INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road Ann Arbor MI 48106-1346 USA
313 761-4700  800 521-0600
Estimating and distinguishing between moderator and quadratic models in covariance structure modeling

Li, Fan, Ph.D.
The Ohio State University, 1992
ESTIMATING AND DISTINGUISHING BETWEEN
MODERATOR AND QUADRATIC MODELS IN
COVARIANCE STRUCTURE MODELING

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Fan Li, B.A., M.A.

* * * * *

The Ohio State University
1992

Dissertation Committee:

Robert C. MacCallum

Michael W. Browne

James T. Austin

Approved by

Robert MacCallum

Adviser

Department of Psychology
To My Parents
I would like to express deep appreciation to my adviser, Dr. Robert C. MacCallum, for his guidance, patience, and support throughout the research. Thanks also go to other members of my advisory committee, Dr. Michael W. Browne and James T. Austin, for their insightful comments and suggestions on this research. I would also like to thank Dr. Karl G. Jöreskog, Dr. Dag Sörbom, and Dr. Werner Wothke for making the testing versions of LISREL 8 and PRELIS 2 available for this project. To my husband, my parents, and my entire family, I would like to let you all know that I could not have done it without your love, encouragement, patience, and support.
VITA

August 11, 1960 ............. Born - Shanghai, P.R. China

1982 ........................ B.A., Psychology, East China Normal University, Shanghai, China

1985 ........................ M.A., Psychology, East China Normal University, Shanghai, China

1989 ........................ M.A., Social Psychology, The Ohio State University, Columbus, Ohio

PUBLICATIONS


FIELDS OF STUDY

Major Field: Psychology

Studies in Quantitative Psychology and Social Psychology
# TABLE OF CONTENTS

**DEDICATION** ........................................... ii  
**ACKNOWLEDGMENTS** ....................................... iii  
**VITA** ............................................... iv  
**LIST OF TABLES** ........................................ ix  
**LIST OF FIGURES** ....................................... xi  
**DISSERTATION ABSTRACT** .................................. xii  

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I. INTRODUCTION</strong></td>
<td></td>
</tr>
<tr>
<td>Linear Models in CSM</td>
<td>1</td>
</tr>
<tr>
<td>Nonlinear Relations in Psychological Research</td>
<td>3</td>
</tr>
<tr>
<td>Objectives of the Present Study</td>
<td>4</td>
</tr>
<tr>
<td><strong>II. LITERATURE REVIEW</strong></td>
<td></td>
</tr>
<tr>
<td>Estimating Nonlinear Models in CSM</td>
<td>6</td>
</tr>
<tr>
<td>Kenny and Judd’s Procedure for Assessing the Quadratic Models</td>
<td>8</td>
</tr>
<tr>
<td>Kenny and Judd’s Procedure for Assessing the Moderator Models</td>
<td>9</td>
</tr>
<tr>
<td>Concerns with Previous Studies</td>
<td>11</td>
</tr>
<tr>
<td><strong>III. STUDY ONE: ESTIMATING NONLINEAR MODELS WITH KENNY-JUDD PROCEDURES IN CSM</strong></td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>15</td>
</tr>
<tr>
<td>Design of the Study</td>
<td>16</td>
</tr>
<tr>
<td>Data Generation</td>
<td>17</td>
</tr>
<tr>
<td>Estimation Methods</td>
<td>24</td>
</tr>
<tr>
<td>Estimating the Quadratic Model</td>
<td>26</td>
</tr>
<tr>
<td>Estimating the Moderator Model</td>
<td>29</td>
</tr>
<tr>
<td>Results</td>
<td>31</td>
</tr>
<tr>
<td>Discussion</td>
<td>45</td>
</tr>
</tbody>
</table>
IV. STUDY TWO: ESTIMATING NONLINEAR MODELS WITH ALTERNATIVE PROCEDURES IN CSM

Introduction ................................................................. 47
Estimating the Quadratic Model ........................................ 48
Estimating the Moderator Model ....................................... 51
Method .......................................................... 54
Results ......................................................... 57
Discussion .................................................. 65

V. STUDY THREE: DISTINCTION BETWEEN NONLINEAR MODELS

Introduction ................................................................. 70
Distinguishing Between Quadratic and Moderator Models .......... 71
Method .......................................................... 73
Results ......................................................... 73
Discussion .................................................. 71

VI. ESTIMATING NONLINEAR MODELS WITH PROJECT TALENT DATA

Introduction ................................................................. 81
Project TALENT Data .................................................. 82
Estimating Nonlinear Models ....................................... 84
Distinguishing Nonlinear Models .................................. 85
Results ......................................................... 86
Discussion .................................................. 93

VII. GENERAL DISCUSSION ................................................. 95

APPENDICES

A. THE FORTRAN PROGRAM FOR GENERATING NORMAL DATA ........ 97
B. THE FORTRAN PROGRAM FOR GENERATING NON-NORMAL DATA .. 101
C. PRELIS INPUT .................................................. 106
D. LISREL 8 INPUT FILE FOR ESTIMATING THE QUADRATIC EFFECT WITH KENNY-JUDD APPROACH ... 108
E. LISREL 8 INPUT FILE FOR ESTIMATING THE MODERATOR EFFECT WITH KENNY-JUDD APPROACH ... 110

vii
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>$A_x$ Loading Matrix for the Quadratic Effects of $\xi_1^2$</td>
<td>28</td>
</tr>
<tr>
<td>2.</td>
<td>$A_x$ Loading Matrix for the Moderator Effects of $\xi_1\xi_2$</td>
<td>30</td>
</tr>
<tr>
<td>3.</td>
<td>Descriptive Results for the LV's of Normal Data (N=1000) and Non-normal Data (N=1000)</td>
<td>32</td>
</tr>
<tr>
<td>4.</td>
<td>Descriptive Results for the MV's of Normal Data (N=1000) and Non-normal Data (N=1000)</td>
<td>34</td>
</tr>
<tr>
<td>5.</td>
<td>Covariance Matrix of MV's Used in Estimating $\xi_1\xi_2$ Effect for Normal Data</td>
<td>36</td>
</tr>
<tr>
<td>6.</td>
<td>Covariance Matrix of MV's Used in Estimating $\xi_1^2$ Effect for Normal Data</td>
<td>36</td>
</tr>
<tr>
<td>7.</td>
<td>Covariance Matrix of MV's Used in Estimating $\xi_1\xi_2$ Effect for Non-normal Data</td>
<td>37</td>
</tr>
<tr>
<td>8.</td>
<td>Covariance Matrix of MV's Used in Estimating $\xi_1^2$ Effect for Non-normal Data</td>
<td>38</td>
</tr>
<tr>
<td>9.</td>
<td>True Values and Estimates by Different Methods -- Quadratic Data and Quadratic Model for Normal and Non-normal Data</td>
<td>39</td>
</tr>
<tr>
<td>10.</td>
<td>True Values and Estimates by Different Methods -- Moderator Data and Moderator Model for Normal and Non-normal Data</td>
<td>44</td>
</tr>
<tr>
<td>11.</td>
<td>Descriptive Results for the LV's (N=100,000)</td>
<td>58</td>
</tr>
</tbody>
</table>
12. Covariance Matrix for the LV's .......... 59
13. Descriptive Results for the MV's  
   (N=100,000) ............................. 60
14. The Sample Covariance Matrix of MV's .... 61
15. The Expected Covariance Matrix of MV's  
    in Population .......................... 61
16. True Values and Estimates -- Quadratic Data  
    and Quadratic Model .................... 66
17. True Values and Estimates -- Moderator Data  
    and Moderator Model .................... 68
18. Estimates by GLS -- Comparison of Moderator  
    and Quadratic Model on the Quadratic Data  
    (N=100,000) ............................. 75
19. Estimates by GLS -- Comparison of Moderator  
    and Quadratic Model on the Moderator Data  
    (N=100,000) ............................. 76
20. The Summary of Changes in Estimates ........ 79
21. Descriptive Results of the Project TALENT  
    Data (N=3323) .......................... 86
22. Covariance Matrix for the MV's of the  
    Project TALENT Data .................... 87
23. Results for the Linear Model on the  
    Project TALENT Data .................... 88
24. Results for the Project TALENT Data ........ 91
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Comparison of data matrices used in Kenny-Judd's procedures and the new procedures</td>
<td>48</td>
</tr>
</tbody>
</table>
ESTIMATING AND DISTINGUISHING BETWEEN
MODERATOR AND QUADRATIC MODELS IN
COVARIANCE STRUCTURE MODELING

