Applications of Structural Equation Modeling (SEM) in Humanities and Science Researches

Dr. Ayed Al Muala  
Department of Marketing, Applied Science University  
aied_muala@yahoo.com

Dr. Mamdouh Al Ziadat  
Department of Marketing, Applied Science University

Dr. Abbas N. Albarq  
Business Department, Imam Muhammad Ibn Saud Islamic University

Dr. Malek AL-Majali  
Department of Marketing, Mu’tah University

Abstract
This study tries to provide a plan to step from regression to a path analysis. Regression analysis sometimes provides and gives weak results using a default model. Path analysis allows you to specify a model and relationships between variables, SEM provides a method to specify a path analysis model, examine relationships between variables, and provides a comprehensive analysis of your data. In addition, SEM can measures any complex relationships, also can analysis and account more than mediating effect in many relationships. Furthermore, SEM approach that was discussed as new statistical technique for the empirical studies. Moreover, SEM still in start using from empirical researches, and the researcher wish using these new method for give fit model and good results. SEM can give results of all constructs in direct and indirect impact in model of study and results of mediating effect in any study. Additionally, if the component that is needed to be measured and evaluated is extremely hypothetical and conceptual such as social science research (perceptive measures such as satisfaction, happiness, and tiredness), SEM is highly recommended. This contrasts with multiple regressions that are designed to measure metric scales (such as price, cost, and temperature). Moreover, Confirmatory Factor Analysis CFA is employed to reduce the measurement of instrument error but SEM techniques are deployed to perform the CFA. To be clear, SEM employs a set of measures to achieve the model fit, used CFA to test convergent validity since CFA is stronger than EFA. This is supported by some previous studies such as Gerbing and Anderson (1988) who argued that CFA can supply a stricter explanation of unidimensionality than other techniques like coefficient alpha, correlation, and EFA (Exploratory Factor Analysis). Additionally, EFA is mainly an explanatory technique since no control over which variables are indicators of the latent variables. In contrary to EFA technique, CFA or measurement model in SEM provides an ability to have a complete control over specification for each construct’s indicators (McDonals & Ho, 2002). Additionally, Goodness of fit is "the degree to which the actual or observed input matrix (Covariances or
correlations) is predicted by the estimated model" (Hair et al., 2006, p. 580). According to Bollen (1989), the χ² likelihood ratio test, the Standardized Root Mean Residual (SRMR), the Goodness-of-Fit Index (GFI, CFI, and IFI) are the most frequently achieved measures. The following sections provide an overview of each of the achieved measures to explain the decisions obtained with regards to the models. The χ² likelihood ratio test, which is highly important a “badness-of-fit” test, is the most identified and apparent measure correlated with CFA.

**Keywords:** structural equation modelling (SEM), Humanities and Science Researches

**Introduction**

For more information from your data and more than regression analysis this study will show some Path analysis that can identify of relationships between variables. This study will review the basics of path analysis with structural equation modeling methodology and then present an example comparing results from regression analysis and path analysis. Additionally, this study tries to offer an alternative approach to regression analysis and a step-by-step approach to path analysis with structural equation modeling methodology (SEM). The goals are to present a powerful, flexible and comprehensive technique for investigating relationships between measured variables and to challenge you to design and plan research where SEM is an appropriate analysis tool. Furthermore, two goals in SEM are:

1) To understand the patterns of correlation/covariance among a set of variables and
2) To explain as much of their variance as possible with the model specified (Kline, 1998).

**STRUCTURAL EQUATION MODELLING (SEM)**

Structural equation modeling (SEM) is considered a family of statistic models that seem for explaining the relationship amongst multiple variables (Hair et al., 2006). Structural equation modeling process includes two main steps as follows;

i) Validating the measurement model
ii) Fitting the structural model (Kenis & Knoke, 2002).

The former is accomplished primarily through path analysis with latent variables. SEM is based on causal relationships (Hair et al., 1998, p. 592). These causal relationships explain how changes in variables (exogenous constructs) will result in changes in other variables (endogenous constructs). The causation among variables can be asserted through theoretical determination. Additionally, in SEM the theoretical based models are the structural model and the measurement model.

The SEM has developed to be one of the well known aspects in selecting a research methodology for investigating issues related to social and behavioural sciences (Baumgartner et al., 1996). Therefore, this methodology contains two major issues: i) the measurement (i.e. what are the things that need to be measured; how to measure them; and how are the reliability and validity conditions met) and ii) causal relationship among variables and the explanation because the variable is complex and unobserved. This is a function of SEM (Hair et al., 2006).
**Why SEM? Why not Multiple Regression?**

According to past literature and previous studies there are many reasons for selecting SEM as an instrument of analysis in a research study. Firstly, SEM analysis techniques are employed when the researcher includes a variety of factors or variables. Secondly, it is felt favourable and a choice of analysis instrument when a questionnaire is constructed to facilitate interval scales (Hair, et al., 2006; 1995).

