Estimating Three-Way Latent Interaction Effects in Structural Equations Modeling

Birhanu Worku Urge
MS300-0008/2015

A thesis submitted to the Pan African University Institute for Basic Sciences, Technology and Innovation in partial fulfillment of the requirements for the award of Degree of Master of Science in Mathematics (Statistics Option)

–2017–

DECLARATION

This Research is my original work and has not been submitted to any other University for examination.

Signature····················· Date : 16/2/2017
Urge Birhanu Worku

This research thesis has been submitted to the Pan African University Institute for Basic Sciences, Technology and Innovation with our approval as the supervisors:

Signature ·················· Date : ······

Prof. Kepher H. Makambi,
Georgetown University, USA

Signature ·················· Date : ······

Dr. Anthony K. Wanjoya
JKUAT, KENYA

DEDICATION

Dedicated to my wonderful brother (Gezaw Muleta), my sisters (Deribe Muleta, Aberash Worku, Adanu Muleta, Ima Worku). To my friends, Megersa Jirata, Daseleng Petros, Betiglu Mezgebu, Willy Attikey and Zelalem Hailu. Also dedicated to Caaltu Birhanu for standing with me throughout my graduate studies.

ACKNOWLEDGMENTS

First, give glory to God for everything. My heartfelt gratitude goes to my respective Supervisor Prof. Kepher H. Makambi for his unlimited constructive advice, suggestions, ideas and comments in all phases of the thesis. I cannot forget to appreciate my second supervisor Dr. Anthony K. Wanjoya for his instruction and guidance. Finally, I would like to extend my appreciation and thanks to Pan African University and Jomo Kenyata University for kind assistance in many ways.

ABSTRACT

A Monte Carlo simulation was performed for estimating and testing hypotheses of three-way interaction effect in latent variable regression models. A considerable amount of research has been done on estimation of simple interaction and quadratic effect in nonlinear structural equation. However no work has been done in estimating three-way continuous latent interaction. The present study extended to three-way continuous latent interaction in structural equation model. The latent moderated structural equation (LMS)approach was used to estimate the parameters of the three-way interaction in structural equation model and investigate the properties of the method under different conditions though simulations. The approach showed least bias, standard error,and root mean square error as indicator reliability and sample size increased. The power to detect interaction effect and type I error control were also manipulated showing that power increased as interaction effect size, sample size and latent covariance increased.

Contents

List of Tables

List of Abbreviations

SEM- Structural equation modeling

LMS- Latent moderated structural

RMSE-Root mean square error

LISREL- Linear structural relations

GAPI-Generelazed appended product indicator

2SLS-Two stage least square

EM- Expectation -maximization

SE-Bias- standard error bias

Chapter 1

Introduction

1.1 Background

Structural Equation Modeling (SEM) is a statistical method used for building models, making inference and quantify the relationship among latent variables that are not observable or cannot be measured precisely. But, measurement on the indicator variable related to those unobservable variables are available. This relationship began its bases as a method for modeling linear relationship. However, because of many of the models for observable variables in the social and behavioral sciences involves nonlinearity, it is unlikely that linear models are always enough to describe the relationship between latent variables. In fact, there has been growing literature (some of which described later in this thesis) developing different kinds of nonlinear structural equation models (polynomial or cross-product) and estimation methods for them.

Structural equation models are regression models comprised of measurement and structural models. The measurement model is the part which relates measured variables to latent variables. The structural model is the part that relates latent variables to one another. The variables in SEM are measured (observed, manifest) variables (indicators) and factors (latent variables). Variables and factors in SEM may be classified as independent variables or dependent variables where the names of measured variables are within rectangles and the names of factors in ellipses. Rectangles and ellipses are connected with lines having an arrowhead on one (unidirectional causation) or two (no specification of direction of causality) ends. Dependent variables are those which have one-way arrows pointing to them and independent variables are those which do not. Dependent variables have residuals (are not perfectly related to the other variables in the model) indicated by $\delta's$ (errors) pointing to measured variables and εs pointing to latent variables(see figure 1).

Extending SEM to include nonlinear functions allows researchers meaningfully and accurately model the relationship underlying their data. Kenny and Judd (1984) introduced the first statistical method aimed at producing estimates of parameters in a nonlinear structural equation model (specifically a quadratic or cross-product structural model with a linear measurement model) and their method attracted methodological discussions and alterations by a number of papers (For instance,Hayduck,(1987),Algina and Moulder, (2001),Wall and Amemiya, (2001),Marsh, Wen, and Hau, (2004),JÃűreskog and Yang,(1996)).

Techniques of Kenny-Judd and their extensions are workable solution for estimation of quadratic or interaction SEM but, the form of nonlinearity are not enough to account for complicated problems (Arminger and Muthen 1998). This thesis focused on the extension of Kenny-Judd model to three-way continuous latent interaction using latent moderated structural method.

1.2 Statement of the problem

It was found that, available literature are only for the specific quadratic and simple cross-product structural model and with certain assumptions on the form of the measurement model or the latent variable distribution. In contrast, no work has been done on extending to models with three-way continuous latent interaction in structural equation model. Accordingly, our research focused on estimating three-way interaction effect in nonlinear SEM using latent moderated structural equation method and investigate the statistical properties of the approach through simulation study.

1.3 Objective

1.3.1 General Objective

- To estimate three-way interaction effects in nonlinear structural equation models.

1.3.2 Specific objective

I To derive the variances-covariances for the nonlinear structural equation model consisting three-way interaction.

II To investigate the properties of the latent moderated structural equation method under different conditions (magnitude of the interaction,sample size,the correlation between the three first-order latent variables and indicator reliability.

III To test the hypothesis of three-way interaction effects.

1.4 Significance of the study

This study extended SEM analysis of interaction effects to three-way continuous latent interaction and can be applied in different areas especially in social and behavioral sciences. Hence, the extension to model three-way interaction with SEM will be helpfully for researchers to deal with interaction effects.

Chapter 2

Literature Review

2.1 Measurement model

In structural equation modeling, there are measurement models for the measured independent and dependent variables. The model for measured independent model is

$$x = \tau_x + \lambda_x \xi + \delta \tag{2.1}$$

and for measured dependent variable

$$y = \tau_y + \lambda_y \eta + \varepsilon \tag{2.2}$$

Where x is a qx1 vector of independent indicator variables,and y is a px1 vector of dependent indicator variables. λ_x is qxn matrix of regression coefficients predicting x by ξ and λ_y is pxm matrix of regression coefficients predicting y by η. τ_x is a qx1 vector of x-intercept and τ_y is a px1 vector of y-intercept. δ is a q x 1 vector of measurements errors of x and ε is a p x 1 vector of measurements error of y. The assumptions under measurement model are

- $E(\delta) = E(\varepsilon) = 0$
- $\delta \sim$ N(0,θ_δ)

- $Cov(\xi, \delta) = Cov(\eta, \varepsilon) = Cov(\delta, \varepsilon) = Cov(\delta, \eta) = 0$

The co-variance structure under measurement model is

- $Cov(X) = \Sigma_{qxq}$
- $Cov(\xi) = \phi_{nxn}$
- $Cov(\delta) = \theta_{qxq}$ assumed diagonal matrix

The mean and covariances of the indicators can be found using model equations. Let $E(\xi) = \kappa$ be n x 1 vector of the means of the latent exogenous variables. Then the mean of x given by $\mu_x = \tau_x + \lambda_x \kappa$. By centering the variables, the covariances of x is

$$\begin{aligned} Cov(x) &= Cov(\lambda_x \xi + \delta) \\ &= \lambda_x Cov(\xi) \lambda_x^T + Cov(\delta) \\ &= \lambda_x \phi \lambda_x^T + \theta_\delta \end{aligned}$$

Where T is transpose. The model identification under measurement model is depending on the total number of parameters to be estimated. Accordingly, there are $q \times n$ of λ_x, $\frac{n(n+1)}{2}$ of ϕ, q (assuming diagonal matrix) of θ_δ. The total number of parameters to be estimated $t = q \times n + \frac{n(n+1)}{2} + q$

The covariance of the predictors X ($Cov(X) = \Sigma_{qxq}$ is symmetric matrix and the number of redundant element is $\frac{q(q+1}{2}$. The decision to model identification depends on value of t.
$t > \frac{q(q+1)}{2}$ unidentified
$t = \frac{q(q+1)}{2}$ uniquely identified
$t < \frac{q(q+1)}{2}$ over identified

Hence, the decision is based on the value of t (see wiley,1973). Assuming that the model is identified and the predictor variable X multivariate normal distributed, maximum likelihood estimation method can be used to estimate model parameters in measurement model. $X \sim N_p(0, \Sigma)$

$f(x_i) = \frac{1}{(2\pi)^{p/2}\Sigma^{1/2}} e^{\frac{-1}{2}[x_i^T \Sigma^{-1} x_i]}$

The likelihood function $L(\Sigma) = \prod_{i=1}^{n} f(x_i)$ $L(\Sigma) = \frac{1}{(2\pi)^{np/2}\Sigma^{n/2}} e^{\frac{-1}{2}\sum_{i=1}^{n}(x_i^T \Sigma^{-1} x_i)}$ taking natural logarithm

$$lnL(\Sigma) = \frac{-np}{2} ln2\pi - \frac{n}{2} ln(\mid \Sigma \mid) - \frac{1}{2}[\sum_{i=1}^{n} x_i^T \Sigma^{-1} x_i]$$

$$byworkingwithparamterofinterest$$

$$\begin{aligned}
&= -\frac{n}{2} ln \mid \Sigma \mid - \frac{1}{2}[\sum_{i=1}^{n} x_i^T \Sigma^{-1} x_i] \\
&= -\frac{n}{2} ln \mid \Sigma \mid - \frac{n}{2}[\sum_{i=1}^{n} \frac{1}{n} x_i^T \Sigma^{-1} x_i] \\
&= -\frac{n}{2} ln \mid \Sigma \mid - \frac{n}{2}[tr(\frac{1}{n} x_i^T x_i \Sigma^{-1})] \\
&= -\frac{n}{2} ln \mid \Sigma \mid - \frac{n}{2} tr(S\Sigma^{-1}), S = \frac{1}{n} x_i^T x_i \\
&= -\frac{n}{2}(ln \mid \Sigma \mid + tr(S\Sigma^{-1}))
\end{aligned} \tag{2.3}$$

From the model,the covariance matrix $Cov(x) = \lambda_x \phi \lambda_x^T + \theta_\delta = \Sigma_{qxq}$ and let S_{qxq} be the sample covariance matrix. The aim of estimation is to match $S_{qxq} \approx \Sigma_{qxq}$. For perfect fit $S = \Sigma$ and then substituting S in the place of Σ in the log likelihood we have

$$\begin{aligned}
lnL(S) &= -\frac{n}{2}(ln(S) + tr(SS^{-1})) \\
&= -\frac{n}{2}(ln(S) + tr(I)) \\
&= -\frac{n}{2}(ln(S) + q)
\end{aligned} \tag{2.4}$$

Where $SS^{-1} = I$ and $tr(I) = q$

Creating a function $F(\theta) \approx S - \Sigma(\theta)$ and subtracting equation (2.3) from (2.4)

$$
\begin{aligned}
F(\theta) &= lnL(S) - lnL(\Sigma) \\
&= -\frac{n}{2}[ln|S| + q] + \frac{n}{2}[ln|\Sigma| + tr|S\Sigma^{-1}|] \\
&= \frac{n}{2}[ln|\Sigma(\theta)| + tr(S\Sigma(\theta)^{-1}) - ln|S| - q
\end{aligned} \tag{2.5}
$$

Using Newton Rapson or Gauss Newton algorithm, equation (2.5) was minimized for numerical approximation (see Wiley 1973).

2.2 Structural Model

The structural model in structural equation modeling is

$$\eta = \alpha + \beta_0\eta + \Gamma_1\xi + \zeta \tag{2.6}$$

η is mx1 vector of latent endogenous variables and ξ is n x 1 vector of exogenous variables. β_0 is mxm matrix of regression coefficients for the latent endogenous variables with zeros on the diagonal and Γ_1 is mxn matrix of regression coefficients predicting η by ξ. α is mxn matrix of intercepts and ζ is a vector of disturbance. The structural model (2.6) is said to be in implicit form, implicit because it has endogenous variables on both sides of the equations, i.e. it is not solved for the endogenous variables. It is assumed that the diagonal of β_0 is zero so that no element of η is a function of itself. A sufficient condition for solving (2.6) is that$(I - B_0)$ is invertible, then (2.6) can be solved for the endogenous variables and written as

$$\eta = \beta_0^* + \Gamma_1^*\xi^* + \zeta^* \tag{2.7}$$

Where, $\beta_0^* = (I - \beta_0)^{-1}\alpha$, $\Gamma_1^* = (I - \beta_0)^{-1})\Gamma_1$, $\zeta^* = (I - \beta_0)^{-1})\zeta$

The structural model (2.7) is said to be in reduced form as the η now appear only on the left hand side of the equations. It is important to note the assumption that the equation errors ζ were additive and independent of the ξ in the implicit form (2.6) results in the equation errors ζ^* in the reduced form (2.7) also being additive and independent of the ξ.

Assumptions under structural model are

- $\xi \sim N(0, \phi)$
- $\zeta \sim N(0, \psi)$
- $E(\xi\zeta) = 0$
- $I - \beta_0$ is non singular

From the above assumptions covariance matrices associated with structural model, ϕ is a n x n covariance matrix of ξ which is $E(\xi\xi^T)$ and ψ is a m x m covariance matrix of ζ which is $E(\zeta\zeta^T)$. Moreover,the covariance structure in structural model are

$$\Sigma = \begin{bmatrix} \Sigma_{\eta\eta} & \Sigma_{\eta\xi} \\ \Sigma_{\xi\eta} & \Sigma_{\xi\xi} \end{bmatrix}$$

$$= \begin{bmatrix} E(\eta\eta^T) & E(\eta\xi^T) \\ E(\xi\eta^T) & E(\xi\xi^T) \end{bmatrix}$$

By centering the variables, it is found that $\eta = (I - \beta_0)^{-1}\Gamma_1\xi + (1 - \beta_0)^{-1}\zeta$ $\eta = \beta\Gamma_1\xi + \beta\zeta$ where β =$(I - \beta_0)^{-1}$

$$\begin{aligned}
E(\eta\eta^T) &= E((\beta\Gamma_1\xi + \beta\zeta)(\beta\Gamma_1\xi + \beta\zeta)^T) \\
&= E\{\beta\Gamma_1\xi\xi^T\Gamma_1^T\beta^T + \beta\Gamma_1\xi\zeta^T\beta^T + \beta\zeta\xi^T\Gamma_1^T\beta^T + \beta\zeta\zeta^T\beta^T\} \\
&= \beta\Gamma_1 E(\xi\xi^T)\Gamma_1^T\beta^T) + \beta\Gamma_1 E(\xi\zeta^T)\beta^T + \beta E(\zeta\xi^T)\Gamma_1^T\beta^T + \beta E(\zeta\zeta^T)\beta^T \\
&= \beta\Gamma_1\phi\Gamma_1^T\beta^T + 0 + 0 + \beta\psi\beta^T \\
&= \beta(\Gamma_1\phi\Gamma_1^T + \psi)\beta^T \\
&= (I - \beta_0)^{-1}(\Gamma_1\phi\Gamma_1^T + \psi)(I - \beta_0)^{-1}
\end{aligned}$$

Similarly,

$$\begin{aligned} E(\eta\xi^T) &= E\{(\beta(\Gamma_1\xi\xi^T + \beta\zeta\xi^T\} \\ &= \beta(\Gamma_1 E(\xi\xi^T) + \beta E(\zeta\xi^T) = \beta\Gamma_1\phi \\ &= (I-\beta_0)^{-1}\Gamma_1\phi \end{aligned}$$

And $E(\xi\eta^T) = E\{\xi\xi^T\Gamma_1^T\beta^T + \xi\zeta^T\beta^T\} = \Gamma_1^T((I-\beta_0)^{-1})^T\phi$

$E(\xi\xi^T) = \phi$

The covariance matrix,$\Sigma(\theta)$ produced by model is

$$\Sigma(\theta) = \begin{bmatrix} \Sigma_{xx^T}(\theta) & \Sigma_{yx^T}(\theta) \\ \Sigma_{xy^T}(\theta) & \Sigma_{yy^T}(\theta) \end{bmatrix}$$

θ consists of all parameters in measurement and structural model. By centering the variables,

$E(yy^T) = E\{(\lambda_y\eta + \delta)(\lambda_y\eta + \delta)^T\}$

$E(yy^T) = \lambda_y E(\eta\eta^T)\lambda_y^T + E(\delta\delta^T)$ by substituting $E(\eta\eta^T)$

$E(yy^T) = \lambda_y\{(I-\beta_0)^{-1})(\Gamma_1\phi\Gamma_1^T + \psi)(I-\beta_0)^{-1})^T\}\lambda_y^T + \theta_\delta$

$E(xx^T) = E\{(\lambda_x\xi + \delta)(\lambda_x\xi + \delta)^T\}$

$E(xx^T) = \lambda_x E(\xi\xi^T)\lambda_x^T + E(\delta\delta^T)$ other expectations are zero

$E(xx^T) = \lambda_x\phi\lambda_x^T + \theta_\delta$

$E(yx^T) = E\{(\lambda_y\eta + \varepsilon)(\lambda_x\xi + \delta)^T\}$ $E(yx^T) = \lambda_y E(\eta\xi^T)\lambda_x^T$, substituting$E(\eta\xi^T)$

$E(yx^T) = \lambda_y(I-\beta_0)^{-1})\Gamma_1\phi\lambda_x^T$

$E(xy^T) = E\{(\lambda_x\xi + \delta)(\lambda_y\eta + \varepsilon)^T\}$

$E(xy^T) = \lambda_x E(\xi\eta^T\lambda_y^T) = \lambda_x\Gamma_1^T(I-\beta_0)^{-1})^T\phi\lambda_y$

The objective of the structural equation is to model the relationship between observed and latent variables in assuming that observed covariance matrix is reproduced by the model covariance matrix defined above. The null hypothesis is the covariance matrix produced by the model is equal to the population covariance matrix, H0:$\Sigma = \Sigma(\theta)$. where Σ is the true population covariance matrix, and its unknown. S is the unbiased sample covariance matrix found using the observed data. The number of parameter to be estimated in the structural model are:

β: m x m parameter to be estimated

Γ_1: m x n parameter to be estimated

ϕ: n x n, and symmetric so that $\frac{n(n+1)}{2}$ parameter to be estimated

ψ: $\frac{m(m+1)}{2}$ parameter to be estimated

The total parameter to be estimated t= m x m + m x n + $\frac{m(m+1}{2}$ + $\frac{n(n+1}{2}$ and the available parameter is $\Sigma_{(m+n)(m+n)}$ in which the number of non-redundant parameter is $\frac{(m+n)(m+n+1)}{2}$. Hence the identification of the model is depending on the values of t as mentioned for measurement model. Similar to measurement model, the parameters in the structural model can be estimated by using maximum likelihood. That is:

$F(\theta) = ln|\Sigma| + tr(S\Sigma^{-1}) - ln|S| - tr(SS^{-1})$. The only diffrence is the order of Σ which is (m+n)(m+n) and the order of sample covariance S (m+n)+(m+n).