By
Fan Li, Ph.D.
The Ohio State University, 1992
Professor Robert C. MacCallum, Adviser

This dissertation investigated the issue of estimating nonlinear effects of latent variables in covariance structure modeling. In particular, this dissertation studied the effect of distribution assumptions on Kenny and Judd's procedures (1984). This dissertation also proposed simple alternatives to the Kenny and Judd procedures of estimating quadratic and moderator effects of latent variables in covariance structure modeling. Using a Monte Carlo approach, it was demonstrated that nonlinear effects could be estimated with proper constraints on the covariance structure and mean structure of latent variables and that no additional created indicators were necessary. Results also
demonstrated that the quadratic model and the moderator model are equivalent models for the alternative procedures. Thus, the goodness of fit indices and the squared multiple correlation for the structural equations are not informative for distinguishing between the quadratic model and the moderator model. Analysis on Project TALENT Data showed that the procedures were easy to apply. In addition, the assumptions and appropriateness of different estimation methods in the context of estimating nonlinear effects of LV's were discussed.
CHAPTER I
INTRODUCTION

Linear Models in Covariance Structure Modeling

Covariance structure modeling (CSM) is widely used in social sciences because it allows modeling relationships among measured variables (MV's) and latent variables (LV's). However, its applications are often limited because of the assumption that MV's are linear functions of LV's, and dependent (endogenous) LV's are linear functions of independent (exogenous) and other dependent LV's (Bentler & Weeks, 1980; Jöreskog & Sörbom, 1989; McArdle & McDonald, 1984).

Based on the LISREL framework (Jöreskog & Sörbom, 1989), a complete covariance structure model consists of two parts. The measurement model represents relations between MV's and LV's, whereas the structural model represents relations among dependent and independent LV's. In addition, MV's are divided into two types: y's being indicators of dependent LV's and x's being indicators of independent LV's. The LISREL model is given as (Jöreskog & Sörbom, 1989):

1
\[ y = \Lambda_y \eta + \epsilon, \quad (1) \]

\[ x = \Lambda_x \xi + \delta, \quad (2) \]

\[ \eta = B\eta + \Gamma \xi + \zeta, \quad (3) \]

with the assumptions of \( \zeta, \epsilon, \) and \( \delta \) being mutually uncorrelated, \( \xi \) being uncorrelated with \( \xi \), \( \epsilon \) being uncorrelated with \( \eta \), \( \delta \) being uncorrelated with \( \xi \), and \( I - B \) being non-singular,

where

\( y = (p \times 1) \) observed MV's, indicators of dependent LV's;

\( x = (q \times 1) \) observed MV's, indicators of independent LV's;

\( \Lambda_y = (p \times m) \) coefficient matrix, showing relation of \( y \) to \( \eta \);

\( \Lambda_x = (q \times n) \) coefficient matrix, showing relation of \( x \) to \( \xi \);

\( \eta = (m \times 1) \) dependent LV's;

\( \xi = (n \times 1) \) independent LV's;

\( B = (m \times m) \) coefficient matrix with zeroes on diagonal, showing relation among \( \eta \)'s;

\( \Gamma = (m \times n) \) coefficient matrix, showing relation of \( \eta \) to \( \xi \);

\( \epsilon = (p \times 1) \) errors of measurement for \( y \);

\( \delta = (q \times 1) \) errors of measurement for \( x \);

\( \zeta = (m \times 1) \) residual terms in the dependent LV's.

Equations (1) - (2) complete the measurement model, in which Equation (1) specifies linear relations between dependent LV's and their indicators, and Equation (2) indicates linear relations between independent LV's and
their indicators. Equation (3) represents the structural model, showing linear relationships among dependent LV's and relationships between dependent LV's and independent LV's. Let $\Phi$ ($n \times n$) and $\Psi$ ($n \times n$) be the covariance matrices of $\xi$ and $\zeta$ respectively, and let $\Theta_e$ and $\Theta_\delta$ be the covariance matrices of $\epsilon$ and $\delta$ respectively. From the assumptions, the covariance matrix of $z = (y', x')'$ can be derived as:

$$
\Sigma = \begin{pmatrix}
\Lambda_y (I-B)^{-1}(\Gamma \Phi \Gamma' + \Psi) (I-B')^{-1} \Lambda_y' + \Theta_e & \Lambda_y (I-B)^{-1} \Gamma \Phi \Lambda_x' \\
\Lambda_x \Phi \Gamma' (I-B')^{-1} \Lambda_y' & \Lambda_x \Phi \Lambda_x' + \Theta_\delta
\end{pmatrix}.
$$

Clearly, the LISREL model represents linear relations between MV's and LV's and linear relations among the latent variables.

**Nonlinear Relations in Psychological Research**

In psychology, however, nonlinear relations among variables are very common. One type of common nonlinear relationship is the moderator effect, in which a variable interacts with another so as to enhance predictability of a criterion (Cohen & Cohen, 1983). For instance, Atkinson (1964) proposed a model for the prediction of behavior, in which a behavioral response ($B$) is multiplicatively related to drive strength ($D$), habit strength ($H$), the incentive value of the anticipated reinforcement ($R$), and the intensity of the stimulus ($S$).
Another type of common nonlinear relation among variables is the quadratic effect (Cohen & Cohen, 1983), where a dependent variable is a function of the first- and second-order trends of an independent variable. For example, this type of function can represent the relationship between immediate recall and serial position of stimuli (Craik, 1970).

As for latent variables with multiple indicators, their relationships may also be realistically represented as moderator or quadratic effects in theory. Unfortunately, these nonlinear models cannot be studied directly by the LISREL model expressed by Equations (1) - (3). Researchers in sociology and psychology have investigated moderator and quadratic effects of latent variables in covariance structural modeling (Stolzenberg, 1979; Kenny & Judd, 1984; Hayduk, 1987). However, the procedures developed to study these issues tend to be very complicated and not easily applied in practice.

**Objectives of the Dissertation**

This dissertation had three general objectives. The first goal was to study the effect of distribution assumptions such as normality and non-zero means on the modeling of nonlinear effects of LV's. The second goal was to propose simple alternative procedures for estimating moderator effects and quadratic effects of LV's. Lastly,
the present study investigated whether and how one could distinguish between moderator and nonlinear models for LV's for the same data. These issues were first studied in a Monte Carlo approach and empirical data were used for subsequent demonstrations.
CHAPTER II
LITERATURE REVIEW

Estimating Moderator and Quadratic Models in CSM

Although the issue of nonlinear models has not been an unfamiliar one in regression analysis (Althauser, 1971; Cohen & Cohen, 1983), nonlinear models in CSM have been relatively neglected. Before Kenny and Judd (1984) published their paper in the Psychological Bulletin, sociologists and econometricians had paid attention to some aspects of nonlinear models in path analysis and structural equation modeling (Stolzenberg, 1979; Mariano, 1983; Marsden, 1983).

Stolzenberg (1979) focused on interpretations of causal effects in different nonlinear and nonadditive models in path analysis. Stolzenberg suggested that the partial derivative, \( \partial Y/\partial X \), which is the rate at which dependent variable \( Y \) changes per change in independent variable \( X \), should be used to explain the effects. For example, if a quadratic effect is involved,

\[
Y = a + b_1X + b_2X^2 + \sum_{j=1}^{r} c_jZ_j + \epsilon, \tag{5}
\]

6
where $a$, $b$'s, and $c$'s are parameters, $X$, $Y$, $Z$'s are variables, and $\epsilon$ is the error, the partial derivative, $\frac{\partial Y}{\partial X} = b_1 + 2b_2X$, shows that the effect of $X$ on $Y$ can be expressed as a linear function of $X$.

Marsden (1983) investigated cases where interactions are between one unobserved block variable and one measured variable, and between two block variables. A block variable is a special type of unobserved variable with empirical meaning and it is defined as a linear function of observed indicators (Marsden, 1983). For an interaction effect that involves one block variable and one measured variable, Marsden showed that both nonlinear regression and LISREL could be used in the estimation of such effects. On the other hand, Marsden claimed that, when the interaction effect involves two block variables, only nonlinear regression can be used for estimation, because LISREL V did not allow for the setting of certain kinds of necessary constraints.

