criteria (AIC)	Smaller positive value indicate a better fit and more parsimony	AIC close to 0 indicates a very good model fit

Source: (Davey & Savla, 2010; Hair et al. 2010; Tabachnick & Fidell, 2007)

Achievement of Goodness of Fit Indices

Although, a mixed of GOF indices from different classes are used not all the indices are normally being achieved. The achievement of the required value for GOF indices sometimes depends on the sample size and model complexity (Hair et al., 2010). For instance, from table 2, it could be seen that a study with the sample size in excess of 250 and observes variables more than 12, p value is not likely to be achieved i.e. it will be significant ($p < 0.05$). On the other hand, if the sample size is less than 250 and the items equal to or less than 12, then the p value is expected to be achieved i.e. it will be not significant ($p > 0.05$) (Hair et al., 2010). See table 2.

Table 2: ***Possibility of achieving GOF indices vis-à-vis sample size and model complexity***

	N < 250			N > 250		
variables Indices	$M \leq 12$	$12<M<30$	$M \geq 30$	$M < 12$	$12<M<30$	$M \geq 30$
X^2	Insignificant p-values expected	Significant p-values even with good fit	Significant p-values even with good fit	Insignificant p-values expected	Significant p-values expected	Significant p-values expected
CFI or TLI	0.97 or better	0.95 or better	Above 0.92	0.95 or better	Above 0.92	Above 0.9
RNI	May not diagnose misspecification well	0.95 or better	Above 0.92	0.95 or better, not used with N>1,000	Above 0.92, not used with N>1,000	Above 0.9, not used with N>1,000
SRMR	Biased upward, use other indices	0.08 or less (with CFI of 0.95 or higher)	Less than 0.09 (with CFI above 0.92)	Biased upward; use other indices	0.08 or less (with CFI above 0.92)	0.08 or less (with CFI above 0.92)
RMSEA	Values < 0.08 with CFI = 0.97 or higher	Values < 0.08 with CFI of 0.95 or higher	Values < 0.08 with CFI of 0.92 or higher	Values < 0.07 with CFI of 0.097 or higher	Values < 0.07 with CFI of 0.92 or higher	Values < 0.07 with CFI of 0.9 or higher

Note: m=number of observed variables; N= number of observations

Source: Hair et al. (2010)

Advantages of Structural Equation Modeling

SEM offers a number of advantages as compared to other analytical techniques. For example, when relationships among factors are assessed, the relationships are free from measurement error because it is estimated and removed. Similarly, reliability of measurement can be accounted for clearly in the analysis after removing the measurement error. With complex and multidimensional model, SEM is the only analysis that allows complete and simultaneous test of all the relationships (Tabachnick & Fidell, 2007).

Missing data is a common problem in the social and behavioural sciences research to the extent that it is considered to be inevitably perennial problem more especially in longitudinal and multi-informant studies (Loehlin, 2004). An incomplete or missing data have significant effect in research related to the fields of psychology, sociology, human development, education, gerontology, nursing, and health sciences to the extent that it could affect the study design and implementation (Davey & Savla, 2010). Despite the fact that missing data is expected to reduce the statistical power, not all the missing data have equal impact and hence, some lead to greater loss of statistical power than others. In SEM, it is possible to answer questions such as how missing data may have affected the statistical power in a specific study, how to increase the power of a design in the presence or expectation of missing data, and how to identify the more statistically powerful design in the presence of missing data (Davey & Savla, 2010).

SEM is an extension of factor analysis and multiple regressions. According to Hair et al. (2010), SEM is useful more especially in testing theories that involve multiple equations in dependence relationships. For example, in a single model a variable can serve as an endogenous variable and also exogenous variable with regards to some other variable. Among the other multivariate techniques none can be used to test both measurement properties and assess theoretical relationships in a single technique.

Disadvantages of SEM

Despite all the aforementioned advantages of using SEM as an analytical technique, there are however some disadvantages attributed to it. Firstly, with the complex relationships in a model under investigation, complexity and ambiguity of analysis is resulted. The challenge of how the model that best represents the data reflects underlying theory/ model fit is yet unresolved (Hooper et al., 2008). Similarly, the difference in opinions regarding the use of fit indices. According to Hair et al. (2010) perhaps there is nothing more debated in SEM than what constitute goodness of fit. For instance, the large number of model-fit indices and lack of consistent guideline could lead to subjective selection of indices by researchers. In the same vein, the problem of cut off values for fit indices is still lingering.

Conclusions

In line with contemporary trend in social and behavioural science research, advance analytical techniques such as Structural Equation Modeling are continuously gaining ground. Employing such statistical techniques would play an important role in ensuring effective and efficient research outcomes. Although, some issues regarding the threshold and the adequacy of the Goodness of fit indices are not yet standardized, the tremendous benefits of the technique over other multivariate techniques like Analysis of Variance, Multivariate Analysis of Variance, Factor Analysis and Regressions made it to become superior and endeared by the contemporary academic researchers.

Reference

Byrne, B. M. (2010). *Structural Equation Modelling with AMOS: Basic concepts, application, and programming (2nd ed.).* New York: Rouledge Taylor & Francis Group

Davey, A. & Savla, J. (2010). *Statistical power analysis with missing data: A structural equation modelling approach*. New York: Rouledge Taylor & Francis Group

Hair, Jr., J. F., Bush, R. P., & Ortinau, D. J. (2006). *Marketing research within a changing information environment (3rd ed.).* New York: McGraw-Hill/Irwin

Hair, Jr., J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Uppersaddle River, New Jersey: Prentice Hall

Hooper, D., Coughlan, J. & Mullen, M. R. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *The Electronic Journal of Business Research Methods,* 6(1), 53 - 60

Kline, R. B. (2011). *Principles and practice of Structural Equation Modelling* (3rd ed.). New York: The Guilford Press

Loehlin, J.C. (2004). *Latent variables models: An introduction to factor, path and structural equation analysis (4th ed.).* Lawrence Erlbaum Associates, Inc.: Mahwah, New Jersey

Muhamad, N. (2008). *Muslim consumers' motivation towards Islam and their cognitive processing of performing taboo behaviours*. Unpublished doctoral dissertation, University of Western Australia.

Nel, D., Heerden, G., Chan, A., Ghazisaeedi, M., Halvorson, W. & Steyn, P. (2011). Eleven years of scholarly research in the Journal of Services Marketing. *Journal of ServicesMarketing,* 25(1), 4–13

O'Rourke, N., Hatcher, L. & Stepanski, E. J. (2005). *A step-by-step approach to using SAS for univariate and multivariate statistics (2nd ed.).* Cary, NC: SAS Institute Inc.

Tabachnick, B.G. & Fidell, L.S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson Education Inc.

Tanaka J. S. & Huba G. J. (1989). A general coefficient of determination for covariance structure models under arbitrary GLS estimation. *British Journal of Mathematical and Statistical Psychology,* 42, 233-239.

Salim, L. (2007). *Store loyalty determinants of modern retailing in Indonesia: A Structural Equation Modelling Approach*. Unpublished doctoral thesis, Universiti Utara Malaysia

Sekaran, U. & Bougie, R. (2010). *Research methods for business: A skill building approach (5th ed.).* Chichester: John Willey & Sons Ltd