$F(\theta) = ln|\Sigma| + tr(S\Sigma^{-1}) - ln|S| - tr(I)$, where $I = SS^{-1}$

$F(\theta) = ln|\Sigma| + tr(S\Sigma^{-1}) - ln|S| - (m+n)$. Hence Newton Rapson or Gaus Newton algorithm can be used to minimize $F(\theta)$ (Pauk and Maiti ,2008)

2.3 Two-way Latent Interaction and Estimation methods

Structural equation modeling allows researchers to model not only observed interaction effects but allows for interactions involving latent variables which have become increasingly popular in social science research. Since the seminal work of Kenny and Judd (1984), considerable effort has been devoted to the estimation of these cross-product effects. More discussion is presented under this chapter with it's estimation methods. In the first two section product indicator approach and approaches utilizing the likelihood for estimating two-way latent interactions effect is presented.

2.3.1 Product indicator approach

Constraint Approach

Kenny and Judd (1984) first formulated a way to estimate and test latent interaction effects in SEM as expressed below

$$y = \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta \tag{2.8}$$

where y is a mean centered observed dependent variable. Where γ_1,γ_2, and γ_3 are regression coefficients, ξ_1 and ξ_2 are exogenous latent variables, $\xi_1\xi_2$ is the latent interaction term between ξ_1 and ξ_2 and ζ is the residual. In this model the intercept term is not included assuming that the measured variable were centered and the model contains two indicators for the latent variables ξ_1 and ξ_2. This indicators were expressed as follows

$$x_1 = \lambda_1\xi_1 + \delta_1$$

$$x_2 = \lambda_2\xi_1 + \delta_2$$

$$x_3 = \lambda_3\xi_2 + \delta_3$$

$$x_4 = \lambda_4\xi_2 + \delta_4$$

Which can be written as

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ \lambda_2 & 0 \\ 0 & \lambda_3 \\ 0 & \lambda_4 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{pmatrix}$$

If we put the equations for all of the indicators of the latent exogenous variables together, we have

$$x = \lambda_x \xi + \delta \tag{2.9}$$

When first introducing the fully-latent approach to estimating latent variable interactions, Kenny and Judd limited their model to effects on a measured variable y. Jaccard ,Wan(1995) and Hayduk, 1987 introduced a latent variable η instead of observed variable y as the dependent variable in their model. Then equation (2.8) became

$$\eta = \gamma_1 \xi_1 + \gamma_2 \xi_2 + \gamma_3 \xi_1 \xi_2 + \zeta \tag{2.10}$$

For the latent endogenous variable model, the measurement portion of the model would be expanded to include

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \lambda_{y1} \\ \lambda_{y2} \end{bmatrix} \begin{bmatrix} \eta \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

The Kenny-Judd approach imposed several types of constraints upon the model. The main characteristic of the constrained approach(es) examination of an interaction between two latent variables is the specification of nonlinear constraints which determine that the parameters of the measurement model of the product variable (i.e., loadings and error (co)variances) are not freely estimated but expressed in terms of the parameters of the measurement models of the first-order effect variables. First, the loadings of the indicators on the interaction term were constrained to be equal to the product of the loadings for the two indicators that created the interaction indicator. According to their approach, all four cross products of the observed variables were used as indicators of the latent interaction. That is

$$\begin{aligned} x_1x_3 &= \lambda_1\lambda_3\xi_1\xi_2 + \lambda_1\xi_1\delta_3 + \lambda_3\xi_2\delta_1 + \delta_1\delta_3 \\ x_1x_4 &= \lambda_1\lambda_4\xi_1\xi_2 + \lambda_1\xi_1\delta_4 + \lambda_4\xi_2\delta_1 + \delta_1\delta_4 \\ x_2x_3 &= \lambda_2\lambda_3\xi_1\xi_2 + \lambda_2\xi_1\delta_3 + \lambda_3\xi_2\delta_2 + \delta_2\delta_3 \\ x_2x_4 &= \lambda_2\lambda_4\xi_1\xi_2 + \lambda_2\xi_1\delta_4 + \lambda_4\xi_2\delta_2 + \delta_2\delta_4 \end{aligned} \tag{2.11}$$

are the indicators of the product latent variables $\xi_1\xi_2$.

There are a total of 23 variables in their model, 15 latent variables ($\xi_1, \xi_2, \xi_1\xi_2, \delta_1, \delta_2, \delta_3, \delta_4, \delta_1\delta_3,$ $\delta_1\delta_4, \delta_2\delta_3, \delta_2\delta_4, \lambda_1\xi_1\delta_3, \lambda_1\xi_1\delta_4, \lambda_3\xi_2\delta_1, \lambda_3\xi_2\delta_2$) and 8 observed variables ($x_1, x_2, x_3, x_4, x_1x_3, x_1x_4, x_2x_3, x_2x_4$).

The measurement portion for the above four product indicators (2.11) of the interaction can be shown as

$$\begin{bmatrix} x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix} = \begin{bmatrix} x_1x_3 \\ x_1x_4 \\ x_2x_3 \\ x_2x_4 \end{bmatrix} = \begin{bmatrix} \lambda_{x5} \\ \lambda_{x6} \\ \lambda_{x7} \\ \lambda_{x8} \end{bmatrix} \begin{bmatrix} \xi_1\xi_2 \end{bmatrix} + \begin{bmatrix} \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \end{bmatrix}$$

Where

$\delta_5 = \lambda_1\xi_1\delta_3 + \lambda_3\xi_2\delta_1 + \delta_1\delta_3$

$\delta_6 = \lambda_1\xi_1\delta_4 + \lambda_4\xi_2\delta_1 + \delta_1\delta_4$

$\delta_7 = \lambda_2\xi_1\delta_3 + \lambda_3\xi_2\delta_2 + \delta_2\delta_3$

$\delta_8 = \lambda_2\xi_1\delta_4 + \lambda_4\xi_2\delta_2 + \delta_2\delta_4$

$\lambda_{x5} = \lambda_{x1}\lambda_{x3}$

$\lambda_{x6} = \lambda_{x1}\lambda_{x4}$

$\lambda_{x7} = \lambda_{x2}\lambda_{x3}$

$\lambda_{x8} = \lambda_{x2}\lambda_{x4}$

In the product $x_1x_3 = \lambda_1\lambda_3\xi_1\xi_2 + \lambda_1\xi_1\delta_3 + \lambda_3\xi_2\delta_1 + \delta_1\delta_3$ which can be written as

$$x_1x_3 = \lambda_1\lambda_3\xi_1\xi_2 + \delta_5 \tag{2.12}$$

The loading of x_1x_3 on the latent interaction term $\xi_1\xi_2$ is equal to $\lambda_{x1}\lambda_{x3}$ which is the product of the loadings for x_1 and x_3. From equation (2.12), the variance decomposition of the indicator product x_1x_3 is give by $var(x_1x_3) = \lambda_{x5}^2 var(\xi_1\xi_2)+var(\delta_5)$, assuming that ξ_1, ξ_2,$\delta_1, \delta_2, \delta_3, \delta_4$ and ζ are all in mean deviation form, multivariate normal and mutually uncorrelated with the exception of ξ_1 and ξ_2.

Then one of the constraint imposed by kenny and juddy on their model is.

$\phi_{33} = var(\xi_1\xi_2) = var(\xi_1)va(\xi_2) + cov^2(\xi_1, \xi_2)$. This can be written as $\phi_{33} = \phi_{11}\phi_{22} + \phi_{21}^2$. Under the same assumptions , the errors of the indicators for the interaction latent variable were constrained as:

$\theta_{\delta 5} = var(\delta_5) = \lambda_{1x}^2 var(\xi_1)var(\delta_3) + \lambda_{x3}^2 var(\xi_2)var(\delta_1) + var(\delta_1)va(\delta_3)$

$\theta_{\delta 6} = var(\delta_6) = \lambda_{1x}^2 var(\xi_1)var(\delta_4) + \lambda_{x4}^2 var(\xi_2)var(\delta_1) + var(\delta_1)va(\delta_4)$

$\theta_{\delta 7} = var(\delta_7) = \lambda_{2x}^2 var(\xi_1)var(\delta_3) + \lambda_{x3}^2 var(\xi_2)var(\delta_2) + var(\delta_2)va(\delta_3)$

$\theta_{\delta 8} = var(\delta_8) = \lambda_{2x}^2 var(\xi_1)var(\delta_4) + \lambda_{x4}^2 var(\xi_2)var(\delta_2) + var(\delta_2)va(\delta_4)$

Therefore, the errors of each of the indicators for the interaction latent variable were equal to

$\theta_{\delta 5} = \lambda_{1x}^2\phi_{11}\theta_{\delta 3} + \lambda_{x3}^2\phi_{22}\theta_{\delta 1} + \theta_{\delta 1}\theta_{\delta 3}$

$\theta_{\delta 6} = \lambda_{1x}^2\phi_{11}\theta_{\delta 4} + \lambda_{x4}^2\phi_{22}\theta_{\delta 1} + \theta_{\delta 1}\theta_{\delta 4}$

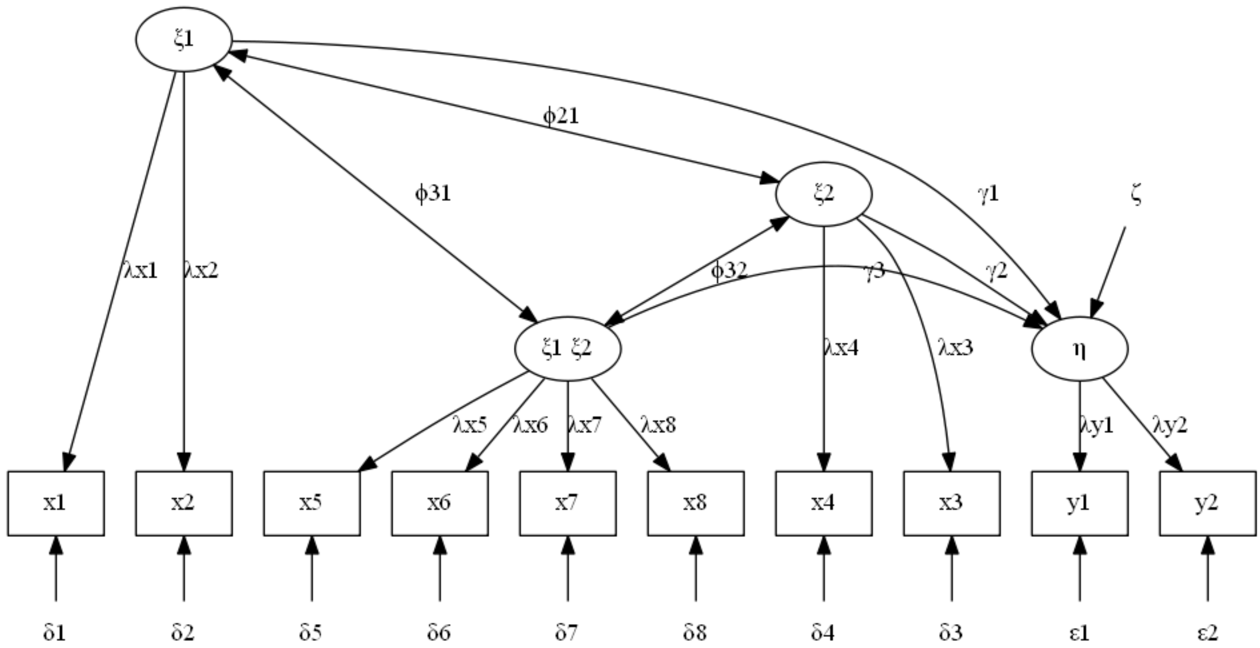
$\theta_{\delta 7} = \lambda_{2x}^2\phi_{11}\theta_{\delta 3} + \lambda_{x3}^2\phi_{22}\theta_{\delta 2} + \theta_{\delta 2}\theta_{\delta 3}$

$\theta_{\delta 8} = \lambda_{2x}^2\phi_{11}\theta_{\delta 4} + \lambda_{x4}^2\phi_{22}\theta_{\delta 2} + \theta_{\delta 2}\theta_{\delta 4}$

Based on the normality assumption, $cov(\xi_1\xi_2, \xi_1) = cov(\xi_1\xi_2, \xi_2) = 0$. That is $\phi_{31} = \phi_{32} = 0$. This indicates that the correlation between $\xi_1\xi_2$ and its components are fixed to be zero in the constrained approach, but freely to estimated in the unconstrained approach as discussed next section. It is noticed that, in the derivation of the variance and covariance of the latent product $\xi_1\xi_2$, kenny and Judd(1984) choose to mean center the observed variables in their model. However, Joreskog and Yang (1996) provided a general model thorough treatment of this constrained approach. They argued that even if the observed variables were mean centered, their products would not necessarily be mean centered. That is, the latent interaction variable will not be mean centered, and thus mean structure is necessary, and the intercept will not necessarily be zero. This implies that mean structure must always be used when specifying the latent interaction model. Accordingly, under the assumption that $\xi_1, \xi_2, \delta_1, \delta_2, \delta_3, \delta_4$ and ζ are in mean-deviation form, multivariate normal, and uncorrelated (except ξ_1 and ξ_2 are allowed to covary), Joreskog and Yang (1996) noted that the mean of the interaction term would be equal to the covariance between ξ_1 and ξ_2 and thus another constraint imposed upon the model was

$$\kappa_3 = \phi_{21}$$

Where κ_3 represents the mean of latent product variable, and ϕ_{21} represents the covariance between the first-order latent exogenous latent variables. Consequently, the Kenny-Judd model (without mean structure) is only appropriate when the covariance between ξ_1 and ξ_2 is approximately zero. Figure 1 contains a graphical depiction corresponding to the model



For observed variables that are not mean centered, Joreskog and Yang (1996) considered kenny-Judd model with intercepts in the structural and measurement models. That is, Joreskog and Yang (1996) and Jonsson (1998) model does not require the observed variables to be mean centered and includes their intercepts in a vector α. Hence equation (2.8) became

$$y = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta \tag{2.13}$$

where α is the constant term. They pointed out that, the value of α impacts the value of the coefficient γ_3 and the disturbance ζ sot that it should not necessary omitted from the structural model. Kenny and Judd (1984) proposed their method for testing for interaction effects using structural equation modeling and the concept of mean centering was carried over from the measured variable. However, Joreskog and Yang (1996), suggested that mean centering was not necessary. Moreover,(Jonsson, 1998)indicated that, even if y, ξ_1,ξ_2 and ζ are all centered, α will not be zero and the means of observed indicators are functions of other parameters and need to be jointly estimated with those parameters. Thus the measurement portion for the observed indicators of the first-order latent variables of the Joreskog and Yang (1996) model that corresponds to equation (2.13) is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \tau_{x1} \\ \tau_{x2} \\ \tau_{x3} \\ \tau_{x4} \end{bmatrix} + \begin{bmatrix} \lambda_{x1} & 0 \\ \lambda_{x2} & 0 \\ 0 & \lambda_{x3} \\ 0 & \lambda_{x4} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} \tag{2.14}$$