Additionally, if the component that is needed to be measured and evaluated is extremely hypothetical and conceptual such as social science research (perceptive measures such as satisfaction, happiness, and tiredness), SEM is highly recommended. This contrasts with multiple regressions that are designed to measure metric scales (such as price, cost, and temperature).

Thirdly, the SEM consists of two models: measurement model and structural model (Hair, et al., 2006; 1995). The measurement model deals with the relationships between measured and latent variables, which specify indicators/items/scales for each construct and the assessment of construct validity. SEM is related to two kinds of errors as the result of the measurement and structural model known as measurement error and structural error, respectively. Structural error will be considered and added to the structural model because independent/dependent latent variables cannot predict the dependent variables perfectly.

Use of confirmatory factor analysis (CFA) will reduce the error effect that each latent having multiple indicators, and take all the constructs that are contained in the model as stimulus testing rather than individual coefficients. Also, CFA is useful to test models with multiple dependents (to model mediating variables and to deal with complicated data (such as non-normal data and incomplete data).

Linear regression analysis assumes that variables are evaluated with no errors. SEM includes multiple regression and factor analyses. According to Hair et al. (1995; 1998), SEM is an effective estimation instrument for a number of separate multiple regression equations evaluated simultaneously.

**CONFIRMATORY FACTOR ANALYSIS (CFA)**

Confirmatory Factor Analysis CFA is employed to reduce the measurement of instrument error but SEM techniques are deployed to perform the CFA. To be clear, SEM employs a set of measures to achieve the model fit, used CFA to test convergent validity since CFA is stronger than EFA. This is supported by some previous studies such as Gerbing and Anderson (1988) who argued that CFA can supply a stricter explanation of unidimensionality than other techniques like coefficient alpha, correlation, and EFA (Exploratory Factor Analysis). Additionally, EFA is mainly an explanatory technique since no control over which variables are indicators of the latent variables. In contrary to EFA technique, CFA or measurement model in SEM provides an ability to have a complete control over specification for each construct’s indicators (McDonals & Ho, 2002). On the other hand, the statistical test of the goodness of fitness for the anticipated confirmatory factor solution (which is not achievable with using EFA technique) can be done by using CFA through...
SEM (McDonald & Ho, 2002). Structural equation modeling process includes two main steps: validating the measurement model and structural model (Hair et al., 2006). CFA measurement model estimation is the first step of structural equation modeling (SEM). The purpose of the measurement model estimation is to specify the pattern by which each measure (indicator or item) loads on a particular factor (construct or variable) in original model in this study, and to assure the reliability and validity of measures and constructs.

The measurement model represents the degree to which the indicator (item) variables capture the essence of the latent factor. Moreover, a valid measurement model is the model which meets the requirements of psychometric soundness both reliability and validity of measures and constructs. The purpose of testing reliability and validity of measurement is to assure multiple items measure the hypothesized latent variables but not others. It is accomplished primarily through CFA (Byrne, 2001).

Some of studies used SPSS 14.0 and AMOS 6.0 to analyze the data. Following Anderson and Gerbing’s (1988) two-step approach, a measurement model was first estimated using Confirmatory Factor Analysis (CFA). After the assessment of the adequacy of the measurement model, Structural Equation Modeling (SEM) was utilized to find the best-fitting model and to test causal relationships. SEM, multivariate technique, combines aspects of multiple regression and factor analysis to assess a series of dependent relationships simultaneously, which is not possible using other multivariate techniques (e.g., multivariate analysis of variance, multiple regression, discriminant analysis, factor analysis, etc.) (Hair et al., 1998). This multivariate technique is particularly useful for modeling tests including several independent/dependent variables and mediators/moderators (Hair et al, 1998).

Modification indices are "the values calculated for each unestimated relationship possible in a specified model" (Hair et al., 2006, p. 581). Arbuckle (2005) emphasizes that the indices of each model modified fits pretty badly so we might look at modification indices (MI), for achieve P-value. Additional fit statistics from Amos recommended the model’s fit be able to be enhanced by using MI. Firstly, we ought to ascertain the MI for each item, and then delete the highest, after deletion the highest value we have to examine the text output for the second model. Thus, change the model and repeat the analysis, the model fit Chi-square will be decreased to achieve p-value and to fit the model. In addition, GFI, CFI and will be increased and RMSEA will be dropped. There are three types of CFA to testing the model which are individual, measurement, and structural model. Firstly, CFA for individual measures of each the construct such as (independents variables). Secondly, CFA for endogenous variables (dependent variables) and mediating effect with dependent variable. Thirdly, CFA for three structural models are hypothesis model (all exogenous and endogenous variables), underpinning theory (original theory), and generating model (all exogenous and endogenous variables).
SEM PROCEDURE