Clearly, although previous studies have attempted to generalize CSM into nonlinear models, none of them offered an overall framework in terms of estimation of the effects. Thus, Kenny and Judd's (1984) paper was the first influential one on this issue. They developed procedures to estimate parameters for models including moderator and quadratic effects of LV's using the computer program COSAN (Fraser, 1980). Their procedures involve including
additional LV's and corresponding indicators and setting constraints on parameters and variances and covariances of LV's, which can be accomplished in COSAN.

**Kenny and Judd's procedure for assessing quadratic models.** To illustrate their procedure for assessing the quadratic effects of LV's, consider a case with one independent LV, \( \xi \), and two corresponding indicators, \( X_1 \) and \( X_2 \), and one dependent variable, \( Y \), as a single indicator of the dependent LV, \( \eta \), where

\[
Y = a \xi + b \xi^2 + \zeta, \tag{6}
\]

\[
X_1 = \xi + \delta_1, \tag{7}
\]

\[
X_2 = f \xi + \delta_2, \tag{8}
\]

with \( \delta_1, \delta_2, \zeta, \) and \( \xi \) all uncorrelated, and \( E(\delta_1) = E(\delta_2) = E(\zeta) = E(\xi) = 0 \). Coefficients \( a, b, \) and \( f \) are parameters being estimated. Indicators, \( X_1^2, X_2^2, \) and \( X_1X_2 \), are used to estimate the effects of \( \xi^2 \) with the following relations:

\[
X_1^2 = \xi^2 + 2 \xi \delta_1 + \delta_1^2, \tag{9}
\]

\[
X_2^2 = f^2 \xi^2 + 2 f \xi \delta_2 + \delta_2^2, \tag{10}
\]

\[
X_1X_2 = f \xi^2 + f \xi \delta_1 + \xi \delta_2 + \delta_1 \delta_2. \tag{11}
\]

Thus, totally there are 6 indicators (\( X_1, X_2, X_1^2, X_2^2, X_1X_2, \) and \( Y \)) and 10 LV's (\( \xi, \xi^2, \delta_1, \delta_2, \delta_1^2, \delta_2^2, \xi \delta_1, \xi \delta_2, \delta_1 \delta_2 \) and \( \zeta \)). Because of the assumption that \( \xi, \delta_1, \) and \( \delta_2 \) are
normally distributed with mean zero, variances of $\xi^2$, $\delta_1^2$, $\delta_2^2$, $\xi\delta_1$, $\xi\delta_2$, and $\delta_1\delta_2$ can be derived as:

\[
\text{Var}(\xi^2) = 2(\text{Var}(\xi))^2,
\]

\[
\text{Var}(\delta_1^2) = 2(\text{Var}(\delta_1))^2,
\]

\[
\text{Var}(\delta_2^2) = 2(\text{Var}(\delta_2))^2,
\]

\[
\text{Var}(\xi\delta_1) = \text{Var}(\xi)\text{Var}(\delta_1),
\]

\[
\text{Var}(\xi\delta_2) = \text{Var}(\xi)\text{Var}(\delta_2),
\]

\[
\text{Var}(\delta_1\delta_2) = \text{Var}(\delta_1)\text{Var}(\delta_2).
\]

With these specific constraints on the coefficients, Kenny and Judd (1984) were able to estimate the parameters $a$, $b$, $f$, and residual variances, and to evaluate goodness of fit of the model, using the generalized least squares (GLS) estimation method in COSAN (Fraser, 1980).

**Kenny and Judd’s procedure for assessing moderator models.** Consider two LV’s ($\xi_1$ and $\xi_2$) where each LV has two indicators ($X_1$ and $X_2$ for $\xi_1$, $X_3$ and $X_4$ for $\xi_2$), with relations:

\[
Y = c\xi_1 + d\xi_2 + e\xi_1\xi_2 + \zeta,
\]

\[
X_1 = \xi_1 + \delta_1,
\]
\begin{align}
X_2 &= g \xi_1 + \delta_2 , \\
X_3 &= \xi_2 + \delta_3 , \\
X_4 &= h \xi_2 + \delta_4 .
\end{align}

The indicators of the product latent variable \( \xi_1 \xi_2 \) are:

\begin{align}
X_1X_3 &= \xi_1 \xi_2 + \xi_1 \delta_3 + \xi_2 \delta_1 + \delta_1 \delta_3 , \\
X_1X_4 &= h \xi_1 \xi_2 + \xi_1 \delta_4 + h \xi_2 \delta_1 + \delta_1 \delta_4 , \\
X_2X_3 &= g \xi_1 \xi_2 + g \xi_1 \delta_3 + \xi_2 \delta_2 + \delta_2 \delta_3 , \\
X_2X_4 &= g h \xi_1 \xi_2 + g \xi_1 \delta_4 + h \xi_2 \delta_2 + \delta_2 \delta_4 .
\end{align}

Therefore, there are a total of 9 MV’s \( X_1, X_2, X_3, X_4, X_1X_3, X_1X_4, X_2X_3, X_2X_4, \) and \( Y \), and 16 LV’s \( (\xi_1, \xi_2, \xi_1 \xi_2, \xi_1 \delta_3, \xi_1 \delta_4, \xi_2 \delta_1, \xi_2 \delta_2, \delta_1, \delta_2, \delta_3, \delta_4, \delta_1 \delta_3, \delta_1 \delta_4, \delta_2 \delta_3, \delta_2 \delta_4 \) and \( \xi \). Based on the assumption that \( \xi_1, \xi_2, \delta_1, \delta_2, \delta_3, \delta_4, \) and \( \xi \) are multivariate normal with mean zero and mutually uncorrelated with the exception of \( \xi_1 \) and \( \xi_2 \), variances of \( \xi_1 \xi_2, \xi_1 \delta_3, \xi_1 \delta_4, \xi_2 \delta_1, \xi_2 \delta_2, \delta_1 \delta_3, \delta_1 \delta_4, \delta_2 \delta_3, \) and \( \delta_2 \delta_4 \) can be derived as:

\begin{align}
\text{Var}(\xi_1 \xi_2) &= \text{Var}(\xi_1) \text{Var}(\xi_2) + (\text{Cov}(\xi_1 \xi_2))^2 , \quad (27) \\
\text{Var}(\delta_1 \delta_3) &= \text{Var}(\delta_1) \text{Var}(\delta_3) , \quad (28) \\
\text{Var}(\delta_1 \delta_4) &= \text{Var}(\delta_1) \text{Var}(\delta_4) , \quad (29) \\
\text{Var}(\delta_2 \delta_3) &= \text{Var}(\delta_2) \text{Var}(\delta_3) . \quad (30)
\end{align}
\[ \text{Var}(\delta_2 \delta_4) = \text{Var}(\delta_2) \text{Var}(\delta_4), \] (31)

\[ \text{Var}(\xi_1 \delta_3) = \text{Var}(\xi_1) \text{Var}(\delta_3), \] (32)

\[ \text{Var}(\xi_1 \delta_4) = \text{Var}(\xi_1) \text{Var}(\delta_4), \] (33)

\[ \text{Var}(\xi_2 \delta_1) = \text{Var}(\xi_2) \text{Var}(\delta_1), \] (34)

\[ \text{Var}(\xi_2 \delta_2) = \text{Var}(\xi_2) \text{Var}(\delta_2). \] (35)

Again, Kenny and Judd (1984) used the GLS method in COSAN to estimate the parameters, \( c, d, e, g, \) and \( h, \) and residual variances, with constraints on the coefficients.

Hayduk (1987) showed that the Kenny-Judd procedure could also be applied with LISREL VI by creating "convenience variables" that mimic the complex variables, e.g. \( \xi \delta_1, \xi \delta_2, \delta_1 \delta_2, \delta_1^2, \) and \( \delta_2^2 \) in Equations (9)–(11), then constraining the convenience variable to have effects on \( X_1^2, \)

\( X_2^2, X_1 X_2, \) and setting equality constraints.

**Concerns with Previous Studies**

Although the procedures that Kenny and Judd (1984) developed make possible modeling the moderator and quadratic effects, there are several concerns with their approaches. First, Kenny and Judd's procedures were based on the assumption that LV's have zero means, e.g. \( E(\xi_1) = E(\xi_2) = 0. \)

The relations defined by Equations (12) – (17) and Equations (27) – (35) are true only when these latent variables have
zero means. Although this assumption simplified the relations between the latent variables, e.g. $\text{Cov}(\xi_1, \xi_1^2) = 0$ and $\text{Var}(\xi_1^2) = 2*(\text{Var}(\xi_1))^2$, it may not be a realistic and practical assumption. In psychological research, variables often have meaningful means. Therefore, it is desirable to have a more general procedure by including mean structure and represent fully the relations among variables in a model. In addition, including mean structure may also help to avoid identification problems during estimation.