Where τ_{x1} through τ_{x4} represent the means of the observed x variables
Under the assumption that ξ_1, ξ_2, ζ, and all δ terms are multivariate normal with mean zero, and each is uncorrelated with the others (except ξ_1 and ξ_2 allowed to be correlated), they proposed a model with a mean structure. Hence, the mean vector and covariance matrix of ξ_1, ξ_2, and $\xi_1\xi_2$ are

$$\kappa = \begin{bmatrix} 0 \\ 0 \\ \phi_{21} \end{bmatrix}, \Phi = \begin{bmatrix} \phi_{11} & & \\ \phi_{21} & \phi_{22} & \\ 0 & 0 & \phi_{11}\phi_{22} + \phi_{21}^2 \end{bmatrix}$$

Joreskog and Yang (1996) showed three consequences of using product indicators. First, the distribution of y is not normal even if ξ_1 and ξ_2 are, indicating that the joint distribution of $(y, x_1, ..., x_q)'$ is not multivariate normal even if the joint distribution of $(x_1, ..., x_q)$ $(x_1, ..., x_q)'$ is. Second, the mean of y is a function of the y-intercepts, the coefficient of the latent interaction term, and the covariance of the main effects; therefore, a mean vector needs to be estimated along with the covariance matrix. That is,

$E(y) = E(\alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta)$

$E(y) = \alpha + \gamma_1 E(\xi_1) + \gamma_2 E(\xi_2) + \gamma_3 E(\xi_1\xi_2) + E(\zeta))$

$E(y) = \alpha + \gamma_3 E(\xi_1\xi_2)$ but $Cov(\xi_1, \xi_2) = E(\xi_1\xi_2) - E(\xi_1)E(\xi_2 = \phi_{21}$

Hence $E(y) = \alpha + \gamma_3\phi_{21}$

And lastly, if a researcher decides to use the Kenny and Judd method with all product indicators some nonlinear constraints are placed upon the mean vector and covariance matrix (Joreskog and Yang, 1996). In the Joreskog and Yang model all cross-products do not have to be used to estimate the interaction, and in order to identify the model only one is needed. Then the mean vector for equation (2.14) is

$E(x1) = \tau_{x1}$, $E(x2) = \tau_{x2}$,$E(x3) = \tau_{x3}$,$E(x4) = \tau_{x4}$, and $E(y) = \alpha + \gamma_3\phi_{21}$ which can be written as

$$\kappa' = (\alpha + \gamma_3\phi_{21}, \tau_{x1}, \tau_{x2}, \tau_{x3}, \tau_{x4})$$

Under the Kenny and Judd normality assumptions, the variance of the product-indicator x_1x_3 for example is given by

$$var(x_1x_3) = var[(\lambda_{x1}\xi_1 + \delta_1)(\lambda_{x3}\xi_2 + \delta_3)]$$

$var(x_1x_3) = \lambda_{x1}^2\lambda_{x3}^2 var(\xi_1\xi_2) + \lambda_{x1}^2 var(\xi_1)var(\delta_3)$

But $var(\xi_1\xi_2) = var(\xi_1)var(\xi_2) + Cov(\xi_1, \xi_2)^2$

$$var(\xi_1\xi_2) = \phi_{11}\phi_{22} + (\phi_{21})^2 \text{ (Kendall and Stuart, 1958)}$$

Hence the $var(x_1x_3) = \lambda_{x1}^2\lambda_{x3}^2(\phi_{21}^2 + \phi_{11}\phi_{22}) + \lambda_{x1}^2\phi_{11}\theta_{\delta 3}$

The covariance for the observed variables y, x_1, x_2, x_3, x_4 in the Kenny and Judd is

$\delta_{yy} = var(y) = var(\gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta)$

$\delta_{yy} = \gamma_1^2 var(\xi_1) + \gamma_2^2 var(\xi_2) + 2\gamma_1\gamma_2 cov(\xi_2, \xi_1) + \gamma_3^2 var(\xi_1\xi_2) + var(\zeta)$

$\delta_{yy} = \gamma_1^2\phi_{11} + \gamma_2^2\phi_{22} + 2\gamma_1\gamma_2\phi_{21} + \gamma_3^2(\phi_{11}\phi_{22} + (\phi_{21})^2) + \psi_{11}$ Since its assumed that the covariance of the exogenous variable with its interaction term is assumed to be zero.

$var(x_1) = var(\lambda_{x1}\xi_1 + \delta_1)$

$var(x_1) = \lambda_{x1}^2 var(\xi_1) + var(\delta_1) + 2cov(\lambda_{x1}\xi_1, \delta_1)$ but,$cov(\xi_1, \delta_1) = 0$

$var(x_1) = \lambda_{x1}^2\phi_{11} + \theta_{\delta 1}$

In the Kenny and Judd model,the first and third variable's factor loading are fixed to one for identification purpose. Hence $var(x_1) = \phi_{11} + \theta_{\delta 1}$

Similarly, $cov(y, x_1) = cov(\gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta, \lambda_{x1}\xi_1 + \delta_1)$ since the other covariances are zero

$cov(y, x_1) = \gamma_1\lambda_{x1} var(\xi_1) + \gamma_2\lambda_{x1} cov(\xi_2, \xi_1)$ $cov(y, x_1) = \gamma_1\phi_{11} + \gamma_2\phi_{21}$

Following the same procedure, the covariance matrix for the all the observed variables in their model is

$$\Sigma(\widehat{\theta})=\begin{pmatrix} \delta_{yy} \\ \gamma_1\phi_{11}+\gamma_2\phi_{21} & \phi_{11}+\theta_{\delta 1} \\ \gamma_1\lambda_{x2}\phi_{11}+\gamma_2\lambda_{x2}\phi_{21} & \lambda_{x2}\phi_{11} & \lambda_{x2}^2\phi_{11}+\theta_{\delta 2} \\ \gamma_1\phi_{21}+\gamma_2\phi_{22} & \phi_{21} & \lambda_{x2}\phi_{21} & \phi_{22}\theta_{\delta 3} \\ \gamma_1\lambda_{x4}\phi_{21}+\gamma_2\lambda_{x4}\phi_{22} & \lambda_{x4}\phi_{21} & \lambda\lambda_{x4}\lambda_{x2}\phi_{21} & \lambda_{x4}\phi_{22} & \lambda_{x4}^2\phi_{22}+\theta_{\delta 4} \end{pmatrix}$$

Similarly, in the Joreskog and Yang model of equation (2.11), the product term of x_1x_3 with $\lambda_{x1}=\lambda_{x3}=1$ can be expressed as

$$x_1x_3=(\tau_{x1}+\xi_1+\delta_1)(\tau_{x3}+\xi_2+\delta_3)$$

$$x_1x_3=\tau_{x1}\tau_{x3}+\tau_{x1}\xi_2+\tau_{x1}\delta_3+\tau_{x3}\xi_1+\xi_1\xi_2+\xi_1\delta_3+\delta_1\tau_{x3}+\delta_1\xi_2+\delta_1\delta_3$$

which can be simplified to

$$x_1x_3=\tau_{x1}\tau_{x3}+\tau_{x1}\xi_2+\tau_{x3}\xi_1+\xi_1\xi_2+\delta_{x1x3} \tag{2.15}$$

Where the error term for the product indicator is

$$\delta_{x1x3}=\tau_{x1}\delta_3+\xi_1\delta_3+\delta_1\tau_{x3}+\delta_1\xi_2+\delta_1\delta_3$$

The covariance of the product indicator with the other variables are

$$cov(y,x_1x_3)=cov(\gamma_1\xi_1+\gamma_2\xi_2+\gamma_3\xi_1\xi_2+\zeta,\tau_{x1}\xi_2+\tau_{x3}\xi_1+\xi_1\xi_2+\delta_{x1x3})$$

$$cov(y,x_1x_3)=cov(\gamma_1\xi_1,\tau_{x1}\xi_2)+cov(\gamma_1\xi_1,\tau_{x3}\xi_1)+cov(\gamma_2\xi_2,\tau_{x1}\xi_2)+cov(\gamma_2\xi_2,\tau_{x3}\xi_1)+var(\gamma_3\xi_1\xi_2)$$

since other covariances are assumed to be zero.

$$cov(y,x_1x_3)=\gamma_1\tau_{x1}cov(\xi_1,\xi_2)+\gamma_1\tau_{x3}cov(\xi_1,\xi_1)+\gamma_2\tau_{x1}cov(\xi_2,\xi_2)+\gamma_2\tau_{x3}cov(\xi_2,\xi_1)+\gamma_3^2var(\xi_1\xi_2)$$

$$cov(y,x_1x_3)=\gamma_1\tau_{x1}\phi_{21}+\gamma_1\tau_{x3}\phi_{11}+\gamma_2\tau_{x1}\phi_{22}+\gamma_2\tau_{x3}\phi_{21}+\gamma_3^2(\phi_{11}\phi_{22}+\phi_{21}^2) \tag{2.16}$$

Similarly,

$$cov(x_1, x_1x_3) = cov(\tau_{x1} + \xi_1 + \delta_1, \tau_{x1}\tau_{x3} + \tau_{x1}\xi_2 + \tau_{x3}\xi_1 + \xi_1\xi_2 + \delta_{x1x3})$$

$$cov(x_1, x_1x_3) = \tau_{x1}cov(\xi_1, \xi_2) + \tau_{x3}cov(\xi_1, \xi_1) + cov(\delta_1, \delta_{x1x3})$$

$$cov(x_1, x_1x_3) = \tau_{x1}\phi_{21} + \tau_{x3}\phi_{11} + cov(\delta_1, \tau_{x1}\delta_3 + \xi_1\delta_3 + \delta_1\tau_{x3} + \delta_1\xi_2 + \delta_1\delta_3)$$

$$cov(x_1, x_1x_3) = \tau_{x1}\phi_{21} + \tau_{x3}\phi_{11} + cov(\delta_1, \tau_{x3}\delta_1)$$

$$cov(x_1, x_1x_3) = \tau_{x1}\phi_{21} + \tau_{x3}\phi_{11} + \tau_{x3}\theta_{\delta 1}$$

$$cov(x_2, x_1x_3) = \lambda_{2x}\tau_{x1}\phi_{21} + \lambda_{2x}\tau_{x3}\phi_{11}$$

$$cov(x_3, x_1x_3) = \tau_{x1}\phi_{22} + \tau_{x3}\phi_{21}$$

$$cov(x_4, x_1x_3) = \lambda_{4x}\tau_{x1}\phi_{21} + \lambda_{4x}\tau_{x3}\phi_{11}$$

The variance of the product indicator is also

$$var(x_1x_3) = var(\tau_{x1}\tau_{x3} + \tau_{x1}\xi_2 + \tau_{x3}\xi_1 + \xi_1\xi_2 + \delta_{x1x3})$$

$$var(x_1x_3) = \tau_{x1}^2 var(\xi_2) + \tau_{x3}^2 var(\xi_1) + var(\xi_1\xi_2) + var(\delta_{x1x3})$$

$$var(x_1x_3) = \tau_{x1}^2\phi_{22} + \tau_{x3}^2\phi_{11} + \phi_{12}^2 + \phi_{11}\phi_{22} + var(\tau_{x1}\delta_3 + \xi_1\delta_3 + \delta_1\tau_{x3} + \delta_1\xi_2 + \delta_1\delta_3)$$

$$var(x_1x_3) = \tau_{x1}^2\phi_{22} + \tau_{x3}^2\phi_{11} + \phi_{12}^2 + \phi_{11}\phi_{22} + \tau_{x1}^2\theta_{\delta 3} + \phi_{11}\theta_{\delta 3} + \tau_{x3}^2\theta_{\delta 1} + \phi_{22}\theta_{\delta 1} + \theta_{\delta 1}\theta_{\delta 3}$$

And the mean of the product indicator is

$$E(x_1x_3) = E(\tau_{x1}\tau_{x3} + \tau_{x1}\xi_2 + \tau_{x3}\xi_1 + \xi_1\xi_2 + \delta_{x1x3})$$

$E(x_1x_3) = \tau_{x1}\tau_{x3} + E(\xi_1\xi_2) = \tau_{x1}\tau_{x3} + \phi_{21}$, the others expectation is assumed to be zero. The same is true for $E(\delta_{x1x3}) = E(\tau_{x1}\delta_3 + \xi_1\delta_3 + \delta_1\tau_{x3} + \delta_1\xi_2 + \delta_1\delta_3)$. It is assumed that $cov(\delta_1, \delta_3) = E(\delta_1\delta_3) - E(\delta_1)E(\delta_3) = 0$. Which indicates $E(\delta_1\delta_3) = 0$

Accordngly, the mean vector $\kappa' = (\alpha + \gamma_3\phi_{21}, \tau_{x1}, \tau_{x2}, \tau_{x3}, \tau_{x4}, \tau_{x1}\tau_{x3} + \phi_{21})$

This indicates that, adding one product indicator to the model increases the number of elements that need to be estimated including a factor loading, an error variance, and a mean. All of these added elements can be estimated as functions of already existing parameters; therefore the number of parameters is the same.

Jroskong and Yang (1996) model in equation (2.13) and (2.14), can be written for all indicators and product-indicators as

$$
\begin{bmatrix} y \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_1x_3 \\ x_1x_4 \\ x_2x_3 \\ x_2x_4 \end{bmatrix} = \begin{bmatrix} \alpha \\ \tau_{1x} \\ \tau_{2x} \\ \tau_{3x} \\ \tau_{4x} \\ \tau_{1x}\tau_{3x} \\ \tau_{1x}\tau_{4x} \\ \tau_{2x}\tau_{3x} \\ \tau_{2x}\tau_{4x} \end{bmatrix} + \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 \\ 1 & 0 & 0 \\ \lambda_{2x} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda_{4x} & 0 \\ \tau_{3x} & \tau_{1x} & 1 \\ \tau_{4x} & \tau_{1x}\lambda_{4x} & \lambda_{4x} \\ \tau_{3x}\lambda_{2x} & \tau_{2x} & \lambda_{2x} \\ \tau_{4x}\lambda_{4x} & \tau_{2x}\lambda_{4x} & \lambda_{2x}\lambda_{4x} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_1\xi_2 \end{bmatrix} + \begin{bmatrix} \zeta \\ \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \end{bmatrix}
$$

where

$\delta_5 = \delta_{x1x3} = \tau_{x1}\delta_3 + \xi_1\delta_3 + \delta_1\tau_{x3} + \delta_1\xi_2 + \delta_1\delta_3$ as defined in equation (2.15)

similarly

$\delta_6 = \delta_{x1x4} = \tau_{x1}\delta_4 + \xi_1\delta_4 + \delta_1\tau_{x4} + \lambda_4\delta_1\xi_2 + \delta_1\delta_4$

$\delta_7 = \delta_{x2x3} = \tau_{x2}\delta_3 + \lambda_{2x}\xi_1\delta_3 + \delta_2\tau_{x3} + \delta_2\xi_2 + \delta_2\delta_3$

$\delta_8 = \delta_{x2x4} = \tau_{x1}\delta_4 + \xi_1\delta_4 + \delta_1\tau_{x4} + \lambda_4\delta_1\xi_2 + \delta_1\delta_4$

Because the product-indicators are functions of the indicators of the observed variables that were used to create them, it is reasonable to assume the errors of the product indicators to correlate with the errors for the observed variables that created them. Accordingly, the covariance matrix for ζ and δ_1 to δ_8 is

$$\begin{bmatrix}
\psi \\
0 & \theta_{\delta 1} \\
0 & 0 & \theta_{\delta 2} \\
0 & 0 & 0 & \theta_{\delta 3} \\
0 & 0 & 0 & 0 & \theta_{\delta 4} \\
0 & \tau_{3x}\theta_{\delta 1} & 0 & \tau_1\theta_{\delta 3} & 0 & \theta_{\delta 5} \\
0 & \tau_{4x}\theta_{\delta 1} & 0 & 0 & \tau_{1x}\theta_{\delta 4} & \theta_{\delta 65} & \theta_{\delta 6} \\
0 & 0 & \tau_{3x}\theta_{\delta 2} & \tau_{2x}\theta_{\delta 3} & 0 & \theta_{\delta 75} & 0 & \theta_{\delta 7} \\
0 & 0 & \tau_{4x}\theta_{\delta 2} & 0 & \tau_{2x}\theta_{\delta 4} & 0 & \theta_{\delta 86} & \theta_{\delta 87} & \theta_{\delta 8}
\end{bmatrix}$$

where

$\psi = var(\zeta)$

$\tau_{3x}\theta_{\delta 1} = cov(\delta_5, \delta_1)$

$\theta_{\delta 5} = var(\delta_5) = var(\tau_{x1}\delta_3 + \xi_1\delta_3 + \delta_1\tau_{x3} + \delta_1\xi_2 + \delta_1\delta_3)$

$\theta_{\delta 5} = \tau_{x1}^2 var(\delta_3) + var(\xi_1)var(\delta_3) + \tau_{x3}^2 var(\delta_1) + var(\delta_1)var(\xi_2) + var(\delta_1)var(\delta_3)$

$\theta_{\delta 5} = \tau_{x1}^2\theta_{\delta 3} + \phi_{11}\theta_{\delta 3} + \tau_{x3}^2\theta_{\delta 1} + \phi_{22}\theta_{\delta 1} + \theta_{\delta 1}\theta_{\delta 3}$

And

$\theta_{\delta 6} = var(\delta_6) = \tau_{1x}^2\theta_{\delta 4} + \tau_{4x}^2\theta_{\delta 1} + \phi_{11}\theta_{\delta 4} + \lambda_{4x}^2\phi_{22}\theta_{\delta 1} + \theta_{\delta 1}\theta_{\delta 4}$

$\theta_{\delta 7} = var(\delta_7) = \tau_{2x}^2\theta_{\delta 3} + \tau_{3x}^2\theta_{\delta 2} + \phi_{22}\theta_{\delta 2} + \lambda_{2x}^2\phi_{11}\theta_{\delta 3} + \theta_{\delta 2}\theta_{\delta 3}$

$\theta_{\delta 8} = var(\delta_8) = \tau_{2x}^2\theta_{\delta 4} + \tau_{4x}^2\theta_{\delta 2} + \lambda_{2x}^2\phi_{11}\theta_{\delta 4} + \lambda_{4x}^2\phi_{22}\theta_{\delta 2} + \theta_{\delta 2}\theta_{\delta 4}$

Algina and Moulder (2001) extended the Joreskog and Yang model by mean-centering the independently observed variables for exogenous ξ_1 and ξ_2 and mean structure for the latent variables is as in the Joreskog-Yang model. They found that this model was more likely to converge, was less biased and more power than the Joreskog and Yang (1996) uncentered model. Mean centering simplifies both the measurement equations for observed indicators of the exogenous latent variables and the correlations of the measurement errors for the exogenous variables. For the sake of simplicity , let each of the latent variable η, ξ_1 and ξ_2 is associated with three indicators. That is y_1, y_2, y_3 to η, x_1, x_2, x_3 to ξ_1 and x_4, x_5, x_6 to ξ_2.