The SEM is a very popular multivariate approach. The first step is the model conceptualization, which handles constructed hypothesis (based on theory) as the main aspect for the relationships amongst latent variables, and other indicators. During this step, the model is developed in accordance to theory and empirical findings. The model should reflect the latent variables through any measured indicator. The path diagram development is the second step. It is deployed to achieve uncomplicated hypothesis visualization resulting from model conceptualization. Thus, model specification (generating model) is the third step in which it deals with the development of the measurement and structural design of the research problem. Causal relationship found in the latent variables should be discussed during the specification of the structural model.

The fourth step relates to model identification. The data is entirely tested to ensure that information collected is with quality and contains effective parameters for the model. The goal is to validate the specification model is not under-identified, or just-identified or over-identified.

The fifth step of SEM refers to the parameter estimation. It mainly handles the process of achieving evaluation for each parameter in the specified model. The reason is to achieve model-based covariance matrix that matches with the targeted covariance matrix. In order to determine the significance of the final parameter is significantly varied from zero, a significance test is used. Out of the available estimation models based on past literature review, maximum likelihood (ML) by Weighted Lasted (WLS) is the most popular one.

The sixth step discusses the goal of model fit testing. In this regard, the purpose of the model fit testing is to investigate the appropriateness (Goodness-of-fit or GOF) between the information collected and model. A GOF criterion relates to whether or not a model-based covariance matrix is similar to the observed covariance matrix. The GOF as a particular construct validity is an essential component in SEM process since it determines validity of the measurement model (Hair et al, 2006).

Model modification is the seventh phase of SEM. The reason is to do model modification in order to achieve better GOF. Re-specification relies mainly on the given modeling strategy. This is due to the fact that, in these outstanding features, SEM was taken into account to check the research model against the collected data so that it can assist in creating the model in the present research.

There are three main strategic frameworks for testing structural equation models (Jöreskog and Sorbom, 1993):

1. Hypothesis model (HM)
2. Alternative Model (AM)
3. Generating Model (GM)
GOODNESS OF FIT INDEX

Goodness of fit is "the degree to which the actual or observed input matrix (Covariances or correlations) is predicted by the estimated model" (Hair et al., 2006, p. 580). According to Bollen (1989), the $\chi^2$ likelihood ratio test, the Standardized Root Mean Residual (SRMR), the Goodness-of-Fit Index (GFI, CFI, and IFI) are the most frequently achieved measures. The following sections provide an overview of each of the achieved measures to explain the decisions obtained with regards to the models. The $\chi^2$ likelihood ratio test, which is highly important a "badness-of-fit" test, is the most identified and apparent measure correlated with CFA. The proposed model does not meet the requirements of the collected data very precisely if the p-value of $\chi^2$ is significant (i.e., <0.05), whereas it meets the demands of the collected data if p-value is >0.05 is achieved. According to Byrne (2001), there is progressing debate on whether a model that has a significant $\chi^2$ statistic must into consideration as valid.

Measuring data through using SEM usually takes place by, deploying goodness-of-fit (GOF) measures. The CFA contains important functions that may be deployed. These functions involve the following:

i) Examining the loading factors in every dimension/constructs in forming a variance,

ii) Confirming that the instrument themselves, how linked the instrument to the latent variables,

iii) Estimating the measurement error in the framework, and

iv) Validating and generate framework.

Therefore, CFA is most often deployed to determine whether the set of factors and the loading of construct items confirm the expected requirements that are needed to measure what is really measures the scale itself.

The researcher used Amos version 6.0 in this study. For measuring the exogenous variables and endogenous variables, there are many key terms of SEM such as Absolute fit index, Incremental Fit Level and Parsimonious Fit Level as shown in Table 1.1.

According to (Byrne, 2001), structural equation modeling can be divided into two sections: measurement model and structural model, the measurement model can measure the relationship between observed and unobserved variables. Likewise, structural model can measure the relationship between unobserved variables.
<table>
<thead>
<tr>
<th>Measures</th>
<th>Threshold Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Absolute Fit Level</strong></td>
<td></td>
</tr>
<tr>
<td>RMSEA</td>
<td>Less than 0.08</td>
</tr>
<tr>
<td>GFI</td>
<td>0.90 and Above</td>
</tr>
<tr>
<td>P- Value</td>
<td>P- Value ≥0.05</td>
</tr>
<tr>
<td><strong>Incremental Fit Level</strong></td>
<td></td>
</tr>
<tr>
<td>AGFI</td>
<td>0.90 and Above</td>
</tr>
<tr>
<td>CFI</td>
<td>0.90 and Above</td>
</tr>
<tr>
<td>TLI</td>
<td>0.90 and Above</td>
</tr>
<tr>
<td>NFI</td>
<td>0.90 and Above</td>
</tr>
<tr>
<td><strong>Parsimonious Fit Level</strong></td>
<td></td>
</tr>
<tr>
<td>CMIN/df</td>
<td>Less than 2.0</td>
</tr>
<tr>
<td>SMC (R²)</td>
<td>Bigger better</td>
</tr>
</tbody>
</table>