Secondly, in Kenny and Judd’s (1984) paper, the authors recognized that squares ($\xi_1^2$) or products ($\xi_1 \xi_2$) of LV’s are not normally distributed even if the individual LV’s, $\xi_1$ and $\xi_2$, have normal distributions. They suggested the use of GLS estimation instead of ML estimation as an alternative because estimating parameters by minimizing a maximum likelihood loss function assumes multivariate normality of the latent variables. However, the GLS method that they suggested was the normal theory GLS defined as:

$$F = \frac{1}{2} \text{tr}[(I - S^{-1}\Sigma)^2],$$  \hfill (36)

where sum of squares weighted by the inverse of the sample covariance matrix $S$ is minimized. The normal theory GLS has been shown to be inappropriate for non-normal multivariate distributions as well (Browne, 1984). Therefore it may be more appropriate to use the unweighted least squares (ULS)
method or the asymptotically distribution free (ADF) method, which does not require the normality assumption.

Thirdly, inclusion of indicators for the squared term and the product term of LV’s, $X_1^2$, $X_2^2$, $X_1X_2$, $X_1X_3$, $X_1X_4$, $X_2X_3$, and $X_2X_4$, does not add any new information to the model. Instead, these created indicators may cause problems in the estimation process. One of the problems is that the constructed indicators may induce a high degree of multicollinearity among indicators, which may result in the asymptotic covariance matrix, used in the ADF estimation method, being nearly singular. In addition, when the number of MV’s or LV’s is increased, more created indicators and more constraints have to be included. Thus, it is desirable to have a procedure that can estimate the same number of parameters with only the first-order MV’s to avoid the unnecessary complications caused by the inclusion of created indicators.

Lastly, although Hayduk demonstrated that the Kenny and Judd (1984) method could also be applied with the LISREL program, he used the ML method for parameter estimation and his procedure was extremely complicated. Thus, in practice, models with nonlinear and interactive effects of latent variables are rarely studied, and apparently have not been studied at all using an estimation method appropriate for non-normal variables.
For example, Tetrick and LaRocco (1987) applied the Kenny-Judd procedure to moderator effects of understanding, prediction, and control on the relationships between perceived stress, job satisfaction, and psychological well-being. The ML estimation method was used because "none of the observed variables' skewnesses were greater than 1.0" (Tetrick & LaRocco, 1987, p. 540). In fact, when a product term is included in the model, the issue of non-normality still exists even if the observed variables are normal. In addition, Tetrick and LaRocco (1987) did not present detailed information about their procedures for the moderator effects and estimates of parameters for the moderator effects. The lack of detailed information makes it difficult to verify and evaluate their procedures and results.
CHAPTER III
STUDY ONE
ESTIMATING NON-LINEAR MODELS
WITH KENNY-JUDD PROCEDURES IN CSM

Introduction

The procedures developed by Kenny and Judd (1984) were used to estimate moderator and quadratic effects of LV's in CSM. The data analyzed were computer generated with known parameters. The goal of this study was to apply Kenny and Judd's approach and to investigate the effect of the normality assumption on estimates provided by the maximum likelihood estimation method (ML), GLS, ULS, and ADF in terms of accuracy and efficiency.

Because Kenny and Judd's procedures involve a squared term of a variable or product terms of two variables, violation of the normality assumption is inevitable. When the first-order variables are normal, the squared terms or the product terms are not normally distributed. When the first-order MV's are not normal, this violation will be more severe. Thus, corresponding tests of hypotheses can give misleading results, and parameter estimates can have sub-optimal properties for ML and GLS because their usages
assumes multivariate normality among variables. Because the ULS and ADF methods have more relaxed assumptions on the distribution, they may recover the true parameters better than ML and GLS methods do.

Design of the Study

Artificial data with moderator and quadratic effects were generated and a corresponding moderator model or quadratic model was fitted to each data set. Different estimates were compared for the accuracy of parameters being recovered and overall goodness of fit of solutions to the data. These comparisons were studied for artificial data generated both from normal distributions and from non-normal distributions.

Study One was a 2 (types of distribution) × 2 (types of effects) design, and there were four sets of artificial data: two for normal first-order MV's (i.e. \(X_1, X_2, X_3\) and \(X_4\)) and two for slightly non-normal first-order MV's. For data from each distribution, two sets of data were created: one with moderator effects and the other with quadratic effects of LV's. Both the moderator model and the quadratic model were fitted to the data sets with the corresponding effects. Estimation methods ML, GLS, ULS, and ADF were used to estimate parameters for these data sets and their estimates were compared to investigate the accuracy of each method in terms of recovering true parameters and goodness of fit.
For data with non-normal first-order MV's, Kenny and Judd's procedures might not provide good estimates for some estimation methods because the relationships between variances and covariances presented in Equations (12) - (17) and Equations (27) - (35) would not hold (Kendall & Stuart, 1963). That is, the Kenny-Judd approach might work better with an estimation method that does not assume normality of latent variables. In the case where the first-order MV's are slightly non-normal, all methods might provide worse estimation than they do when the first-order MV's are normal.

Data Generation

In the present study, there were two independent LV's ($\xi_1$ and $\xi_2$), each of which had two indicators ($X_1$ and $X_2$ for $\xi_1$; $X_3$ and $X_4$ for $\xi_2$). Similar to Kenny and Judd's study (Kenny & Judd, 1984), latent variables ($\xi_1$ and $\xi_2$) and values for the error terms ($\delta_1$, $\delta_2$, $\delta_3$, $\delta_4$, and $\epsilon$) were generated first. Then four indicators of independent LV's, $X_1$, $X_2$, $X_3$, and $X_4$, and two dependent MV's, $Y_{\text{moderator}}$ (or $Y_m$) and $Y_{\text{quadratic}}$ (or $Y_q$) were derived. Finally, LISREL 8 was used to estimate the nonlinear effects of LV's.

Kaiser and Dickman Procedure. The Kaiser and Dickman procedure (Kaiser & Dickman, 1962) was used to generate variables $\xi_1$ and $\xi_2$ because this procedure allows us to
generate a matrix of scores with specified distribution and, at the same time, maintain specified correlations among the variables.

The Kaiser and Dickman procedure is based on the fundamental postulate of component analysis,

\[ Z = X F' \],

where \( F \) \((n \times n)\) is any factoring of \( R \) \((n \times n)\), a given correlation matrix and \( X \) \((N \times n)\) is a score matrix on the components represented in \( F \). \( F \) is computed from the following equation (Johnson & Wichern, 1988):

\[ F = E V^{1/2} \],

where \( E \) \((n \times n)\) contains the eigenvectors of \( R \) \((n \times n)\) and \( V \) \((n \times n)\) is a diagonal matrix with the corresponding eigenvalues on diagonal. Thus, given \( n \) variables for \( N \) cases and a positive definite correlation matrix \( R \) \((n \times n)\), a data matrix \( Z \) \((N \times n)\) can be obtained with specified relationship among the variables which are represented by \( R \).

For instance, when an arbitrary \( X \) \((N \times n)\), sampling randomly from an uncorrelated population with \( N(0,1) \), is postmultiplied with a \( F'(n \times n) \) factored from a given \( R \) \((n \times n)\), the resulting matrix, \( Z \) \((N \times n)\), contains observations from a multivariate population with zero means, unit standard deviations and population correlations \( R \) among variables. However, observed correlations of variables in \( Z \) will not exactly correspond to \( R \) due to sampling error.
Generating Normal First-Order MV's. For the present study, there were seven latent variables to be generated: \( \xi_1, \xi_2, \delta_1, \delta_2, \delta_3, \delta_4, \) and \( \epsilon. \) Among them, \( \xi_1 \) and \( \xi_2 \) were correlated with each other and \( \delta_1, \delta_2, \delta_3, \delta_4, \) and \( \epsilon \) were uncorrelated. Thus, the Kaiser and Dickman procedure was used to generate \( \xi_1 \) and \( \xi_2 \) only. The sample size was 1000.