Then the LISREL specification for the measurement portion of the mean-centered using 3 matched -pairs for the constrained approach of Algina and Moulder (2001) can be written as:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \tau_{y1} \\ \tau_{y2} \\ \tau_{y3} \end{bmatrix} + \begin{bmatrix} 1 \\ \lambda_{y2} \\ \lambda_{y3} \end{bmatrix} \eta \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix},$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_1x_4 \\ x_2x_5 \\ x_3x_6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \lambda_2 & 0 & 0 \\ \lambda_3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \lambda_5 & 0 \\ 0 & \lambda_6 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & \lambda_8 \\ 0 & 0 & \lambda_9 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_1\xi_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \end{bmatrix},$$

And

$$\eta = \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta,$$

$$\kappa = \begin{bmatrix} 0 \\ 0 \\ \kappa_3 \end{bmatrix}, \Phi = \begin{bmatrix} \phi_{11} & & \\ \phi_{21} & \phi_{22} & \\ 0 & 0 & \phi_{33} \end{bmatrix}$$

$$\Theta_\varepsilon = \begin{bmatrix} \theta_{\varepsilon 1} & 0 & 0 \\ 0 & \theta_{\varepsilon 2} & 0 \\ 0 & 0 & \theta_{\varepsilon 3} \end{bmatrix}$$

$$\Theta_\delta = \begin{bmatrix} \theta_{\delta 1} & & & & & & & & \\ 0 & \theta_{\delta 2} & & & & & & & \\ 0 & 0 & \theta_{\delta 3} & & & & & & \\ 0 & 0 & 0 & \theta_{\delta 4} & & & & & \\ 0 & 0 & 0 & 0 & \theta_{\delta 5} & & & & \\ 0 & 0 & 0 & 0 & \theta_{\delta 6} & & & & \\ \theta_{\delta 71} & 0 & 0 & \theta_{\delta 74} & 0 & 0 & \theta_{\delta 7} & & \\ 0 & \theta_{\delta 82} & 0 & 0 & \theta_{\delta 85} & 0 & 0 & \theta_{\delta 8} & \\ 0 & 0 & \theta_{\delta 93} & 0 & 0 & \theta_{\delta 96} & 0 & 0 & \theta_{\delta 9} \end{bmatrix}$$

We note that when the measurement model for the y variables include the intercept terms (τ_y terms), the structural equation does not need an intercept term. Based on the product-indicators:

$x_1x_4 = (\xi_1 + \delta_1)(\xi_2 + \delta_4) = \xi_1\xi_2 + \xi_1\delta_4 + \xi_2\delta_1 + \delta_1\delta_4$,

$x_2x_5 = (\lambda_2\xi_1 + \delta_2)(\lambda_5\xi_2 + \delta_5) = \lambda_2\lambda_5\xi_1\xi_2 + \lambda_2\xi_1\delta_5 + \lambda_5\xi_2\delta_2 + \delta_2\delta_5$

and $x_3x_6 = (\lambda_3\xi_1 + \delta_3)(\lambda_6\xi_2 + \delta_6) = \lambda_3\lambda_6\xi_1\xi_2 + \lambda_3\xi_1\delta_6 + \lambda_6\xi_2\delta_3 + \delta_3\delta_6$

The necessary constraints were

Constraint 1 - the loadings for each of the interaction effects indicators is constrained to equal the product of their associated loadings on ξ_1 and ξ_2 $\lambda_8 = \lambda_2\lambda_5$

$\lambda_9 = \lambda_3\lambda_6$

Constraint 2 - the variance of the interaction latent variable is equal to the product of the variances of ξ_1 and ξ_2 plus the squared covariance between ξ_1 and ξ_2 [imposed in conjunction with the normality constraint]

$\phi_{33} = \phi_{11}\phi_{22} + \phi_{21}^2$;

Constraint 3 - constrains the errors of the each of the indicators for the interaction latent variable

Mean centering

$\theta_{\delta 7} = \phi_{11}\theta_{\delta 4} + \phi_{22}\theta_{\delta 1} + \theta_{\delta 1}\theta_{\delta 4}$

$\theta_{\delta 8} = \lambda_2^2 \phi_{11} \theta_{\delta 5} + \lambda_5^2 \phi_{22} \theta_{\delta 2} + \theta_{\delta 2} \theta_{\delta 5}$

$\theta_{\delta 9} = \lambda_3^2 \phi_{11} \theta_{\delta 6} + \lambda_6^2 \phi_{22} \theta_{\delta 3} + \theta_{\delta 3} \theta_{\delta 6}$

Normality Constraint âĂŞ The second constraint is based on the assumption that ξ_1 and ξ_2 are normally distributed. If this assumption holds true then the covariance of $\xi_1 \xi_2$ and each of the first-order terms (i.e.ξ_1 and ξ_2) is zero, (i.e., $\phi_{31} = 0$ and $\phi_{32} = 0$). Thus, the second constraint should also be imposed in conjunction with the normality constraint in which ϕ_{31} and ϕ_{32} are constrained to equal zero.

Constraint 4 - the mean of the interaction latent variable is constrained to equal the covariance between ξ_1 and ξ_2

$\kappa_3 = \phi_{21}$;

The constraints specified in the constrained model of Algina and Moulder (2001) are based on the assumption that ξ_1 and ξ_2 are normally distributed. Wall and Amemiya (2001) pointed out that when this assumption is not met then the covariance between ξ_1 and $\xi_1 \xi_2$, and the covariance between ξ_2 and $\xi_1 \xi_2$ are not necessarily zero ($\phi_{31} \neq 0$ and $\phi_{32} \neq 0$), and the constraint on the variance of $\xi_1 \xi_2$ does not necessarily hold true ($\phi_{33} = \phi_{11} \phi_{22} + \phi_{21}^2$). Based on this premise, Wall and Amemiya (2001) proposed a generalized appended product indicator (GAPI) approach in which the second constraint was relaxed. That is the GAPI procedure constructs the model covariance matrix with no assumption on the distributional form of any variables in the model. This model is also referred to as the partially constrained approach.

Unconstrained Approach

Marsh et al.(2004) questioned whether all the constraints in the constrained approach were necessary or even appropriate. In their approach the product of observed variables is used to form indicators of the latent interaction term, as in the constrained approach. However, they did not impose any complicated nonlinear constraints to define relations between product indicators and the latent interaction factor. The mean of the latent product variable (κ_3) was only constrained to be equal to the covariance of the two latent predictor variables ($cov(\xi_1, \xi_2)$). They demonstrated that the unconstrained model is identified when there are at least two product indicators of the latent interaction effect.

Similarly to the partially-constrained model, this model allows $\xi_1\xi_2$ to covary with ξ_1 and ξ_2,and does not require the stringent assumption that ξ_1 and ξ_2 are normally distributed. Marsh et al. (2006) noted that this unconstrained model was much easier for researchers to implement than the constrained model because it does not necessitate the specification of nonlinear constraints.

Methods for creating product indicators

In their model, Algina and Moulder used a similar approach to Kenny and Judd (1984) for forming indicators for the interaction term, by using all possible products of indicators for ξ_1 and ξ_2 to form the indicators for the interaction term. In three-indicator model this would yield nine indicators ($x_1x_4, x_1x_5, x_1x_6, x_2x_4, x_2x_5, x_2x_6, x_3x_4, x_3x_5, x_3x_6$) for the latent variable interaction. In their model, JÃűreskog and Yang (1996) used a single product (x_1x_4) to form an indicator for the interaction term. In another study, Yang (1998) used a matched-pairs(x_1x_4, x_2x_5, x_3x_6) approach in which each indicator of ξ_1 was paired with another indicator of ξ_2 . In the matched-pairs approach, each first-order indicator was used in only one product-indicator of the latent variable interaction. In the model with the three indicators for each of ξ_1 and ξ_2 , this would yield only three indicators for the latent interaction term, $\xi_1\xi_2$.

Marsh, Wen, and Hau (2004) conducted a simulation study to compare these three methods and found that the matched-pairs method yielded the most precise parameter estimates. Based on this finding, two recommendations were made. First, researchers should use all information that is available (i.e., all observed variables that are indicators of ξ_1 and ξ_2 should be used to form the interaction indicators). Second, information should not be reused (i.e., once an observed variable has been used to form an indicator of the interaction term, that indicator should not be used to form a second indicator of the same interaction term). This second recommendation was made to avoid inducing correlations between the error variances of the indicators for ξ_1 and ξ_2 , and $\xi_1\xi_2$.

Robustness to Violations of Normality Assumptions in Product Indicator Approaches

There are different problems associated with violating the assumption of multivariate normality that must be considered in the evaluation of the product indicator approaches. First, the maximum likelihood (ML) estimation typically used with each of these approaches is based on the assumption of the multivariate normality of all indicators. Second, even when the indicators of ξ_2 and ξ_2 are normally distributed, the distributions resulting from the products of these indicators ($\xi_1\xi_2$) are known to be non-normal (JÃűreskog and Yang, 1996). The distribution analytic approaches discussed next, model this non-normality explicitly and are not affected by it. However, both the constrained and the unconstrained product indicator approaches are negatively affected when ML estimation is used. Fortunately, ML estimation tends to be robust to violations of normality such as these in terms of parameter estimates (Hau and Marsh, 2004). However, previous research suggests that the likelihood ratio test is biased, and that ML standard errors are too small under some conditions of non-normality (Hu, Bentler, and Kano, 1992; West, Finch, and Curran, 1995). Although there are alternative estimators that do not assume multivariate normality (e.g., asymptotically distribution-free or weighted least squares estimators; Bollen, 1989), simulation studies of latent interaction effects suggest that ML estimation outperforms these alternative estimation procedures under most conditions (Jaccard and Wan, 1995; Wall and Amemiya, 2001; see also Marsh et al., 2004).

The third problem associated with violations of the assumptions of normality is specific to the constrained approach but does not apply to the unconstrained approach. As emphasized by Wall and Amemiya (2001) and Marsh and colleagues (2004), the nature of the constraints imposed in the constrained approach depends fundamentally on the assumption that ξ_2 and ξ_2 are normally distributed. Importantly, estimates of the interaction effect based on these constraints are not robust in relation to violations of this assumption of normality; neither the size nor even the direction of this bias is predictable a prior, and the size of the bias does not decrease systematically with increasing N. However, the unconstrained approach provides relatively unbiased estimates of the latent interaction effects under the widely varying conditions of non-normality considered by Marsh and colleagues (2004).

Indicator centering and mean structure in product indicator method

Suppose that the endogenous latent variable η has three indicators:y_1, y_2, y_3, and the exogenous latent variables ξ_1 and ξ_2 have two indicators, respectively: $x_1, x_2; x_3, x_4$. The indicators of the product term $\xi_1\xi_2$ are the matched pairs: x_1x_3, x_2x_4. When uncentered indicators are used, the measurement equations of the product indicators are very complicated involving intercept, product, and error terms. Thus , for example,

$$x_2x_4 = (\tau_2 + \lambda_2\xi_1 + \delta_2)(\tau_4 + \lambda_4\xi_2 + \delta_4) = \tau_2\tau_4 + \tau_4\lambda_2\xi_1 + \tau_2\lambda_4\xi_2 + \lambda_2\lambda_4\xi_1\xi_2 + \delta_{24}$$

where δ_{24}=$\tau_4\delta_2 + \lambda_4\xi_2\delta_2 + \tau_2\delta_4 + \lambda_2\xi_1\delta_4 + \delta_2\delta_4$

The loading of x_2x_4 on ξ_1 and ξ_2 are generally not zero.Importantly, a mean structure is always needed for the structural and measurement models. As pointed out by JÃűreskog and Yang (1996), even if ξ_1,ξ_2 and ζ are so as to have means of zero, $\kappa_3 = E(\xi_1\xi_2) = cov(\xi_1, \xi_2)$ will typically not be zero.

Consequently, as in the case in the analyses of interaction in multiple regression, centering the observed variables is routinely carried out to simplify the model (Aiken and West, 1991; Algina and Moulder,2001;marsh et al.,2004). Centering these indicators simplifies the model considerably. Denote x^C as the mean-centered variable of x, that is, $x^C = x - E(x)$. The measurement models become

$x_1^C = \xi_1 + \delta_1$, $x_2^C = \lambda_2\xi_1 + \delta_2$

$x_3^C = \xi_2 + \delta_3$, $x_4^C = \lambda_4\xi_2 + \delta_4$

The matched product indicators are $x_1^C x_3^C$, $x_2^C x_4^C$. After the indicators of the latent predictors have been centered, the intercept terms of the measurement equations of the original and the product indicators are no longer necessary. The intercept terms of the measurement equations of indicators of the latent outcome variable, however, are still necessary even if they have been centered (see Algina and Moulder, 2001; Marsh et al., 2004). Furthermore, even if ξ_1 and ξ_2 are centered so as to have means of zero, $E(\xi_1\xi_2) = cov(\xi_1, \xi_2)$ will typically not be zero. Hence, the latent mean of $E(\xi_1\xi_2)$ must be included in the model (see Algina and Moulder, 2001; Marsh et al., 2004). Therefore, although the latent interaction model is simpler after indicators are centered, the mean structure of the model is still necessary. More precisely,intercepts of indicators of the latent outcome variable and the mean of $E(\xi_1\xi_2)$ have to be included in the model.

2.3.2 Approaches utilizing the likelihood

The product-indicators approach discussed above can be viewed as artificially measured variables because they are not unique observed variables, instead they are created by the researcher and are thus ad hoc. Another problem in this approach is the distributional assumptions imposed by the models. For instance the assumptions that $\xi_1, \xi_2, \xi_3, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8, \delta_9$ and ζ are multivariate normal, and uncorrelated (except ξ_1,ξ_2 and ξ_3 are allowed to relate to one another). When data are non-normal, then the constraints imposed upon the variance of the interaction term (constraint 2) and it's covariance with the first-order terms (normality constraint) do not hold, and thus the constrained approach is not appropriate.

Similarly, its found that even if ξ_1,ξ_2 and ξ_3 are assumed to be normally distributed,the interaction is known to be non-normally distributed (Joreskog and Yang, 1996). The product-indicator models use maximum-likelihood estimation which is based on the assumption that all indicators in the model are multivariate normally distributed. It indicates that, the indicators for the interaction are known to be non-normally distributed and this assumption is violated when maximum-likelihood is used.

One of the solution to this violation of multivariate normality when estimating the interaction effect is the latent moderated structural equations (LMS) method proposed by Klein and Moosbrugger (2000). This approach does not require the creation of indicators for the interaction and recognizes the non-normal distribution of the interaction. The LMS method utilizes a mixture of multivariate normal distributions that are implied by the interaction model. The Expectation-Maximization (EM) algorithm (Dempster, Laird, and Rubin, 1977) is used to compute maximum-likelihood parameter estimates.