**Source:** Hair et al. (2006)

According to Hair et al. (2006, p. 753), as shows in Table 1.1 above, point out that the recommendation values of fit model are following:

i) Absolute Fit Index (AFI) assess whether a specific model leaves appreciable unexplained variance. Alkhaldi and AL-Faoury (2007), Indicates such as Chi-square (X²) accompanied by the model's degree of freedom and its probability, goodness of fit index (GFI), and the root mean square error of approximation (RMSEA) are usually utilized here. As following: RMESA <0.08, GFI > 0.90, P-value > 0.05.

ii) Incremental Fit Index (IFI) compares the (Generating Model (GM)) specific model to possible baseline or null models estimated using the same data. Indices such as Tucker-Lewis index (TLI), comparative fit index (CFI), and the incremental fit index (IFI) are commonly used GFI > 0.90, CFI > 0.90, TLI >0.90, NFI >0.90.
Parsimonious Fit Index (PFI) also it is called as adjusted measure, ask how well the model measures both fit and parsimony, taking into account the degree of freedom used in the model specification. Indices such as Normed fit index (the adjusted chi−square by the degree of freedom) can be used $CMIN/df < 3, SMC (R^2) >0.00$.

The main of this section is to investigate and examine the relationships between exogenous and endogenous variables. Firstly, the researcher measured the individual variable related to measurement model.

**HYPOTHESIS TESTING**

Most of studies meant to test the ten direct hypotheses, and two indirect hypotheses through mediating effect as mentioned earlier in chapter four.

**Direct Effect**

Direct effects are the relationship between two constructs with a single path (Hair et al., 2006). In other words, a direct effect is the effect variables have on one another in a direct relationship, in this study there are ten direct effects as mentioned earlier. To make sure that all paths in the model whether supported or rejected we should be certain with recommendation values of (C.R and $P$ value). Critical ratio (C.R) refers to the parameter estimate divided by an estimate of its standard error. C.R should be more than 1.96 to achieve recommendation value (Hair et al., 2006). This means that if C.R is more than 1.96, it supports this path, but if C.R is less than 1.96 that means it does not support the path or rejects the hypothesis.

However, probability level ($P$ value) provides a cut-off beyond which we assert that the findings are ‘statistically significant’ (by convention, this is $p<0.05$). Furthermore, "smaller $p$-values ( $p<0.01$) are sometimes called (highly significant) because they indicate that the observed difference would happen less than once in a hundred times if there was really no true difference" (Davies and Crombie, 2009, p. 4).

**Indirect Effect**

Regarding the indirect effects, there are those relationships that involve a sequence of relationships with at least one intervening construct involved (Baron and Kenny, 1986). However, this study shown that how examines the mediating effect. According to Brown (1996) to examine the indirect paths there are some steps to follow:

1. Total indirect effect, which consists of all paths from one variable to another that are intervened or mediated by at least one additional variable.

2. Second type is the total effect, which is the sum of the direct and total indirect effects in the model.

3. Third type is the standardized indirect effect, which is the decomposition of the total indirect effect into standardized indirect paths.
After comparing between indirect effect and direct effect, we can confirm if this path is a full mediator or not. This is through obtaining the values of both direct and indirect effect as in the example below.

Thus, there is an example to explain the mediating effect, as shown in figure 1.1, the mediating effect only can happen when there are three variables (independent variable is perceived behavior control (PBC), mediator variable is revisit intention (INT), and dependent variable is actual visit behavior (ACT), we can calculate the result of mediating effect during estimate each variable from output of analysis. Therefore, if the indirect effect (PBC→INT→ACT) is more than direct effect (PBC→ACT) and all paths are significant, then it considered as full mediator. In contrast, if indirect effect is less than direct effect, it is not considered a mediator.

**Figure 1.1:**

*A SEM Model with an Example of Direct and Indirect Effects*

**Conclusions**

the SEM approach that was discussed as new statistical technique for the quantities' studies. Moreover, structural equation modeling still in start using from empirical researches, and the researcher wish using these new method for give fit model and good results. SEM can give results of all constructs in model of study and results of mediating effect in any study.

**References**