First, eigenvalues and eigenvectors were computed from the given correlation matrix \( R \) (2 x 2) with 0.5 being the correlation coefficient to obtain the loading matrix, \( F \) (2 x 2). Secondly, a matrix \( X \) (1000 x 2) was generated, containing random numbers drawn independently from two normal distributions with zero means and variances of 1.0. Lastly, the matrix \( F'(2 \times 2) \) was postmultiplied by the random number matrix \( X \) (1000 x 2) and, therefore, \( Z \) (1000 x 2) was computed, which contained scores on \( \xi_1 \) and \( \xi_2 \) for 1000 cases.

The remaining latent variables, \( \delta_1, \delta_2, \delta_3, \delta_4, \) and \( \epsilon, \) were created by a random number generator. They were drawn independently from normal distributions with zero means and variances of 0.3.

The measured variables were derived from these seven latent variables defined by the following equations:

\[
X_1 = \xi_1 + \delta_1, \tag{39}
\]
\[ X_2 = 0.6 \xi_1 + \delta_2, \quad (40) \]
\[ X_3 = \xi_2 + \delta_3, \quad (41) \]
\[ X_4 = 0.6 \xi_2 + \delta_4, \quad (42) \]
\[ Y_{\text{moderator}} = 0.8 \xi_1 + 0.8 \xi_2 + 0.5 \xi_1 \xi_2 + \epsilon, \quad (43) \]
\[ Y_{\text{quadratic}} = 0.8 \xi_1 + 0.8 \xi_2 + 0.5 \xi_1^2 + \epsilon. \quad (44) \]

\( X_1, X_2, X_3, \) and \( X_4 \) were also from normal distributions because \( \xi_1 \) and \( \xi_2 \) were from normal distributions and all first-order MV's were linear functions of \( \xi_1 \) and \( \xi_2 \).

Thus, to generate data for the present study, a computer program was written in Microsoft FORTRAN to follow the procedures described previously (see Appendix A for the source codes). To summarize the process of data generation, the program computed the eigenvalues and eigenvector first for a given correlation matrix \( R \) (2 \times 2). Then the program generated two uncorrelated random variables from normal distributions with zero means and unit variances. By applying Equation (37), a matrix \( Z \) (1000 \times 2) was obtained, which contained values on \( \xi_1 \) and \( \xi_2 \) for 1000 cases. After generating five more variables as \( \delta_1, \delta_2, \delta_3, \delta_4, \) and \( \epsilon \) independently from normal distributions with zero means and variances of 0.3, the computer program also derived MV's: \( X_1, X_2, X_3, X_4, Y_{\text{moderator}}, \) and \( Y_{\text{quadratic}} \). Finally the program
computed product and squared terms $X_1X_2, X_1^2, X_2^2, X_1X_3, X_1X_4,$ $X_2X_3,$ and $X_2X_4,$ which were used for estimation of non-linear effects.

**Generating Non-Normal First-Order MV's.** The procedures used to generate normal first-order MV's could also be used to generate multivariate non-normal first-order MV's with two exceptions. First, the loading matrix $F (n \times n)$ was obtained based on an intermediate correlation matrix $R' (n \times n)$. Secondly, variables in the resulting data matrix $Z (N \times n)$ were transformed with a power function.

Fleishman (1978) developed a simple procedure to generate univariate non-normal variables. This method uses a polynomial transformation of the form:

$$Y = a + bX + cX^2 + dX^3,$$

or

$$Y = ((dX + c)X + b)X + a.$$  \hspace{1cm} (45)

Fleishman (1978) showed that, with mean=0 (c=-a) and variance=1.0, skew and kurtosis of a distribution could be expressed as functions of $b$, $c$, and $d$:

$$skew = 2c \left( b^2 + 24bd + 105d^2 + 2 \right),$$

and kurtosis=$$
24(bd+c^2[1+b^2+28bd]+d^2[12+48bd+141c^2+225d^2]).$$

Fleishman (1978) also presented a table of $b$'s, $c$'s, and $d$'s for skewness less than or equal to 1.75 and kurtosis of less
than or equal to 3.75. For example, to obtain data with mean = 0, variance = 1.0, skew = 0.25, and kurtosis = 2.5, coefficients: \( a = -0.0299 \), \( b = 0.8108 \), \( c = 0.0299 \), and \( d = 0.0593 \) should be used (cf. Table 1 of Fleishman, 1978).

Vale and Maurelli (1983) extended Fleishman's procedure to generate multivariate random numbers with specified intercorrelations and univariate means, variances, skews, and kurtoses. The method they suggested is a combination of the Kaiser-Dickman procedure and the Fleishman procedure. That is, multivariate normal random numbers with specified intercorrelations are generated and then univariately transformed to the desired shapes. However, instead of using the desired correlation \( R (n \times n) \) for the Kaiser-Dickman procedure, Vale and Maurelli (1983) pointed out that an intermediate correlation matrix \( R' (n \times n) \) should be used in order to generate non-normal variables with the desired intercorrelations. The intermediate correlation between two variables \( X_1 \) and \( X_2 \) is determined by solving the following polynomial:

\[
\rho_{Y_1Y_2} = \rho_{X_1X_2}(b_1b_2 + 3b_1d_2 + 3d_1b_2 + 9d_1d_2) + \rho_{X_1X_2}^2(2c_1c_2) + \rho_{X_1X_2}^3(6d_1d_2),
\]

where \( Y \)'s are transformed non-normal variables, \( b \)'s, \( c \)'s, and \( d \)'s are the coefficients in the power functions, \( \rho_{Y_1Y_2} \) is the desired post-transformation correlation, \( \rho_{X_1X_2} \) is the
intermediate correlation between two normal variables, and the subscripts indicate the number of variables.

Thus, for the desired correlation $r = 0.50$, an intermediate correlation $\rho = 0.51$ was obtained using Equation (49). To generate multivariate normal variables with 0.5 correlation, the computer program (see Appendix B) provided eigenvalues and eigenvectors for the intermediate correlation 0.51. Then the program generated two uncorrelated random variables from normal distributions with zero means and unit variances. By applying Equation (37), a matrix $Z$ ($1000 \times 2$) was obtained. Finally, in order to generate the multi-vari­ate non-normal variables with mean $= 0$, variance $= 1.0$, skew $= 0.25$, and kurtosis $= 2.5$, coefficients: $a = -0.0299$, $b = 0.8108$, $c = 0.0299$, and $d = 0.0593$ were used for the transformation of the both variables. The resulting matrix contained values on non-normal $\xi_1$ and $\xi_2$ for 1000 cases.

The remaining latent variables, $\delta_1$, $\delta_2$, $\delta_3$, $\delta_4$, and $\epsilon$, were also created by a random number generator. They were still drawn independently from normal distributions with zero means and variances of 0.3. The measured variables were again derived from these seven latent variables defined by the Equations (39) - (44). As linear functions of $\xi_1$ and $\xi_2$, $X_1$, $X_2$, $X_3$, and $X_4$ were also non-normal distributed because $\xi_1$ and $\xi_2$ were from non-normal distributions with skewnesses being .25 and kurtoses being 2.5.
Before LISREL 8 was used to estimate the non-linear effects of the latent variables, the computer generated data were analyzed by PRELIS 2 in order to obtain descriptive statistics (see Appendix C for PRELIS input). PRELIS is a computer program for multivariate data screening and data summarization and is often used as a preprocessor for the LISREL program (Jöreskog & Sörbom, 1988). Results will be presented showing that the data generation procedure produced data with the desired properties.

**Estimation Methods**

Estimation methods ML, GLS, ULS, and ADF were used for the present study. One of assumptions of using ML and GLS is multivariate normality of variables, while ULS and ADF make no such distribution assumption. Each method minimizes its own fitting function and fits $\Sigma$, a covariance matrix implied by the model, to the sample covariance matrix $S$. The fitting function for ULS is

$$F_{ULS} = \frac{1}{2} \text{tr}[(S - \Sigma)^2], \quad (50)$$

i.e., half the sum of squares of all the elements of the residual matrix $(S - \Sigma)$.

The fitting function for ML is

$$F_{ML} = \log|\Sigma| + \text{tr}(S\Sigma^{-1}) - \log|S| - (p + q), \quad (51)$$

where $q$ is the number of indicators of independent LV's and $p$ is the number of indicators of dependent LV's. The ML
method is asymptotically equivalent to normal theory GLS (Browne, 1974).