The general structural equation for an interaction model using the LMS approach can be written as

$$\eta = \alpha + \Gamma\xi + \xi'\Omega\xi + \zeta \tag{2.17}$$

where η is a endogenous latent variable, α is an intercept term, Γ is a (1 x k) vector of coefficients, ξ is a (k x 1) vector of latent exogenous variables, Ω is an upper triangular (k x k) matrix, and ζ is a disturbance term. The matrix Ω contain the non-linear effects and Γ contains linear effects in equation (2.17). In the case of the model with two exogenous latent variables and their interactions, the Ω matrix is an upper triangular matrix, and is specified as

$$\begin{bmatrix} 0 & \gamma_3 \\ 0 & 0 \end{bmatrix}$$

where γ_3 represents the interaction effect. The diagonal elements are zero unless we want to deal with quadratic effects. If one wanted to simultaneously estimate quadratic effects with the interaction effect, then the parameters on the diagonal could be freed. With two exogenous latent variables interaction and affect a single endogenous latent variable, the structural equation model in equation (2.28) can be written as

$$\eta = \alpha + \begin{bmatrix} \gamma_1 & \gamma_2 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \xi_1 & \xi_2 \end{bmatrix} \begin{bmatrix} 0 & \gamma_3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \zeta$$

The measurement portion of this model can be written as

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda_{y2} \\ \lambda_{y3} \end{bmatrix} \eta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_{2x} & 0 \\ \lambda_{x3} & 0 \\ 0 & 1 \\ 0 & \lambda_{5x} \\ 0 & \lambda_{6x} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{bmatrix}$$

The LMS method is based on the assumption that $\xi_1, \xi_2, \xi_3, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8, \delta_9, \varepsilon_1, \varepsilon_2$ and ε_3 are multivariately normally distributed. It also assumed that $\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8, \delta_9 \varepsilon_1, \varepsilon_2$ and ε_3 have expected values of zero and are uncorrelated with ξ_1,ξ_2 and ξ_3.

Finally, ζ has an expected value of zero and is assumed to be uncorrelated with$\xi_1, \xi_2, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \varepsilon$ and ε_3. Additionally, ζ is not assumed to be normal as that of product-indicator methods. In the case of two exogenous latent variables interact and affect a single endogenous latent variable, and there are three indicators for each of ξ_1, ξ_2 and ζ) has a nine-dimensional indicator vector $(x, y) = (x_1, x_2, x_3, x_4, x_5, x_6, y_1, y_2, y_3)$ can be represented as a finite mixture of multivariate normal distributions. The indicator x is assumed to be normally distributed, whereas indicator y is not assumed to be normally distributed because the product term $\xi_1\xi_2$ is in the structural equation. Thus linear and non-linear effects are separated and decomposed into independent random z variables using the Cholesky decomposition of the covariance matrix Φ.

z is made up of vectors $z1$ and $z2$ which represent the nonlinear and linear effects, respectively. From this, a continuous mixture of normal densities with $z1$ as the mixing vector can be derived. Then the partitioned mean vector and covariance matrix can be obtained. If an interaction exists, and thus γ_3 differs from zero, then the integral of the mixture cannot be solved analytically. But approximated by Hermite-Gaussian quadrature formulas of numerical integration, which are used to calculate mixture probabilities and mixture components (Klein and Moosbrugger, 2000). More discussion is presented under the next section.

Chapter 3

Methods

3.1 Model

3.1.1 Three way latent interaction

In this study we use a model with three latent variables with three observed indicators each for both endogenous and exogenous latent variables. For the identification purposes we chose to set a single factor loading to 1 for η, ξ_1, ξ_2 and ξ_3. By three-way interaction, we mean the interaction of three continuous exogenous latent variables (ξ_1, ξ_2, ξ_3). Following the LISREL specification, the structural equation considered was:

$$\eta = \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_3 + \gamma_4\xi_1\xi_2 + \gamma_5\xi_1\xi_3 + \gamma_6\xi_2\xi_3 + \gamma_7\xi_1\xi_2\xi_3 + \zeta \tag{3.1}$$

The measurement equations for each models were given by

$$
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \end{bmatrix}
=
\begin{bmatrix}
1 & 0 & 0 & 0 \\
\lambda_{2y} & 0 & 0 & 0 \\
\lambda_{3y} & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & \lambda_{x2} & 0 & 0 \\
0 & \lambda_{x3} & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & \lambda_{x5} & 0 \\
0 & 0 & \lambda_{x6} & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & \lambda_{x8} \\
0 & 0 & 0 & \lambda_{x9}
\end{bmatrix}
\begin{bmatrix} \eta \\ \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix}
+
\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \\ \delta_9 \end{bmatrix}
$$

The following assumptions were used in the current study.

- $\xi_1, \xi_2, \xi_3, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8, \delta_9, \varepsilon_1, \varepsilon_2,$and ε_3 are multivariate normally distributed.
- $\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8, \delta_9, \varepsilon_1, \varepsilon_2,$and ε_3 have expected values of zero and are uncorrelated with $\xi_1, \xi_2$and ξ_3.
- Finally, ζ has an expected value of zero and assumed to be uncorrelated with $\xi_1, \xi_2, \xi_3, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta$
  ε_3.

Based on these assumptions the mean vector and covariance matrix of $(\xi_1, \xi_2, \xi_3,$ $\xi_1\xi_2, \xi_1\xi_3, \xi_2\xi_3, \xi_1\xi_2\xi_3)$ were derived as follows:

$$cov(\xi_1, \xi_1\xi_2) = E(\xi_1)cov(\xi_1, \xi_2) + E(\xi_2)cov(\xi_1, \xi_1) + E[(\xi_1 - E(\xi_1))(\xi_1 - E(\xi_1))(\xi_2 - E(\xi_2))]$$

By centering ξ_1 and ξ_2, the expected value of both becomes ,$E(\xi_1) = 0$ and $E(\xi_2) = 0$.

Hence $cov(\xi_1, \xi_1\xi_2) = E[(\xi_1 - E(\xi_1))(\xi_1 - E(\xi_1))(\xi_2 - E(\xi_2))]$. Under multivariate normality all third moments vanish (see Bohnstedt and Goldberger 1969). This indicates that $cov(\xi_1, \xi_1\xi_2) = E(\xi_1\xi_1\xi_2) = 0$. Accordingly, all the covariance of the main effects with their two-way interaction is zero under normality condition and the given assumptions.

Following the same procedure, $cov(\xi_1\xi_2, \xi_1\xi_3) = E(\xi_1)E(\xi_1)cov(\xi_2, \xi_3) + E(\xi_1)E(\xi_3)cov(\xi_2, \xi_1) + E(\xi_2)E(\xi_1)cov(\xi_1, \xi_3) + E(\xi_2)E(\xi_3)cov(\xi_1, \xi_1) + cov(\xi_1, \xi_1)cov(\xi_2, \xi_3) + cov(\xi_1, \xi_3)cov(\xi_2, \xi_1)$. Centering ξ_1, ξ_2, ξ_3 the first four terms are zero and we have

$cov(\xi_1\xi_2, \xi_1\xi_3) = cov(\xi_1, \xi_1)cov(\xi_2, \xi_3) + cov(\xi_1, \xi_3)cov(\xi_2, \xi_1) = \phi_{11}\phi_{23} + \phi_{13}\phi_{12}$ and

$cov(\xi_1\xi_2, \xi_2\xi_3) = cov(\xi_1, \xi_2)cov(\xi_2, \xi_3) + cov(\xi_1, \xi_3)cov(\xi_2, \xi_2) = \phi_{12}\phi_{23} + \phi_{13}\phi_{22}$

The covariance of the main effects with the product of the three exogenous variables can be found in the similar manner

$$
\begin{aligned}
cov(\xi_1, \xi_1\xi_2\xi_3) &= cov(\xi_1, \xi_1)cov(\xi_2, \xi_3) + cov(\xi_1, \xi_2)cov(\xi_1, \xi_3) + cov(\xi_1, \xi_3)cov(\xi_1\xi_2) \\
&= \phi_{11}\phi_{23} + \phi_{12}\phi_{13} + \phi_{13}\phi_{12}
\end{aligned} \tag{3.2}
$$

$$
\begin{aligned}
cov(\xi_2, \xi_1\xi_2\xi_3) &= cov(\xi_2, \xi_1)cov(\xi_2, \xi_3) + cov(\xi_2, \xi_2)cov(\xi_1, \xi_3) + cov(\xi_2, \xi_3)cov(\xi_1, \xi_2) \\
&= \phi_{21}\phi_{23} + \phi_{22}\phi_{13} + \phi_{23}\phi_{12}
\end{aligned} \tag{3.3}
$$

$$cov(\xi_3, \xi_1\xi_2\xi_3) = cov(\xi_3, \xi_1)cov(\xi_2, \xi_3) + cov(\xi_3, \xi_2)cov(\xi_1, \xi_3) + cov(\xi_3, \xi_3)cov(\xi_1, \xi_2)$$
$$= \phi_{31}\phi_{23} + \phi_{32}\phi_{13} + \phi_{33}\phi_{12} \tag{3.4}$$

The covariance between two and three product is zero since the covariances involving five variables, for example $cov(\xi_1\xi_2, \xi_1\xi_2\xi_3) = 0$ under normality.

Following Bohnstedt and Goldberger (1969) and under normality, the variance of the latent product is

$$var(\xi_1\xi_2) = E^2(\xi_1)var(\xi_2) + E^2(\xi_2)var(\xi_1) + 2E(\xi_1)E(\xi_2)cov(\xi_1, \xi_2) + var(\xi_1)var(\xi_2)$$
$$+cov(\xi_1, \xi_2)^2$$

And under the given assumptions it reduces to

$$var(\xi_1\xi_2) = var(\xi_1)var(\xi_2) + cov(\xi_1, \xi_2)^2$$
$$= \phi_{11}\phi_{22} + \phi_{12}^2$$

Similarly,

$$var(\xi_1\xi_3) = var(\xi_1)var(\xi_3) + cov(\xi_1, \xi_3)^2$$
$$= \phi_{11}\phi_{33} + \phi_{13}^2$$

$$var(\xi_2\xi_3) = var(\xi_2)var(\xi_3) + cov(\xi_2, \xi_3)^2$$
$$= \phi_{22}\phi_{33} + \phi_{23}^2$$

For a normally distributed random variables $\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6$ with mean zero, the fourth and six moment is

$E(\xi_1\xi_2\xi_3\xi_4) = cov(\xi_1\xi_2, \xi_3\xi_4) + cov(\xi_1\xi_3, \xi_2\xi_4) + cov(\xi_1\xi_4, \xi_2\xi_3)$ and

$E(\xi_1\xi_2\xi_3\xi_4\xi_5\xi_6) = cov(\xi_1, \xi_2)E(\xi_3\xi_4\xi_5\xi_6) + cov(\xi_1, \xi_3)E(\xi_2\xi_4\xi_5\xi_6) + cov(\xi_1, \xi_4)E(\xi_2\xi_3\xi_4\xi_5) + cov(\xi_1, \xi_5)E(\xi_2\xi_3\xi_4\xi_6) + cov(\xi_1, \xi_6)E(\xi_2\xi_3\xi_4\xi_5)$(Kendall and Stuart, 1958).

Then we can find $var(\xi_1\xi_2\xi_3)$. That is ;

$$\begin{aligned} var(\xi_1\xi_2\xi_3) &= E(\xi_1^2\xi_2^2\xi_3^2) - E(\xi_1\xi_2\xi_3) \\ &= E(\xi_1^2\xi_2^2\xi_3^2) \\ &= var(\xi_1)E(\xi_2^2\xi_3^2) + cov(\xi_1, \xi_2)E(\xi_1\xi_2\xi_3^2) + cov(\xi_1, \xi_2)E(\xi_1\xi_2\xi_3^2) + cov(\xi_1, \xi_3)E(\xi_1\xi_2^2\xi_3) + cov(\xi_1 \\ &= var(\xi_1)E(\xi_2^2\xi_3^2) + 2cov(\xi_1, \xi_2)E(\xi_1\xi_2\xi_3^2) + 2cov(\xi_1, \xi_3)E(\xi_1\xi_2^2\xi_3) \end{aligned}$$

Using the fourth moment, it can be shown that;

$$\begin{aligned} E(\xi_2^2\xi_3^2) &= var(\xi_2)var(\xi_3) + 2cov((\xi_2, \xi_3) \\ &= \phi_{22}\phi_{33} + 2\phi_{23}\phi_{23} \end{aligned}$$

$$\begin{aligned} E(\xi_1\xi_2\xi_3^2) &= cov(\xi_1, \xi_2)var(\xi_3) + 2cov(\xi_1, \xi_3)cov(\xi_2, \xi_3) \\ &= \phi_{12}\phi_{33} + 2\phi_{13}\phi_{23} \end{aligned}$$

and

$$E(\xi_1\xi_2^2\xi_3) = 2\phi_{12}\phi_{23} + \phi_{13}\phi_{22}$$

Hence,

$$var(\xi_1\xi_2\xi_3) = \phi_{11}(\phi_{22}\phi_{33} + 2\phi_{23}\phi_{23}) + 2\phi_{12}(\phi_{12}\phi_{33} + 2\phi_{13}\phi_{23}) + 2\phi_{13}(2\phi_{12}\phi_{23} + \phi_{13}\phi_{22})$$

We can also find the mean vectors for $(\xi_1\xi_2, \xi_1\xi_3, \xi_2\xi_3, \xi_1\xi_2\xi_3)$

Centering ξ_1, ξ_2, ξ_3, the mean for the two products is

$$E(\xi_1\xi_2) = cov(\xi_1, \xi_2) = \phi_{12}$$
$$E(\xi_1\xi_3) = cov(\xi_1, \xi_3) = \phi_{13}$$
$$E(\xi_2\xi_3) = cov(\xi_2, \xi_3) = \phi_{23}$$
$$E(\xi_1\xi_2\xi_3) = 0$$

Then the mean and variance of the endogenous latent variable in equation (3.1) is

$$E(\eta) = \gamma_4\phi_{12} + \gamma_5\phi_{13} + \gamma_6\phi_{23}$$

and

$$\begin{aligned} var(\eta) = \gamma_1^2 var(\xi_1) + \gamma_2^2 var(\xi_2) + \gamma_3^2 var(\xi_3) + \gamma_4^2 var(\xi_1\xi_2) + \gamma_5^2 var(\xi_1\xi_3) + \gamma_6^2 var(\xi_2\xi_3) \\ + \gamma_7^2 var(\xi_1\xi_2\xi_3) + 2[cov(\gamma_1\xi_1, \gamma_2\xi_2) + cov(\gamma_1\xi_1, \gamma_3\xi_3) \\ + cov(\gamma_2\xi_2, \gamma_3\xi_3) + cov(\gamma_1\xi_1, \gamma_7\xi_1\xi_2\xi_3) \\ + cov(\gamma_2\xi_2, \gamma_7\xi_1\xi_2\xi_3) + cov(\gamma_3\xi_3, \gamma_7\xi_1\xi_2\xi_3)] + var(\zeta) \end{aligned}$$

That is

$$
\begin{aligned}
var(\eta) = \sigma_\eta^2 = {} & \gamma_1^2\phi_{11} + \gamma_2^2\phi_{22} + \gamma_3^2\phi_{33} + \gamma_4^2(\phi_{11}\phi_{22} + \phi_{12}^2) + \gamma_5^2(\phi_{11}\phi_{33} + \phi_{13}^2) \\
& + \gamma_6^2(\phi_{22}\phi_{33} + \phi_{23}^2) + \gamma_7^2[\phi_{11}(\phi_{22}\phi_{33} + 2\phi_{23}\phi_{23}) \\
& + 2\phi_{12}(\phi_{12}\phi_{33} + 2\phi_{13}\phi_{23}) + 2\phi_{13}(2\phi_{12}\phi_{23} \\
& + \phi_{13}\phi_{22})] + 2[\gamma_1\gamma_2\phi_{12} + \gamma_1\gamma_3\phi_{13} + \gamma_2\gamma_3\phi_{23} \\
& + \gamma_1\gamma_7(\phi_{11}\phi_{23} + \phi_{12}\phi_{13} + \phi_{13}\phi_{12}) \\
& + \gamma_2\gamma_7(\phi_{21}\phi_{23} + \phi_{22}\phi_{13} + \phi_{23}\phi_{12}) \\
& + \gamma_3\gamma_7(\phi_{31}\phi_{23} + \phi_{23}\phi_{13} + \phi_{33}\phi_{12})] \\
& + \psi
\end{aligned}
$$

(see Gerry Gray (1999))

Hence the variance covarince matrix for the latent variables in the model become:

$$
\Phi = \begin{pmatrix}
\phi 11 & & & & & & \\
\phi_{21} & \phi_{22} & & & & & \\
\phi_{31} & \phi_{32} & \phi_{33} & & & & \\
0 & 0 & 0 & \phi_{11}\phi_{22} + \phi_{12}^2 & & & \\
0 & 0 & 0 & \phi_{11}\phi_{23} + \phi_{13}\phi_{12} & \phi_{11}\phi_{33} + \phi_{13}^2 & & \\
0 & 0 & 0 & \phi_{12}\phi_{23} + \phi_{13}\phi_{22} & \phi_{12}\phi_{33} + \phi_{13}\phi_{23} & \phi_{22}\phi_{33} + \phi_{23}^2 & \\
\phi_{11}\phi_{23} + 2\phi_{12}\phi_{13} & \phi_{22}\phi_{13} + 2\phi_{23}\phi_{12} & \phi_{33}\phi_{12} + 2\phi_{32}\phi_{13} & 0 & 0 & 0 & \omega
\end{pmatrix}
$$

where ω stands for

$$
var(\xi_1\xi_2\xi_3) = \phi_{11}(\phi_{22}\phi_{33} + 2\phi_{23}\phi_{23}) + 2\phi_{12}(\phi_{12}\phi_{33} + 2\phi_{13}\phi_{23}) + 2\phi_{13}(2\phi_{12}\phi_{23} + \phi_{13}\phi_{22})
$$

3.2 Estimation Method

Let $f_i = (\eta, \xi_1, \xi_2, \xi_3)'$, $Z_i = (y_1, y_2, y_3, x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9)'$, $\epsilon_i = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta$
Then the full nonlinear structural equation model can be specified as follows

$$Z_i = \Lambda f_i + \epsilon_i \tag{3.5}$$

Following the notation in Wall(2009), let θ_m represent the measurement model parameters (i.e., parameters in Λ, Θ and θ_s denote the nonlinear structural parameters (i.e.,γ_1 to γ_7, Ψ). Where Θ is varance- covariance matrix for ϵ_i in equation (3.5). Note that $\theta = ((\theta m)', (\theta_s)')'$.

For individual i, the joint distribution of the observed data and the latent variables conditional on the parameter vector θ can be written under the nonlinear structural equation model in equation (3.1) and (3.5) as follows.

$$\begin{aligned} P(Z_i, f_i; \theta) &= P(Z_i|f_i, \theta_m)P(f_i, \theta_s) \\ &= P(Z_i|\eta_i, \xi_{1i}, \xi_{2i}, \xi_{3i}; \theta_m)P(\eta_i, \xi_{1i}, \xi_{2i}, \xi_{3i}; \theta_s) \\ &= P(Z_i|\eta_i, \xi_{1i}, \xi_{2i}, \xi_{3i}; \theta_m)P(\eta_i|\xi_{1i}, \xi_{2i}, \xi_{3i}; \theta_s)P(\xi_{1i}, \xi_{2i}, \xi_{3i}; \theta_\xi) \end{aligned} \tag{3.6}$$

Where θ_ξ is describing the distribution of ξ_i. However, the latent variables are not observable. Therefore, one must integrate the latent variables out of the joint distribution to obtain the marginal density of Z_i. That is:
$P(Z_i; \theta_m, \theta_s, \theta_\xi) = \int P(Z_i|\eta_i, \xi_{1i}, \xi_{2i}, \xi_{3i}; \theta_m)P(\eta_i|\xi_{1i}, \xi_{2i}, \xi_{3i}; \theta_s)P(\xi_{1i}, \xi_{2i}, \xi_{3i}; \theta_\xi)d\xi_i$

Hence, the likelihood function is

$$L(\Theta) = \prod \int P(Z_i|\eta_i, \xi_{1i}, \xi_{2i}, \xi_{3i}; \theta_m)P(\eta_i|\xi_{1i}, \xi_{2i}, \xi_{3i}; \theta_s)P(\xi_{1i}, \xi_{2i}, \xi_{3i}; \theta_\xi)d\xi_i \tag{3.7}$$

Rather than directly approximate the integral in equation (3.7) Klein and Moosbrugger (2000) proposed the latent moderated structural equation method, which does not require the creation of indicators for the interaction of latent variable. LMS uses numerical integration methods for approximating the integrals in Equation (3.7) and uses a finite mixture of normal distributions to approximate the nonnormal distribution.Then they develop an EM algorithm to find the MLEs of this distribution(see Klein and Moosbrugger, 2000)more in details.

3.3 Simulation Design

We designed the simulation study to achieve two goals. The first is to examine the performance of the estimation method in terms of parameter bias, root-mean-square error (RMSE), and standard error. The second is to test whether the regression coefficient for the three-way interaction term is statistically significant. There are twelve observed variables. Nine indicators,$x_1, \cdots, x_6$, for the three latent exogenous variables, ξ_1, ξ_2 and ξ_3. Three observed indicators , $y_1, \cdots, y_3$, for the latent endogenous variable η. The observed variable covariance matrix contains ($\frac{12(12+1)}{2} = 78$) unique elements. The model contains 44 parameters to be estimated : eight of the the twelve factor loading , twelve error variances, eight factor variances , three covariance between main effects, three covariance betweeen main effects and the product of the latent variables three way interaction term ,and three covariance between two way interactions term. All variables were simulated to come from the following population parameters

$$\eta = 0.3\xi_1 + 0.4\xi_2 + 0.5\xi_3 + 0.1\xi_1\xi_2 + 0.2\xi_1\xi_3 + 0.2\xi_2\xi_3 + \gamma_7\xi_1\xi_2\xi_3 + \zeta \qquad (3.8)$$

Where ξ_1, ξ_2 and ξ_3 are standard normal variables. The values of γ_1 to γ_6 paths were chosen based on values used by Klein and MuthÃĩn (2007). The values of γ_7 varied depending on the magnitude of the interaction effect size.

The errors for the 12 indicators in the measurement model were generated with the variances of the errors chosen so that the reliability of each indicator is 0.64. These population values are chosen so that the variances of the factor indicators are one which makes the parameter values more easily interpretable. Reliability is calculated as the ratio of the variance of the factor indicator explained by the factor to the total variance of the factor indicator using the following expression,

$$\frac{\lambda^2 * \psi}{\lambda^2 * \psi + \theta}$$

where λ is the factor loading, ψ is the factor variance, and θ is the residual variance of the factor indicators. We have used indicator reliability because of it has been shown to affect power to detect interaction effect in a latent variable interaction model (Weiss 2010). We chose the indicator reliabilities to be equal across the 12 indicator variables. The latent factor, ξ_1, ξ_2, ξ_3, were generated under the distributional study conditions with mean 0 and variance 1.

The error term ζ was generated from a normal distribution with mean 0 and variance 0.4 which is the same value used by Klein and MuthÃľn (2007).

Sample size (n=50 n=100, n=250,n=500)were used in the current study. Past simulation studies investigating interactions between two latent variables have used similar sample sizes (Klein and MuthÃľn, 2007; Marsh et al., 2004). The loading of 0.8 was selected to represent adequate loading size and is comparable to what has been used in previous studies (Klein and MuthÃľn, 2007; Little et al., 2006; Marsh et al., 2004).

In the first simulation study,the correlation between the two first-order latent variables ξ_1, ξ_2 and ξ_3 were set equal to the values used by Klein and MuthÃľn(2007):$\phi_{11} = \phi_{22} = \phi_{33} = 1, \phi_{12} = 0.3, \phi_{13} = 0.1, \phi_{32} = 0.2$. When first-order latent variables are strongly related, the standard errors associated with the gamma estimates will become very large (Cohen et al., 2003). Thus, for the current study a larger value for $\phi_{12}, \phi_{13}, \phi_{32}$ were selected to investigate the robustness of the standard errors when the covariance of the latent exogenous factors were high.

The distributions of ξ_1, ξ_2 and ξ_3 were manipulated to be either normal or non-normal. Previous research that examined skewed distributions for indicators, used a $\chi^2_{df=6}$, with skewness ≈ 1.15 (Marsh et al., 2004). The current study manipulated the distribution of ξ_1, ξ_2 and ξ_3 to be either normal or non-normal with value of skewness and Kurtosis equal to the value used by Klein and Muthen(2007)simulated using (skweness, Kurtosis)=(-1.5, 4),(-1.5,5), and (0.5, 0.5) respectively. The variance of the interaction term varied depending on the correlation between the two first-order latent variable ξ_1, ξ_2 and ξ_3 as shown under section (3.4).

The effect size represents the additional variance that the three way interaction effect term explains in η above and beyond that which can be explained by the first-order effects and the other three two way interaction term (Marsh et al., 2004) as shown below.

$$R^2_{\gamma 7} = \gamma_7^2 [\frac{\phi_{11}(\phi_{22}\phi_{33} + 2\phi_{23}\phi_{23}) + 2\phi_{12}(\phi_{12}\phi_{33} + 2\phi_{13}\phi_{23}) + 2\phi_{13}(2\phi_{12}\phi_{32} + \phi_{13}\phi_{22})}{\sigma^2_\eta}]$$

Jaccard and Wan (1995) did a review of the social science literature and found that interaction effect sizes typically accounted for 0.05 and 0.1 of the variance in the dependent variable in the case of two-way latent interaction effects. In the case of three-way interaction effects, the current study chose similar effect sizes for interaction effects in which the proportion of variance in η accounted for by the interaction effect was set equal to .0 (to investigate Type I error rates), .05, and .10 (to investigate power)

The squared multiple correlation R^2 is

$$
\begin{aligned}
R^2 = \gamma_1^2\phi_{11} + \gamma_2^2\phi_{22} + \gamma_3^2\phi_{33} + \gamma_4^2(\phi_{11}\phi_{22} + \phi_{12}^2) + \gamma_5^2(\phi_{11}\phi_{33} + \phi_{13}^2) \\
+ \gamma_6^2(\phi_{22}\phi_{33} + \phi_{23}^2) + \gamma_7^2[\phi_{11}(\phi_{22}\phi_{33} + 2\phi_{32}) \\
+ 2\phi_{12}(\phi_{12}\phi_{33} + 2\phi_{13}\phi_{23}) + 2\phi_{13}(2\phi_{12}\phi_{32} \\
+ \phi_{13}\phi_{22})] + 2[\gamma_1\gamma_2\phi_{12} + \gamma_1\gamma_3\phi_{13} + \gamma_2\gamma_3\phi_{23} \\
+ \gamma_1\gamma_7(\phi_{11}\phi_{23} + \phi_{12}\phi_{13} + \phi_{13}\phi_{12}) \\
+ \gamma_2\gamma_7(\phi_{21}\phi_{23} + \phi_{22}\phi_{13} + \phi_{23}\phi_{12}) \\
+ \gamma_3\gamma_7(\phi_{31}\phi_{23} + \phi_{32}\phi_{13} + \phi_{33}\phi_{12})] \\
/\sigma_\eta^2
\end{aligned}
$$

For the interaction effect size 0,0.05,0.1 and the population variance covariance matrix defined above, squared multiple correlation is ,65.95%, 71.61%, 74.65%,72.86%,79.93%,and 83.01 respectively.

Hence the design of our study is 3(effect size)x 4 (sample size)x 2(latent variable distribution levels)x 2(latent covariance) completely crossed factorial design resulting in 48 possible combinations.Once the data were generated, they were analyzed with Mplus 7.4.

For each of the 48 possible condition combinations, 500 data sets were generated with Mplus version 7.4. This decision was based on the number of replications used in previous studies, and factors that are known to influence the number of necessary replications for Monte Carlo simulations. For instance, Powell and Schafer (2001) conducted a meta analysis of 219 simulation studies in structural equation modeling and reported that the number of replications used in these studies ranged from 20 to 1,000, with the median number of replications being 200. Similarly, Bandalos (2006) suggested that 500 replications were large for SEM Monte Carlo simulation studies. She argued that this number of replications would provide stable standard error estimates even when data were generated to come from a non-normal distribution. To check the stability of the model estimation, we have used different seeds to implement the same Monte Carlo simulations, and the model results basically remain unchanged. Thus we conclude that Monte Carlo simulation results are stable.

Criteria for Evaluating Models

In this study, the performance of Latent moderated structural equations(LMS) was evaluated on the basis of different conditions. Bias for each parameter was calculated as the difference between the mean of the estimates obtained from the 500 replications and the true value. Hoogland and Boomsma (1998) suggested as a criterion that the absolute value of bias be less than .05 for parameter estimates in a particular condition, or across conditions in a study, to be considered unbiased.

Let M be the number of replications in a condition. Absolute bias (or raw bias) of point estimates, defined as the Monte Carlo average of the point estimates minus the true parameter value,

$$B(\theta) = M^{-1}\Sigma_{i=1}^{M}(\widehat{\theta}_i - \theta)$$

,can tell the difference between the true parameter value and the mean of parameter estimates across replications. Relative bias of point estimates, defined as raw bias divided by the true parameter value (when the ratio is well-defined),

$$B_r(\theta) = M^{-1}\Sigma_{i=1}^{M}(\frac{\widehat{\theta}_i - \theta)}{\theta}$$

,is the proportion of absolute bias relative to the true parameter value. In the conditions in which the population-generating parameter value γ_7 was equal zero the bias cannot be divided by zero, therefore the proportion relative bias was equal to the difference between the average parameter estimates over the 500 replications and the population-generating.There are varying guidelines in the literature for interpreting bias values. For example, Hoogland and Boomsma (1998) suggested that absolute values of Bias less than .05 could be considered to represent a lack of bias. MuthÃľn, Kaplan, and Hollis (1987) offered the more lenient criterion that bias of less than .10 to .15 might be considered negligible.In the current study, we adopted Hoogland and Boomsma's (1998) suggested value of .05 to demarcate bias and unbiased parameter estimates.

To evaluate the standard errors, the mean of the estimated standard errors for that parameter across the replications should be compared with the Monte Carlo standard deviation of a given parameter estimate.
Let

$$SE(\widehat{\theta}) = M^{-1}\Sigma_{i=1}^{M} SE(\widehat{\theta}_i)$$

be the mean of the estimated standard errors, where $SE(\widehat{\theta}_i)$ is the estimated standard error from replication i, and let

$$SD(\widehat{\theta}) = \{\frac{1}{M-1}\Sigma_{i=1}^{M}(\widehat{\theta}_i - \overline{\widehat{\theta}})\}^{\frac{1}{2}}$$

be the Monte Carlo standard deviation of the point estimates, where $\overline{\widehat{\theta}}$ is the mean of point estimates. As a comparable measure of accuracy of the standard error estimates, the relative bias of the standard errors is also reported. It is defined as:

$$B_r(SE) = \frac{SE(\widehat{\theta}) - SD(\widehat{\theta})}{SD(\widehat{\theta})}$$

Values close to zero are desirable because they indicate that the standard error values that are computed based on the model, are representative of what is in the population. Values less than zero indicate that the model underestimates the standard error estimates, while values greater than zero indicate that the model overestimates the standard error estimates.

Root Mean Square Error (RMSE), which can provide information of both the distance of each parameter estimate from the true value and the variability of such distances is also calculated. For a generic parameter θ, RMSE is defined as

$$RMSE = \{\frac{1}{M}\Sigma_{i=1}^{M}(\widehat{\theta}_i - \theta)\}^{\frac{1}{2}}$$

, where $\widehat{\theta}_i$ is the estimate from replication i, θ is the true value. Convergence to a proper solution was also tracked, as the estimation method demonstrated a lack of convergence to a proper solution at the more extreme levels of the condition (small sample size). In situations in which replicates did not converge, replacement replications were not generated. Thus, the computations of the outcome criteria are based on only those converged replications.

Chapter 4

Results and Discussion

4.1 Results of the Simulation Study

4.1.1 Parameter recovery and Accuracy of standard error estimate

Bias for main effects' regression coefficients

While the bias of the γ_7 parameter was the primary interest, bias was also examined for the main effects and two-way interaction term. Bias of the main effects,γ_1, γ_2, and γ_3, were examined across different conditions (see table 5.2). With small sample size(i.e, n=50)and moderate reliability (reliability=0.64), this bias was very high.The resulting overestimation decreased as reliability of the indicators and sample size increased, but kept increasing as the interaction effect size for the three-way interaction term ($(R^2_\gamma 7)$) and co-variance between latent exogenous variables increased. That is, bias decreased as $\phi_{12}, \phi_{13}, \phi_{23}$ decreased. In reference to the criterion of .05, the estimation method in this study (LMS) produced unbiased estimates for sample size 500 and also for n=250 with high reliability (0.84). Therefore. with moderate reliability and small sample size(i.e, n=50), the bias estimates for γ_1, γ_2, and γ_3 resulting from LMS approach cannot be trusted.

The column labeled SE-Bias in table 5.2 stands for standard error bias for the estimates of γ_1, γ_2, and γ_3. It was found that, this bias is very large (in absolute value) with small sample size(n=50) and moderate reliability indicators, indicating that the LMS approach underestimated standard errors. With the same sample size(n=50), the standard error bias for the estimate of first-order effects decreased as $((R^2_\gamma 7))$ and reliability increased. For all sample size under study,this bias increased as the covariance of latent exogenous increased which is consistent with result of Cohen et al., 2003. In reference to the criterion of 0.1, the LMS produced unbiased estimates for main effects with sample size 100 and greater but underestimated standard errors because most values were negatives. However, the standard error estimates were fairly accurate when the sample size was 500.

Bias for two-way interaction term regression coefficients

Biases of the latent two-way interaction effects,γ_4, γ_5 and γ_6, were examined across the conditions of interaction effect size, reliability of indicators , sample size, and covariances of latent exogenous variables (see table 5.3 through table 5.8). The bias for these interaction effects was also very high when sample size 50 and with moderate reliability. It has been shown that increasing the interaction effect size $(R^2_\gamma 7)$ increased the bias for the two-way interaction term. Similarly, this biased increased with the increase of the covariances. But the increment of indicators reliability and sample sized reduced the biases.
The standard errors for the estimates of two-way interactions term was shown in table 5.3 through 5.8. As indicated in tables this value was decreased as the covariance between factors decreased. Increasing the reliability of the indicators and sample size reduced the standard errors of these estimates. But the increment of interaction effect size increased the standard errors.

The variability of the two-way interaction term parameter estimates as measured by root mean square error showed changes across different conditions under study. Because the RMSE comprises both parameter estimate variability and bias, it's values were high for small sample size(ie., n=50) and moderate reliability, as it indicated by high values for bias of the parameter estimates for this conditions(see table 5.3 through 5.8). Hence, as predicted, the parameter was more precisely estimated as the sample size and indicator reliability increased.