The fitting function for the normal theory GLS is defined as:

\[ F = \frac{1}{2} \text{tr}[(I - S^{-1} \Sigma)^2], \]  

where a sum of squares weighted by the inverse of the sample covariance matrix \( S \) is minimized.

Browne (1984) showed that all these fitting functions are special cases of a general family of fitting functions, which may be written as:

\[ F(\theta) = (s-\sigma)' \tilde{W}^{-1} (s-\sigma), \]  

where \( \tilde{W} \) is a positive definite weight matrix. For example, when \( \tilde{W} = S \), where \( S \) is the sample covariance matrix, Equation (53) is equivalent to Equation (52). An assumption for the ML and GLS fitting functions is multivariate normality. Thus, when the assumption is violated, corresponding tests of hypotheses can give misleading results, and parameter estimates can have sub-optimal properties. However, the ADF method that Browne (1984) proposed requires a less restrictive assumption of a distribution from the elliptical class. Let \( x_r = (x_{1r}, x_{2r}, \ldots, x_{nr}, \ldots, x_{pr})' \) represent the \( r \)th observation, \( r=1, \ldots, N, \) on \( x \) and let
\[ \bar{X}_i = N^{-1} \sum_{r=1}^{N} X_{ir} , \quad (54) \]

\[ w_{ij} = N^{-1} \sum_{r=1}^{N} (X_{ir} - \bar{X}_i) (X_{jr} - \bar{X}_j) = \frac{n}{N} [S]_{ij} , \quad (55) \]

\[ w_{ijkl} = N^{-1} \sum_{r=1}^{N} (X_{ir} - \bar{X}_i) (X_{jr} - \bar{X}_j) (X_{kr} - \bar{X}_k) (X_{lr} - \bar{X}_l) . \quad (56) \]

When \([W]_{ijkl} = w_{ijkl} - w_{ij} w_{kl}\), Equation (53) becomes the fitting function for the ADF method, which is called WLS in the LISREL program.

**Estimating the Quadratic Model**

LISREL 8 was used to estimate the parameters because of its popularity and new capability for imposing the necessary constraints on parameters. The Kenny-Judd procedure, illustrated in Equations (6) - (17), was applied to the present study which involved two independent LV's (ξ₁ and ξ₂) and four MV's (X₁, X₂, X₃, and X₄) and one dependent variable (Yᵣ).

MV's X₁, X₂, X₃, X₄, X₁², X₂², X₁X₂, and Y_quad were selected from the data set for estimating the quadratic effects. For the purpose of estimating parameters, all MV's were treated as x's in terms of LISREL framework. All LV's: ξ₁, ξ₂, ξ₁², ξ₁δ₁, ξ₁δ₂, δ₁, δ₂, δ₃, δ₄, δ₁², δ₂² and δ₁δ₂, except ε were
included as independent LV's, and ε was considered as an error term. Table (1) shows the resulting form of the Λₜ matrix.

Because a value of 1.0 is fixed to the loading of X₁, X₂, X₃, X₄, X₁², X₂², and X₁X₂ on δ₁, δ₂, δ₃, δ₄, δ₁², δ₂² and δ₁δ₂ respectively, the diagonal matrix Θₓ, containing error variances of x's, had zeroes except for the last element representing the variance of ε. The following equations, Equations (57) - (62), defined constraints in the Φ matrix. Equations (57) - (62) correspond to Equations (12) - (17) and are repeated here for convenience. In addition, the covariance between ξ₁ and ξ₂ was also included as a free parameter being estimated. For the input file for LISREL 8, see Appendix D.

\[
\begin{align*}
Var(ξ₁²) &= 2(Var(ξ₁))², \\
Var(δ₁²) &= 2(Var(δ₁))², \\
Var(δ₂²) &= 2(Var(δ₂))², \\
Var(ξ₁δ₁) &= Var(ξ₁)Var(δ₁), \\
Var(ξ₁δ₂) &= Var(ξ₁)Var(δ₂), \\
Var(δ₁δ₂) &= Var(δ₁)Var(δ₂).
\end{align*}
\]
Table 1

$\Lambda_1$ Loading Matrix for the Quadratic Effects of $\xi_1$

<table>
<thead>
<tr>
<th>MV</th>
<th>$\xi_1$</th>
<th>$\xi_2$</th>
<th>$\xi_1^2$</th>
<th>$\xi_1\delta_1$</th>
<th>$\xi_1\delta_2$</th>
<th>$\delta_1$</th>
<th>$\delta_2$</th>
<th>$\delta_3$</th>
<th>$\delta_4$</th>
<th>$\delta_1^2$</th>
<th>$\delta_2^2$</th>
<th>$\delta_1\delta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_2$</td>
<td>f</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_3$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_4$</td>
<td>0</td>
<td>g</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_1^2$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_2^2$</td>
<td>0</td>
<td>0</td>
<td>$f^2$</td>
<td>0</td>
<td>2$f$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_1X_2$</td>
<td>0</td>
<td>0</td>
<td>$f$</td>
<td>$f$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Y</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Latent Variables
Estimating the Moderator Model

The Kenny and Judd procedure for analyzing moderator effects, illustrated in Equations (18) - (35), was also applied in the present study. MV's, \(X_1, X_2, X_3, X_4, X_1X_3, X_1X_4, X_2X_3, X_2X_4,\) and \(Y_{\text{moderator}}\) were selected from the data set for estimating the moderator effect. Again, for the purpose of estimating parameters, all MV's were treated as x's in terms of LISREL framework, all LV's: \(\xi_1, \xi_2, \xi_1\xi_2, \delta_1, \delta_2, \delta_3, \delta_4, \delta_1\delta_3, \delta_1\delta_4, \delta_2\delta_3, \delta_2\delta_4, \xi_1\delta_3, \xi_1\delta_4, \xi_2\delta_1,\) and \(\xi_2\delta_2,\) except \(e\) were included as independent LV's, and \(e\) was considered as an error term. Table (2) shows the form of the \(\Lambda_x\) matrix.

Because a value of 1.0 was fixed to the loading of \(X_1, X_2, X_3, X_4, X_1X_3, X_1X_4, X_2X_3,\) and \(X_2X_4\) on \(\delta_1, \delta_2, \delta_3, \delta_4, \delta_1\delta_3, \delta_1\delta_4, \delta_2\delta_3,\) and \(\delta_2\delta_4\) respectively, the diagonal matrix \(\Theta_s,\) containing error variances of x's, had zeroes except for the last element representing variances of \(e.\) The following equations, Equations (63) - (71), defined constraints in the \(\Phi\) matrix. Equations (63) - (71) correspond to Equations (27) - (35) and are repeated here for convenience. In addition, covariance between \(\xi_1\) and \(\xi_2\) was also included as a free parameter being estimated. For the input file for LISREL 8, see Appendix E.
### Table 2

$\Lambda_x$ Loading Matrix for the Moderator Effects of $\xi_1 \xi_2$

<table>
<thead>
<tr>
<th>MV</th>
<th>$\xi_1$</th>
<th>$\xi_2$</th>
<th>$\xi_1 \xi_2$</th>
<th>$\delta_1$</th>
<th>$\delta_2$</th>
<th>$\delta_3$</th>
<th>$\delta_4$</th>
<th>$\delta_1 \delta_4$</th>
<th>$\delta_2 \delta_3$</th>
<th>$\delta_2 \delta_4$</th>
<th>$\xi_1 \delta_3$</th>
<th>$\xi_1 \delta_4$</th>
<th>$\xi_2 \delta_1$</th>
<th>$\xi_2 \delta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_2$</td>
<td>f</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_3$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_4$</td>
<td>0</td>
<td>g</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_1 X_3$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$X_1 X_4$</td>
<td>0</td>
<td>0</td>
<td>g</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>g</td>
<td>0</td>
</tr>
<tr>
<td>$X_2 X_3$</td>
<td>0</td>
<td>0</td>
<td>f</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>f</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$X_2 X_4$</td>
<td>0</td>
<td>0</td>
<td>fg</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>f</td>
<td>0</td>
<td>g</td>
<td>0</td>
</tr>
<tr>
<td>$Y$</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
\[ \text{var}(\xi_1,\xi_2) = \text{var}(\xi_1) \text{var}(\xi_2) + [\text{cov}(\xi_1,\xi_2)]^2, \quad (63) \]

\[ \text{var}(\delta_1,\delta_3) = \text{var}(\delta_1) \text{var}(\delta_3), \quad (64) \]

\[ \text{var}(\delta_1,\delta_4) = \text{var}(\delta_1) \text{var}(\delta_4), \quad (65) \]

\[ \text{var}(\delta_2,\delta_3) = \text{var}(\delta_2) \text{var}(\delta_3), \quad (66) \]

\[ \text{var}(\delta_2,\delta_4) = \text{var}(\delta_2) \text{var}(\delta_4), \quad (67) \]

\[ \text{var}(\xi_1,\delta_3) = \text{var}(\xi_1) \text{var}(\delta_3), \quad (68) \]

\[ \text{var}(\xi_1,\delta_4) = \text{var}(\xi_1) \text{var}(\delta_4), \quad (69) \]

\[ \text{var}(\xi_2,\delta_1) = \text{var}(\xi_2) \text{var}(\delta_1), \quad (70) \]

\[ \text{var}(\xi_2,\delta_2) = \text{var}(\xi_2) \text{var}(\delta_2). \quad (71) \]