Bias for three-way interaction term regression coefficients

Table 5.9 shows the latent moderated structural equations (LMS) approach parameter estimates of γ_7 in the all conditions understudy. Perhaps not too surprisingly,bias was greatest for sample size 50 coupled with moderate indicator reliability, but reduced almost by 10% when reliability of the indicators was good (e.g., reliability = 0.84). In the same conditions the standard error and root mean square error reduced by 8% and 10 % respectively for the increment of indicator reliability from 0.64 to 0.84. As anticipated, bias across conditions decreased as sample size increased, but there was a pattern indicative of diminishing returns for sample sizes larger than 500. The increment of interaction effects size resulted increased bias, standard errors and RMSE at small sample size(i.e,n=50), but showed inconstant pattern for the sample size greater than 50. Similarly, for the increase of the covariance between latent factors, the bias, standard error and RMSE reduced for small sample. However, this properties showed inconstant pattern for the others samples size in study.

4.1.2 Testing hypotheses of three -way interaction effects

Type I error rates and empirical power

As previously stated, the proportion of variance in η accounted for by the three-way interaction effect was set equal to .00 (to investigate Type I error rates), .05, and .10 (to investigate power).The empirical Type I error rates of the nominal size Îś = .05 two-sided tests (under the null hypothesis, H0 : $\gamma_7 = 0$) when using the LMS procedure are given in Table 5.11. The Type I error rate was computed as the proportion of converged solutions that had a statistically significant three-way interaction effect (at the .05 level) in the simulated data when H0 was true. In addition, empirical power (probability of rejecting a false null hypothesis, H0 : $\gamma_7 = 0$) was represented by the proportion of converged solutions that have a statistically significant interaction effect in the simulated data when H0 was false (Marsh et al., 2004)and tabulated in table 5.12 under the 5% and 10% effect size conditions.

Type I error rates.

When the sample size was 50,100 and indicator reliability 0.64, type I error rates closest to the desired Îś level, but increased as the covariance between latent factors and indicator reliability increased. Moreover, when sample size 100 and reliability was 0.84 the approach in this study (LMS)had very high Type I error rates, rejecting 10% of true models. In this condition(indicator reliability 0.84), the approach under study rejected the null hypothesis (with all the samples) more frequently than the nominal level would predict, except when coupled with moderate reliability. In general, in this study the type I error increased as latent factor covariance and indicators reliability increased(see table 5.11).

Empirical power

Empirical power is represented by the proportion of converged solutions that have a significant interaction effect in the simulated data when the population interaction effect is not equal to zero. Empirical power rates for effect size $R^2_\gamma 7 = 0.05$ and $R^2_\gamma 7 = 0.1$ were computed using an Îś level of .05, and are shown in Tables 5.12. As anticipated, empirical power increased as the size of the effect increased from 5% to 10% across methods and conditions. That is, when medium to large three-way interaction effects exist in the population, the methods were able to detect them with a great deal of certainty for moderate sample sizes under high reliability. This was the case even when the sample size was extremely small (n = 50) and the indicators were moderate(reliability = .64).Predictably, power increased as reliability and sample size increased. In general, power for the LMS approach under study increased as $R^2_\gamma 7$ increased, sample size increased, and $\phi_{12}, \phi_{13}, \phi_{23}$ increased.

4.2 Discussion

Although many simulation studies have been conducted to study latent interaction effects in nonlinear SEM, majority of these studies has focused on two-way latent interactions and quadratic effects. In current study an examination of three-way continuous latent interaction effects was conducted via monte carlo simulation using latent moderated structural method. The simulated data were varied as a function of the size of the three-way interaction term effect, sample size, indicator reliability and the size of the relation between first-order latent variables.

Our findings in the simulation studies indicate that, when indicator reliability was moderate and three-way interaction effect present in the generating population-generating model(i.e, $R^2_{\gamma 7} \neq 0.00$), the LMS method led to biased estimate of interaction effect. As with past simulation studies in two-way interaction, indicator variable reliability tended to have the greatest impact on the ability of the LMS to accurately and precisely estimate the three-way interaction effect with size of the relation between the first-order latent variable exerting less influence. Moreover,Parameter estimates for the LMS approach became less biased as the size of the interaction effect and the correlation between the first-order latent variables decreased. We observed that this result for three-way interaction was similar to previous findings of two -way interaction in which the LMS approach was found to result in unbiased estimates of the interaction effect across all sizes of the interaction effect (Klein and Moosbrugger, 2000; and Klein and MuthÃĺn, 2007).

This finding suggest that the method appeared to control Type I error fairly by reducing the size of the relation between first-order latent variables. Hence moderate indicator reliability, and small sample sizes appear to have the greatest negative impact on the estimation accuracy, precision, and deflation of standard errors of the three-way interaction parameter. Because the method investigated here performed poorly under these circumstances, if data exhibit these characteristics in practice, statistical conclusions should be made cautiously.

Limitations

Normality

Because of only normal condition was used in the current study, it is unknown how non-normal distribution need to be for parameter estimates and their precision to change. For instance, previous studies investigated the impact that non-normality had on estimating interaction effect in structural equation modeling (Marsh et al., 2004; klein and Moosbrugger, 2004; Klein and Muthen, 2007; and Wall and Amemiya, 2001).

Number of Indicators

Previous simulation studies have used two or three indicators to represent each latent variable for estimating interaction effect in nonlinear structural equation modeling. The current study used three indicators for each latent variable. In the previous latent interaction,it was found that more indicators may be necessary to avoid problems such as non-convergence and inaccurate parameter estimation (Kline,2005).

Limitations of software

Analyses were conducted for the current study in MPlus version 7.4 and default settings within a given software program were used. It is the current software to do three-way interaction in structural equation modeling. The default number of maximum iterations with Mplus is 1000 and the starting values were set equal to the values in the population-generating model.Applied researchers may choose to change the default settings, but this decision would be specific to their particular data set. One limitation of the LMS approach observed in this study was that it becomes increasing slow to converge as the number of latent variables increases and the length of time it took to reach convergence was positively related to the sample size. In the current study the structural model had four nonlinear term(three two-way and one three-way interaction term).

Chapter 5

Conclusion and Recommendation

It is the conclusion of the author that the latent moderated structural approach can be used to study three-way continuous latent interaction in nonlinear structural equation modeling using Mplus software. The approach had no model convergence problems across the conditions in the study and did not produced unrealistic estimates. However, because of the complexity of the model, it took along time to get monte carlo simulation output.

In the conditions considered in the current study, the method led to the least biased estimates of the interaction effect, and accurate standard error estimates, particularly when the sample size was 250 or greater and the indicator reliability was high. Additionally, the latent moderated structural approach accurately estimated first-order effects provided that the sample size was 250 or greater. For the small size(i.e, n=50), the bias for interaction effects and exogenous regression coefficients was high. But for the same sample size, the method had less bias (approximately less than 2%) in estimating the exogenous covariances. This bias increased as the interaction effect size ($R^2_{\gamma 7}$) increased and decrease when sample size increased.

Type I error rates were close to the desired alpha level, particularly when the sample size was 250 or greater. Comparing to other conditions in the study, when indicator reliability were low and the sample size was 50, the method had low power to detect true three-way interaction effects and a sample size of at least 250 was necessary to have acceptable power(greater than

0.8).

Based on these findings,high indicator reliability and a sample size of 250 or more is recommended for use with the latent moderated structural method, although it performs fairly well with sample sizes of 100. It also recommended that,under small sample (i.e,n=50),the method provided sufficient power to detect the three-way interaction effects when high indicator reliability and the covariance of the exogenous latent variable was increased.

REFERENCES

Aiken, L. S., & West, S. G. (1991). Multiple regression Testing and interpreting interactions. Newbury Park: SAGE Publications.

Algina, J., & Moulder, B. C. (2001). A note on estimating the JÃűreskog-Yang model for latent variable interaction using LISREL 8.3. Structural Equation Modeling: A Multidisciplinary Journal, 8, 40-52.

Arminger, G., & Muthen, B. O. (1998). A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the Metropolis-Hastings algorithm. Psychometrika, 63(3), 271-300.

Bandalos, D. L. (2006). The use of Monte Carlo studies in structural equation modeling research. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course (pp. 385 - 426).* Greenwich, CT: Information Age Publishing

Bohnstedt, G. W., & Goldberger, A. S. (1969).On the exact covariance of products of random variables.*Journal of the American Statistical Association, 64, 325-328.*

Bollen, K.A. (1989), Structural Equations with Latent Variables, New York: John Wiley & Sons, Inc.

Bollen, K. A. (1995). Structural equation models that are nonlinear in latent variables: A least-squares estimator. Sociological Methodology, 25, 223-251.

Bollen, K. A. (1996). An alternative two stage least squares (2SLS) estimator for latent variable equations. Psychometrika, 61(1), 109-121.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003).*Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.*

Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. Psychological Methods. 1996; 1:16-29. doi:10.1037/1082-989X. 1.1.16.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of Royal Statistical Society. Ser. B, 39, 1-38.

Gerry Gray (1999). Covariances in Multiplicative Estimates.*Transactions of the American Fisheries Society,128:3, 475-482, DOI: 10.1577/1548-8659(1999)128<0475:CIME>2.0.CO;2.*

Harring, Jeffrey R.; Weiss, Brandi A.; Hsu, Jui-Chen(2012).A comparison of methods for estimating quadratic effects in nonlinear structural equation models.*Psychological Methods, Vol 17(2), 193-214.*

Hayduk, LA.1987. Structural equation modeling with LISREL. *Essentials and advances.* Johns Hopkins University; Baltimore, MD.

Hu L, Bentler PM, Kano Y. Can test statistics in covariance structure analysis be trusted? Psychological Bulletin. 1992; 112:351-362. doi:10.1037/0033-2909.112.2.351. [PubMed:

1454899]

Jaccard, J., & Wan, C. K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin, 117, 348 - 357.*

Jonsson, F. (1998). Modeling interaction and nonlinear effects: A step-by-step LISREL example. In: R. E. Schumacker & G. A. Marcoulides (eds.), Interactions and Nonlinear Effects in Structural Equation Models. Mahwah, NJ: Lawrence Erlbaum, pp. 17-42.

Joreskog, K. G. & Yang, F. (1996). Nonlinear structural equation models: The KennyJudd model with interaction effects. In G. A. Marcoulides & R. E. Schumacker (Eds.), Advanced structural equation modeling: Issues and techniques (pp. 57-88). Mahwah, NJ: Erlbaum.

Kendall,M.G.,and Stuart, A.(1958). *The advanced Theory of Statistics(vol.1).* London:Griffin

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin, 96, 201-210.*

Klein, A. G. & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika, 65: 457-474.*

Klein, A.G. & MuthÃĺn, B. (2007). Quasi-maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research, 42, pp 647-673.*

Little, T. D., Bovaird, J. A., & Widaman, K. F. (2006). On the merits of orthogonalizing powered and product terms: Implications for modeling interactions among latent variables. *Structural Equation Modeling: A Multidisciplinary Journal, 13, 497-519.*

Marsh, H. W., Wen, Z., & Hau, K. T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods, 9, 275 - 300.*

Marsh, H. W., Wen, Z., & Hau, K. T. (2006). Structural equation models of latent interaction and quadratic effects. In G. R. Hancock & R. O. Mueller (Eds.), Structural equation modeling: A second course (pp. 225 - 265). Greenwich, CT: Information Age Publishing.

Powell, D. A., & Schafer, W. D. (2001). The robustness of the likelihood ratio chi-square test for structural equation models: A meta-analysis.*Journal of Educational and Behavioral Statistics, 26, 105-132.*

Schermelleh-Engel, K., Klein, A., & Moosbrugger, H. (1998). Estimating nonlinear effects using a latent moderated structural equations approach. In R. E. Schumacker & G. A. Marcoulides (Eds.), Interaction and nonlinear effects in structural equation modeling (pp. 203-238). Mahwah, NJ: Lawrence Erlbaum Associates.

Wall, M. M., & Amemiya, Y., (2001). Generalized appended product indicator procedure for nonlinear structural equation analysis. Journal of Educational and Behavioral Statistics, 26, 1-29.

Wall, M. M. (2009). Maximum likelihood and Bayesion estimation for nonlinear structural equation models. In R. E. Millsap & A. Maydeu-Olivares (Eds.), The SAGE Handbook of Quantitative Methods in Psychology (pp. 540-567). SAGE Publications.

Weiss, BA(2010) A comparison of methods for testing interaction effects in structural equation modeling. University of Maryland; College Park.

Appendix

Tables

Table 5.1: Summary of Manipulated Features

Factor	1	2	3	4			
Sample size	50	100	250	500			
Indicator reliability	0.64	0.84					
Effect size($(R^2_\gamma 7)$)	0.00	0.05	0.10				
$\phi_{12}, \phi_{13}, \phi_{23}$	0.3	0.1	0.2		0.6	0.4	0.5
Distribution of ξ_1, ξ_2, ξ_3	Normal						
Factor loading	0.8						
Estimation method	LMS						

Table 5.2: Parameter Estimates for the first-order main effects γ_1 to γ_3 when $R^2_{\gamma 7} = 0.05, 0.1$ and with different covarince of exogenous latent variables

	$\phi_{12} = 0.3$ $\phi_{13} = 0.1$ $\phi_{23} = 0.2$						$\phi_{12} = 0.6$ $\phi_{13} = 0.4$ $\phi_{23} = 0.5$			
			$R^2_{\gamma 7} = 0.05$		$R^2_{\gamma 7} = 0.1$		$R^2_{\gamma 7} = 0.05$		$R^2_{\gamma 7} = 0.1$	
Rel.	N	Par.	Bias	SE-Bias	Bias	SE-Bias	Bias	SE-Bias	Bias	SE-Bias
0.64	50	γ_1	33.32	-0.929	55.146	-0.90	95.33	-0.968	142.297	-0.969
		γ_2	44.115	-0.944	48.110	-0.932	3.482	-0.958	-11.977	-0.962
		γ_3	40.259	-0.948	56.899	-0.940	58.602	-0.943	78.215	-0.939
	100	γ_1	0.182	-0.049	0.183	-0.047	0.207	-0.051	0.212	-0.045
		γ_2	0.141	-0.029	0.143	-0.028	0.148	-0.007	0.155	-0.008
		γ_3	0.172	-0.009	0.176	-0.026	0.198	-0.026	0.2	-0.056
	250	γ_1	0.069	-0.035	0.069	-0.027	0.081	-0.029	0.083	-0.026
		γ_2	0.051	-0.055	0.049	-0.058	0.054	-0.059	0.054	-0.063
		γ_3	0.062	-0.015	0.062	-0.017	0.007	-0.029	0.075	-0.025
	500	γ_1	0.039	-0.024	0.039	-0.025	0.045	-0.033	0.046	-0.037
		γ_2	0.023	-0.017	0.024	-0.016	0.022	-0.023	0.023	-0.024
		γ_3	0.031	-0.008	0.030	-0.009	0.036	-0.019	0.036	-0.018
0.84	50	γ_1	0.193	-0.126	0.197	-0.126	0.200	-0.138	0.209	-0.142
		γ_2	0.171	-0.109	0.174	-0.106	0.163	-0.109	0.169	-0.108
		γ_3	0.204	-0.074	0.209	-0.015	0.209	-0.089	0.217	-0.097
	100	γ_1	0.087	-0.052	0.088	-0.049	0.092	-0.057	0.094	-0.058
		γ_2	0.076	-0.097	0.077	-0.096	0.073	-0.066	0.075	-0.067
		γ_3	0.084	-0.006	0.085	-0.005	0.089	-0.024	0.092	-0.026
	250	γ_1	0.038	-0.045	0.038	-0.043	0.042	-0.036	0.043	-0.036
		γ_2	0.031	-0.053	0.031	-0.057	0.031	-0.055	0.032	-0.056
		γ_3	0.034	-0.008	0.034	-0.010	0.038	-0.015	0.039	-0.015
	500	γ_1	0.023	-0.035	0.024	-0.035	0.025	-0.029	0.026	-0.030
		γ_2	0.014	-0.027	0.014	-0.027	0.013	-0.034	0.014	-0.032
		γ_3	0.018	0.011	0.018	0.011	0.021	0.000	0.021	0.000

Table 5.3: Parameter Estimates for nonlinear effects γ_4 to γ_6 when $R^2_{\gamma 7} = 0$,reliability-0.64 and with different covarince of exogenous latent variables

	$\phi_{12} = 0.3$ $\phi_{13} = 0.1$ $\phi_{23} = 0.2$				$\phi_{12} = 0.6$ $\phi_{13} = 0.4$ $\phi_{23} = 0.5$		
N	Parameter	Bias	SE-Bias	RMSE	Bias	SE-Bias	RMSE
50	γ_4	41.491	-0.864	46.808	-12.866	-0.985	75.024
	γ_5	21.034	-0.937	74.840	40.408	-0.903	125.121
	γ_6	54.298	-0.945	99.375	63.581	-0.921	132.965
100	γ_4	0.038	-0.029	0.1334	0.062	-0.051	0.15
	γ_5	0.122	-0.101	0.1664	0.137	-0.115	0.226
	γ_6	0.095	-0.085	0.155	0.093	-0.066	0.189
	γ_7	-0.0154	-0.089	0.150	-0.014	-0.120	0.117
250	γ_4	0.023	-0.0744	0.073	-0.035	-0.06	0.082
	γ_5	0.064	-0.092	0.087	0.065	-0.121	0.113
	γ_6	0.047	-0.086	0.083	0.056	-0.075	0.102
500	γ_4	-0.003	-0.043	0.048	-0.028	-0.053	0.057
	γ_5	0.031	-0.020	0.055	0.037	-0.009	0.072
	γ_6	0.018	-0.050	0.054	0.020	-0.056	0.068

Table 5.4: Parameter Estimates for nonlinear effects γ_4 to γ_6 when $R^2_{\gamma 7} = 0$ and with different covarince of exogenous latent variables and reliability=0.84

	$\phi_{12} = 0.3$ $\phi_{13} = 0.1$ $\phi_{23} = 0.2$				$\phi_{12} = 0.6$ $\phi_{13} = 0.4$ $\phi_{23} = 0.5$		
N	Parameter	Bias	SE-Bias	RMSE	Bias	SE-Bias	RMSE
50	γ_4	0.067	-0.093	0.152	0.052	-0.143	0.178
	γ_5	0.148	-0.212	0.200	0.158	-0.188	0.246
	γ_6	0.115	-0.133	0.177	0.121	-0.129	0.208
100	γ_4	-0.001	-0.089	0.092	0.020	-0.087	0.102
	γ_5	0.052	-0.135	0.110	0.058	-0.141	0.141
	γ_6	0.017	-0.077	0.100	0.025	-0.095	0.122
250	γ_4	0.004	-0.055	0.052	-0.034	-0.051	0.058
	γ_5	0.037	-0.085	0.062	0.033	-0.053	0.076
	γ_6	0.029	-0.092	0.061	0.033	-0.094	0.073
500	γ_4	-0.004	-0.025	0.036	-0.017	-0.042	0.041
	γ_5	0.018	-0.015	0.040	0.019	-0.007	0.051
	γ_6	0.011	-0.059	0.040	0.013	-0.071	0.050

Table 5.5: Parameter Estimates for nonlinear effects γ_4 to γ_6 when $R^2_{\gamma 7} = 0.05$,reliability=0.64 and with different covarince of exogenous latent variables

	$\phi_{12} = 0.3$ $\phi_{13} = 0.1$ $\phi_{23} = 0.2$				$\phi_{12} = 0.6$ $\phi_{13} = 0.4$ $\phi_{23} = 0.5$		
N	Parameter	Bias	SE-Bias	RMSE	Bias	SE-Bias	RMSE
50	γ_4	0.98	-0.90	55.04	16.96	-0.886	85.279
	γ_5	17.05	-0.91	65.02	68.726	-0.941	183.918
	γ_6	63.74	-0.95	118.21	35.610	-0.933	141.246
100	γ_4	0.034	-0.041	0.139	0.042	-0.048	0.163
	γ_5	0.123	-0.090	0.169	0.149	-0.079	0.234
	γ_6	0.110	-0.097	0.163	0.121	-0.085	0.21
250	γ_4	0.03	-0.091	0.079	-0.031	-0.074	0.091
	γ_5	0.07	-0.091	0.090	0.077	-0.080	0.121
	γ_6	0.044	-0.078	0.085	0.049	-0.069	0.108
500	γ_4	0.001	-0.018	0.050	-0.030	-0.052	0.061
	γ_5	0.034	-0.023	0.057	0.043	-0.018	0.077
	γ_6	0.015	-0.033	0.055	0.012	-0.039	0.071

Table 5.6: Parameter Estimates for nonlinear effects γ_4 to γ_6 when $R^2_{\gamma 7} = 0.05$ and with different covarince of exogenous latent variables and reliability=0.84

	$\phi_{12} = 0.3$ $\phi_{13} = 0.1$ $\phi_{23} = 0.2$				$\phi_{12} = 0.6$ $\phi_{13} = 0.4$ $\phi_{23} = 0.5$		
N	Parameter	Bias	SE-Bias	RMSE	Bias	SE-Bias	RMSE
50	γ_4	0.070	-0.064	0.155	0.021	-0.142	0.189
	γ_5	0.158	-0.219	0.209	0.176	-0.197	0.259
	γ_6	0.122	-0.149	0.185	0.137	-0.137	0.22
100	γ_4	-0.002	-0.076	0.096	0.004	-0.083	0.109
	γ_5	0.058	-0.123	0.113	0.068	-0.121	0.145
	γ_6	0.023	-0.083	0.105	0.032	-0.085	0.128
250	γ_4	0.005	-0.074	0.057	-0.034	-0.055	0.063
	γ_5	0.038	-0.086	0.065	0.034	-0.055	0.081
	γ_6	0.028	-0.098	0.064	0.031	-0.094	0.077
500	γ_4	-0.007	0.000	0.037	-0.023	-0.041	0.043
	γ_5	0.017	-0.019	0.042	0.020	-0.015	0.054
	γ_6	0.010	-0.041	0.041	0.010	-0.054	0.051

Table 5.7: Parameter Estimates for estimating nonlinear effects γ_4 to γ_6 when $R^2_{\gamma 7} = 0.1$ and with different covarince of exogenous latent variables and reliability=0.64

	$\phi_{12} = 0.3$ $\phi_{13} = 0.1$ $\phi_{23} = 0.2$				$\phi_{12} = 0.6$ $\phi_{13} = 0.4$ $\phi_{23} = 0.5$		
N	Parameter	Bias	SE-Bias	RMSE	Bias	SE-Bias	RMSE
50	γ_4	0.988	-0.866	64.23	33.021	-0.916	131.136
	γ_5	43.34	-0.897	100.87	140.706	-0.941	297.476
	γ_6	63.18	-0.938	142.597	-9.077	-0.934	205.969
100	γ_4	0.047	-0.035	0.145	0.045	-0.045	0.174
	γ_5	0.126	-0.085	0.1752	0.153	-0.079	0.248
	γ_6	0.118	-0.109	0.173	0.143	-0.146	0.241
250	γ_4	0.031	-0.099	0.084	-0.03	-0.081	0.097
	γ_5	0.072	-0.084	0.093	0.083	-0.084	0.1269
	γ_6	0.042	-0.072	0.088	0.047	-0.068	0.114
500	γ_4	0.001	-0.011	0.053	-0.034	-0.054	0.065
	γ_5	0.035	-0.024	0.059	0.046	-0.032	0.081
	γ_6	0.013	-0.028	0.057	0.009	-0.038	0.074

Table 5.8: Parameter Estimates for nonlinear effects γ_4 to γ_6 when $R^2_{\gamma 7} = 0.1$ and with different covarince of exogenous latent variables and reliability=0.84

	$\phi_{12} = 0.3$ $\phi_{13} = 0.1$ $\phi_{23} = 0.2$				$\phi_{12} = 0.6$ $\phi_{13} = 0.4$ $\phi_{23} = 0.5$		
N	Parameter	Bias	SE-Bias	RMSE	Bias	SE-Bias	RMSE
50	γ_4	0.072	-0.051	0.160	0.006	-0.138	0.198
	γ_5	0.164	-0.216	0.216	0.187	-0.199	0.269
	γ_6	0.126	-0.149	0.191	0.147	-0.137	0.229
100	γ_4	-0.002	-0.068	0.100	0.000	-0.079	0.115
	γ_5	0.060	-0.114	0.117	0.071	-0.114	0.150
	γ_6	0.027	-0.083	0.110	0.035	-0.083	0.133
250	γ_4	0.005	-0.080	0.060	-0.035	-0.057	0.068
	γ_5	0.038	-0.082	0.067	0.035	-0.057	0.084
	γ_6	0.028	-0.099	0.068	0.031	-0.097	0.082
500	γ_4	-0.009	0.002	0.039	-0.027	-0.041	0.047
	γ_5	0.017	-0.023	0.045	0.02	-0.021	0.056
	γ_6	0.010	-0.034	0.043	0.009	-0.047	0.054

Table 5.9: Parameter Estimates for nonlinear effects γ_7 across the study conditions

	$\phi_{12}=0.3$ $\phi_{13}=0.1$ $\phi_{23}=0.2$						$\phi_{12}=0.6$ $\phi_{13}=0.4$ $\phi_{23}=0.5$		
Rel.	$R^2_{\gamma 7}$	N	Parameter	Bias	SE-Bias	RMSE	Bias	SE-Bias	RMSE
0.64	0.0	50	γ_7	2.0698	-0.924	84.206	-0.471	-0.853	53.911
		100	γ_7	-0.0154	-0.089	0.150	-0.014	-0.120	0.117
		250	γ_7	-0.007	-0.029	0.069	-0.007	-0.079	0.055
		500	γ_7	-0.0003	-0.026	0.047	-0.003	-0.092	0.037
	0.05	50	γ_7	58.87	-0.96	138.74	35.704	-0.901	62.404
		100	γ_7	-0.025	-0.158	0.173	-0.004	-0.208	0.152
		250	γ_7	-0.033	-0.068	0.079	-0.029	-0.096	0.069
		500	γ_7	0.002	-0.019	0.053	-0.006	-0.060	0.045
	0.1	50	γ_7	73.838	-0.965	249.624	46.579	-0.925	107.664
		100	γ_7	0.004	-0.186	0.202	0.036	-0.323	0.216
		250	γ_7	-0.02	-0.079	0.092	-0.013	-0.102	0.081
		500	γ_7	0.003	-0.017	0.059	0.000	-0.032	0.052
0.84	0.0	50	γ_7	0.003	-0.171	0.182	0.002	-0.189	0.151
		100	γ_7	-0.004	-0.187	0.104	-0.003	-0.158	0.081
		250	γ_7	-0.003	-0.080	0.053	-0.004	0.088	0.041
		500	γ_7	0.000	-0.065	0.036	-0.0016	-0.077	0.028
	0.05	50	γ_7	0.088	-0.174	0.198	0.102	-0.179	0.171
		100	γ_7	-0.005	-0.176	0.114	0.000	-0.126	0.093
		250	γ_7	-0.012	-0.089	0.060	-0.016	-0.080	0.051
		500	γ_7	0.004	-0.042	0.040	-0.002	-0.041	0.035
	0.1	50	γ_7	0.085	-0.175	0.214	0.102	-0.176	0.193
		100	γ_7	0.002	-0.169	0.125	0.007	-0.114	0.106
		250	γ_7	-0.006	-0.091	0.068	-0.008	-0.075	0.060
		500	γ_7	0.005	-0.029	0.045	0.002	-0.020	0.040

Table 5.10: Bias and Power estimates for covarince of exogenous latent variables with $R^2_{\gamma 7} = 0.05, 0.1$ and the two reliability of the latent indicators

	$\phi_{12} = 0.3$						$\phi_{12} = 0.6$			
	$\phi_{13} = 0.1$						$\phi_{13} = 0.4$			
	$\phi_{23} = 0.2$						$\phi_{23} = 0.5$			
			$R^2_{\gamma 7} = 0.05$		$R^2_{\gamma 7} = 0.1$			$R^2_{\gamma 7} = 0.05$	$R^2_{\gamma 7} = 0.1$	
Rel.	N	Par.	Bias	Power	Bias	Power	Bias	Power	Bias	Power
0.64	50	ϕ_{12}	-0.010	0.474	-0.018	0.470	-0.007	0.962	-0.008	0.964
		ϕ_{13}	-0.197	0.101	-0.199	0.101	-0.043	0.675	-0.043	0.671
		ϕ_{23}	-0.035	0.253	-0.041	0.261	-0.022	0.871	-0.020	0.869
	100	ϕ_{12}	-0.004	0.752	-0.004	0.750	-0.003	1.000	-0.003	1.000
		ϕ_{13}	-0.114	0.134	-0.115	0.138	-0.022	0.936	-0.021	0.940
		ϕ_{23}	0.006	0.422	0.006	0.42	-0.002	0.994	0.248	0.994
	250	ϕ_{12}	-0.007	0.988	-0.006	0.988	-0.003	1.000	-0.003	1.000
		ϕ_{13}	-0.06	0.244	-0.061	0.238	-0.010	1.000	-0.010	1.000
		ϕ_{23}	0.001	0.784	0.001	0.786	-0.001	1.000	0.000	1.000
	500	ϕ_{13}	0.000	1.000	0.000	1.000	0.000	1.000	0.000	1.000
		ϕ_{13}	-0.046	0.408	-0.047	0.412	-0.009	1.000	-0.009	1.000
		ϕ_{23}	0.018	0.964	0.018	0.962	0.004	1.000	0.004	1.000
0.84	50	ϕ_{12}	-0.008	0.582	-0.008	0.583	-0.004	0.994	-0.005	0.994
		ϕ_{13}	-0.197	0.112	-0.200	0.114	-0.045	0.810	-0.045	0.812
		ϕ_{23}	-0.021	0.326	-0.020	0.331	-0.019	0.934	0.226	0.936
	100	ϕ_{12}	-0.005	0.850	-0.005	0.854	-0.004	1.000	-0.004	1.000
		ϕ_{13}	-0.130	0.164	-0.130	0.166	-0.027	0.974	-0.027	0.974
		ϕ_{23}	0.006	0.504	0.007	0.506	-0.004	0.998	-0.004	0.998
	250	ϕ_{23}	-0.006	0.994	-0.006	0.994	-0.004	1.000	-0.004	1.000
		ϕ_{13}	-0.054	0.300	-0.055	0.302	-0.011	1.000	-0.011	1.000
		ϕ_{23}	-0.002	0.876	0.003	0.874	-0.002	1.000	-0.002	1.000
	500	ϕ_{12}	0.000	1.000	0.000	1.000	0.000	1.000	0.000	1.000
		ϕ_{13}	-0.0032	0.520	-0.033	0.522	-0.007	1.000	-0.007	1.000
		ϕ_{23}	0.017	0.982	0.017	0.982	0.004	1.000	0.004	1.000

Table 5.11: Type I error rates for $R^2_{\gamma 7} = 0$ and with different covarince of exogenous latent variables and reliability of the latent indicators

	$\phi_{12} = 0.3$			$\phi_{12} = 0.6$
	$\phi_{13} = 0.1$			$\phi_{13} = 0.4$
	$\phi_{23} = 0.2$			$\phi_{23} = 0.5$
Reliability	N	Type I error	type I error	
0.64	50	0.029	0.034	
	100	0.064	0.068	
	250	0.070	0.070	
	500	0.066	0.082	
0.84	50	0.080	0.092	
	100	0.100	0.108	
	250	0.082	0.076	
	500	0.080	0.080	

Table 5.12: Power estimates for $R^2_{\gamma 7} = 0.05, 0.1$ with different covarince of exogenous latent variables and reliability of the latent indicators

	$\phi_{12} = 0.3$ $\phi_{13} = 0.1$ $\phi_{23} = 0.2$			$\phi_{12} = 0.6$ $\phi_{13} = 0.4$ $\phi_{23} = 0.5$	
		Power		Power	
Reliability	N	$R^2_{\gamma 7} = 0.05$	$R^2_{\gamma 7} = 0.1$	$R^2_{\gamma 7} = 0.05$	$R^2_{\gamma 7} = 0.1$
0.64	50	0.099	0.162	0.151	0.238
	100	0.344	0.526	0.468	0.666
	250	0.746	0.942	0.884	0.982
	500	0.972	1.000	0.996	1.000
0.84	50	0.302	0.507	0.414	0.612
	100	0.602	0.832	0.704	0.908
	250	0.940	0.994	0.982	1.000
	500	1.000	1.000	1.000	1.000

Simulation Syntax

```
TITLE: Monte Carlo simulation for three-way latent interaction
MONTECARLO: NAMES=x1-x9 y1-y3;
NOBSERVATIONS = 250;
NREPS = 500;
SEED = 12345;
ANALYSIS: ESTIMATOR = MLR;
TYPE = RANDOM;
ALGORITHM = INTEGRATION;
MODEL POPULATION:
[x1 − x9@0y1 − y3@0];
xi1 BY x1-x3@0.8;
xi2 BY x4-x6@0.8;
xi3 BY x7-x9@0.8;
eta BY y1-y3@0.8;
xi1@1;
xi2@1;
xi3@1;
eta@0.4;
D | xi1 XWITH xi2;
E | xi1 XWITH xi3;
F | xi2 XWITH xi3;
G | xi1 XWITH F;
eta ON xi1@0.3 xi2@0.4 xi3@0.5 D@0.1 E@0.2 F@0.2 G@0.3;
x1-x9@0.12; y1-y3@0.12;
xi1 WITH xi2@0.3 xi3@0.1;
xi2 WITH xi3@0.2;
MODEL:
```

$[x1 - x9 * 0y1 - y3 * 0]$;

xi1 BY x1-x3*0.8;

xi2 BY x4-x6*0.8;

xi3 BY x7-x9*0.8;

eta BY y1-y3*0.8;

xi1@1;

xi2@1;

xi3@1;

eta@0.4;

D | xi1 XWITH xi2;

E | xi1 XWITH xi3;

F | xi2 XWITH xi3;

G | xi1 XWITH F;

eta ON xi1*0.3 xi2*0.4 xi3*0.5 D*0.1 E*0.2 F*0.2 G*0.3;

x1-x9*0.12; y1-y3*0.12;

xi1 WITH xi2*0.3 xi3*0.1;

xi2 WITH xi3*0.2;

OUTPUT: TECH9;