**Results**

Descriptive statistics, mean, standard deviations, skewnesses, kurtoses, and correlation coefficients obtained from the computer generated data were compared with the true parameters that were used to generate each data set. Table 3 shows descriptive results for correlated LV's, \( \xi_1 \) and \( \xi_2 \), which were generated from a multivariate normal distribution \( N(0, 1) \) with \( r = .5 \), and from a multivariate non-normal distribution, and descriptive results of the product term \( \xi_1 \xi_2 \) and the squared term \( \xi_1^2 \).
Table 3

Descriptive Results for the LV's of Normal Data (N=1000) and Non-Normal Data (N=1000)

<table>
<thead>
<tr>
<th></th>
<th>Normal Data</th>
<th></th>
<th>Non-Normal Data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Skew</td>
<td>Kurt.</td>
</tr>
<tr>
<td>True Value</td>
<td>.00</td>
<td>1.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>$\xi_1$</td>
<td>.03</td>
<td>.99</td>
<td>-.03</td>
<td>.18</td>
</tr>
<tr>
<td>$\xi_2$</td>
<td>-.01</td>
<td>1.01</td>
<td>.08</td>
<td>-.10</td>
</tr>
<tr>
<td>$\xi_1 \xi_2$</td>
<td>.48</td>
<td>1.11</td>
<td>2.21</td>
<td>7.89</td>
</tr>
<tr>
<td>$\xi_1^2$</td>
<td>.98</td>
<td>1.44</td>
<td>3.05</td>
<td>12.65</td>
</tr>
</tbody>
</table>

It was clear that the generated LV's had properties very close to the desired values. For the normal data, $\xi_1$ and $\xi_2$ were normally distributed with 0.48 correlation between them. For the non-normal data, $\xi_1$ and $\xi_2$ had 0.48 correlation between them and were not normally distributed. In both cases, $\xi_1 \xi_2$ and $\xi_1^2$ were not normally distributed with large skew and kurtosis.

The descriptive results also showed that all first-order measured variables could be considered as being distributed normally in the normal data set. Table 4
indicates that skewness of \(X_1, X_2, X_3,\) and \(X_4\) were very close to zero with the largest difference of -.10 for \(X_2\), although there was more variation in kurtosis for these first-order MV's with \(X_2\) having the largest kurtosis 0.21. For the non-normal data, all measured variables were non-normal. Although the skewness of all first-order MV's was small with \(X_3\) having the largest skewness 0.25, the kurtosis of all MV's were higher than zero with \(X_1\) having the largest kurtosis 1.73.

In addition, Table 4 shows that other MV's: \(X_1^2, X_2^2, X_1X_2, X_1X_3, X_1X_4, X_2X_3, X_2X_4, Y_m,\) and \(Y_q\) were not normally distributed in both data sets. They had skewnesses ranging from 1.12 to 3.07 and kurtoses ranging from 1.80 to 14.19 for the normal data, and skewnesses ranging from 1.93 to 5.47 and kurtoses ranging from 7.62 to 27.15 for the non-normal data.

Two covariance matrices for the normal first-order MV's and created indicators used for estimating moderator and quadratic effects of LV's are presented in Table 5 and Table 6.
Table 4

Descriptive Results for the MV’s of Normal Data (N=1000) and Non-Normal Data (N=1000)

<table>
<thead>
<tr>
<th></th>
<th>Normal Data</th>
<th></th>
<th>Non-Normal Data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Skew</td>
<td>Kurt.</td>
</tr>
<tr>
<td>$X_1$</td>
<td>.01</td>
<td>1.13</td>
<td>.03</td>
<td>.09</td>
</tr>
<tr>
<td>$X_2$</td>
<td>.03</td>
<td>.80</td>
<td>-.10</td>
<td>.21</td>
</tr>
<tr>
<td>$X_3$</td>
<td>-.01</td>
<td>1.16</td>
<td>.05</td>
<td>-.14</td>
</tr>
<tr>
<td>$X_4$</td>
<td>-.05</td>
<td>.83</td>
<td>.08</td>
<td>.00</td>
</tr>
<tr>
<td>$X_1^2$</td>
<td>1.27</td>
<td>1.83</td>
<td>2.86</td>
<td>10.89</td>
</tr>
<tr>
<td>$X_2^2$</td>
<td>.64</td>
<td>.95</td>
<td>3.07</td>
<td>14.19</td>
</tr>
<tr>
<td>$X_1X_2$</td>
<td>.58</td>
<td>1.11</td>
<td>2.51</td>
<td>8.16</td>
</tr>
<tr>
<td>$X_1X_3$</td>
<td>.50</td>
<td>1.41</td>
<td>1.92</td>
<td>7.38</td>
</tr>
<tr>
<td>$X_1X_4$</td>
<td>.31</td>
<td>.99</td>
<td>1.72</td>
<td>8.22</td>
</tr>
<tr>
<td>$X_2X_3$</td>
<td>.30</td>
<td>1.01</td>
<td>2.12</td>
<td>9.16</td>
</tr>
<tr>
<td>$X_2X_4$</td>
<td>.16</td>
<td>.71</td>
<td>1.55</td>
<td>6.85</td>
</tr>
<tr>
<td>$Y_m$</td>
<td>.25</td>
<td>1.59</td>
<td>1.12</td>
<td>1.80</td>
</tr>
<tr>
<td>$Y_4$</td>
<td>.50</td>
<td>1.66</td>
<td>1.14</td>
<td>2.39</td>
</tr>
</tbody>
</table>

Among the four first-order MV's, as expected, the sample covariance between two indicators of $x_1$, $X_1$ and $X_2$, and the covariance between two indicators of $x_2$, $X_3$ and $X_4$, were higher than other covariance terms. Sample covariance terms between the first-order MV’s, $X_1$, $X_2$, $X_3$, and $X_4$, and their product terms, $X_1X_3$, $X_1X_4$, $X_2X_3$, $X_2X_4$, and $X_1X_2$ and sample
covariance terms between the first-order MV's and their squared terms, $X_1^2$ and $X_2^2$ were near zero. The reason is that, for variables with zero mean, expected values of any odd moments equal zero in population (Kendall & Stuart, 1963), and computation of the sample covariance between the first-order MV's and their product terms or squared terms involves expected values of third moments. For example, $\text{Cov}(X_i, X_jX_k) = E(X_iX_jX_k) - E(X_i)E(X_jX_k)$ and the value would be near zero because $E(X_iX_jX_k)$, $E(X_i)$, and $E(X_jX_k)$ are all near zero in sample. Because $Y_m$ and $Y_q$ were generated as linear functions of $\xi_1$, $\xi_2$, and $\xi_1\xi_2$ or $\xi_1^2$, and the fact that $X_1$ and $X_3$ had higher loading on $\xi_1$ and $\xi_2$ respectively than $X_2$ and $X_4$, $Y_m$ and $Y_q$ had higher covariances with $X_1$ and $X_3$ than with $X_2$ and $X_4$.

Two covariance matrices for the non-normal MV's and created indicators used for estimating moderator and quadratic effects of LV's are presented in Table 7 and Table 8.

Similar to the normal data, among the four first-order MV's, the sample covariance between two indicators of $\xi_1$, $X_1$ and $X_2$, and the covariance between two indicators of $\xi_2$, $X_3$ and $X_4$, were higher than other covariance terms, and $Y_m$ and $Y_q$ had higher covariance with $X_1$ and $X_3$ than with $X_2$ and $X_4$. 
Table 5
Covariance Matrix of MV's Used in Estimating the $\xi_1\xi_2$ Effect for Normal Data

<table>
<thead>
<tr>
<th></th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_1X_3$</th>
<th>$X_1X_4$</th>
<th>$X_2X_3$</th>
<th>$X_2X_4$</th>
<th>$Y_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_2$</td>
<td>.58</td>
<td>.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_3$</td>
<td>.50</td>
<td>.30</td>
<td>1.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_4$</td>
<td>.31</td>
<td>.16</td>
<td>.64</td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_1X_3$</td>
<td>-.06</td>
<td>-.07</td>
<td>.01</td>
<td>.01</td>
<td>1.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_1X_4$</td>
<td>-.06</td>
<td>-.06</td>
<td>-.01</td>
<td>.03</td>
<td>.97</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_2X_3$</td>
<td>-.06</td>
<td>-.05</td>
<td>.05</td>
<td>.04</td>
<td>.97</td>
<td>.46</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_2X_4$</td>
<td>-.05</td>
<td>-.05</td>
<td>.02</td>
<td>.05</td>
<td>.47</td>
<td>.46</td>
<td>.50</td>
<td>.50</td>
<td></td>
</tr>
<tr>
<td>$Y_m$</td>
<td>1.15</td>
<td>.67</td>
<td>1.26</td>
<td>.77</td>
<td>.60</td>
<td>.33</td>
<td>.42</td>
<td>.23</td>
<td>2.54</td>
</tr>
</tbody>
</table>

Table 6
Covariance Matrix of MV's Used in Estimating the $\xi_1^2$ Effect for Normal Data

<table>
<thead>
<tr>
<th></th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_1^2$</th>
<th>$X_2^2$</th>
<th>$X_1X_2$</th>
<th>$Y_q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_2$</td>
<td>.58</td>
<td>.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_3$</td>
<td>.50</td>
<td>.30</td>
<td>1.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_4$</td>
<td>.31</td>
<td>.16</td>
<td>.64</td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_1^2$</td>
<td>.06</td>
<td>-.02</td>
<td>-.05</td>
<td>-.01</td>
<td>3.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_2^2$</td>
<td>-.03</td>
<td>-.02</td>
<td>-.04</td>
<td>-.02</td>
<td>.74</td>
<td>.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_1X_2$</td>
<td>.01</td>
<td>-.04</td>
<td>-.05</td>
<td>-.02</td>
<td>1.59</td>
<td>.80</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td>$Y_q$</td>
<td>1.20</td>
<td>.68</td>
<td>1.25</td>
<td>.76</td>
<td>1.03</td>
<td>.32</td>
<td>.56</td>
<td>2.76</td>
</tr>
</tbody>
</table>
However, covariance matrices in Table 7 and Table 8 for non-normal data had slightly different structures from those in Table 5 and Table 6, because the relations among variances specified in Equations (12) - (17) or Equations (27) - (35) are not applicable to non-normal distributed data. For example, some sample covariance terms between the first-order MV's, $X_1$, $X_2$, $X_3$, and $X_4$, and their product terms, $X_1X_3$, $X_1X_4$, $X_2X_3$, $X_2X_4$, and $X_1X_2$ and sample covariance terms between the first-order MV's and their squared terms, $X_1^2$ and $X_2^2$ were not near zero.

Table 7

Covariance Matrix of MV's Used in Estimating the $\xi_1\xi_2$ Effect for Non-Normal Data

<table>
<thead>
<tr>
<th></th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_1X_3$</th>
<th>$X_1X_4$</th>
<th>$X_2X_3$</th>
<th>$X_2X_4$</th>
<th>$Y_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_2$</td>
<td>.58</td>
<td>.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_3$</td>
<td>.51</td>
<td>.31</td>
<td>1.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_4$</td>
<td>.31</td>
<td>.17</td>
<td>.64</td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_1X_3$</td>
<td>.07</td>
<td>.00</td>
<td>.12</td>
<td>.08</td>
<td>2.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_1X_4$</td>
<td>.02</td>
<td>-.01</td>
<td>.06</td>
<td>.07</td>
<td>1.26</td>
<td>1.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_2X_3$</td>
<td>.02</td>
<td>-.01</td>
<td>.14</td>
<td>.09</td>
<td>1.24</td>
<td>.63</td>
<td>1.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_2X_4$</td>
<td>.00</td>
<td>-.02</td>
<td>.08</td>
<td>.09</td>
<td>.64</td>
<td>.58</td>
<td>.61</td>
<td>.57</td>
<td></td>
</tr>
<tr>
<td>$Y_m$</td>
<td>1.21</td>
<td>.70</td>
<td>1.32</td>
<td>.81</td>
<td>.99</td>
<td>.58</td>
<td>.68</td>
<td>.40</td>
<td>2.80</td>
</tr>
</tbody>
</table>
Table 8

Covariance Matrix of MV’s Used in Estimating the $\xi_1^2$ Effect for Non-Normal Data

<table>
<thead>
<tr>
<th></th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_1^2$</th>
<th>$X_2^2$</th>
<th>$X_1X_2$</th>
<th>$Y_q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_2$</td>
<td>.58</td>
<td>.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_3$</td>
<td>.51</td>
<td>.31</td>
<td>1.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_4$</td>
<td>.31</td>
<td>.17</td>
<td>.64</td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_1^2$</td>
<td>.33</td>
<td>.11</td>
<td>.08</td>
<td>.07</td>
<td>6.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_2^2$</td>
<td>.03</td>
<td>.01</td>
<td>.00</td>
<td>.01</td>
<td>1.59</td>
<td>1.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_1X_2$</td>
<td>.14</td>
<td>.02</td>
<td>.02</td>
<td>.02</td>
<td>3.13</td>
<td>1.28</td>
<td>2.08</td>
<td></td>
</tr>
<tr>
<td>$Y_q$</td>
<td>1.33</td>
<td>.74</td>
<td>1.30</td>
<td>.80</td>
<td>2.70</td>
<td>.82</td>
<td>1.47</td>
<td>3.70</td>
</tr>
</tbody>
</table>

Quadratic Effects. LISREL 8 was used to estimate the quadratic effects of $\xi_1$ with the $\Lambda_1$ loading matrix specified in Table 1 and constraints defined by relations among LV’s in Equation (57) - (62) for both normal and non-normal data from Table 6 and Table 8. Results are contained in Table 9.

Table 9 compares the estimates provided by different methods with the true parameters that were used to generate the data, both normal and non-normal. Thus, those comparisons indicate the accuracy of the estimation. Results of $\chi^2$ and root mean squared residual (RMR) were included to gauge the goodness of fit of the model.
**Table 9**

True Values and Estimates by Different Methods -- Quadratic Data and Quadratic Model for Normal and Non-Normal Data

<table>
<thead>
<tr>
<th>True Value</th>
<th>Normal Data</th>
<th>Non-Normal Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ML</td>
<td>GLS</td>
</tr>
<tr>
<td><strong>f</strong></td>
<td>.6</td>
<td>.59</td>
</tr>
<tr>
<td><strong>g</strong></td>
<td>.6</td>
<td>.60</td>
</tr>
<tr>
<td><strong>a</strong></td>
<td>.8</td>
<td>.77</td>
</tr>
<tr>
<td><strong>b</strong></td>
<td>.8</td>
<td>.83</td>
</tr>
<tr>
<td><strong>c</strong></td>
<td>.5</td>
<td>.49</td>
</tr>
<tr>
<td><strong>Cov(ξ_1,ξ_2)</strong></td>
<td>.5</td>
<td>.51</td>
</tr>
<tr>
<td><strong>Var(ξ_1)</strong></td>
<td>1.0</td>
<td>1.01</td>
</tr>
<tr>
<td><strong>Var(ξ_2)</strong></td>
<td>1.0</td>
<td>1.07</td>
</tr>
<tr>
<td><strong>Var(δ_1)</strong></td>
<td>.3</td>
<td>.27</td>
</tr>
<tr>
<td><strong>Var(δ_2)</strong></td>
<td>.3</td>
<td>.31</td>
</tr>
<tr>
<td><strong>Var(δ_3)</strong></td>
<td>.3</td>
<td>.29</td>
</tr>
<tr>
<td><strong>Var(δ_4)</strong></td>
<td>.3</td>
<td>.30</td>
</tr>
<tr>
<td><strong>Var(ε)</strong></td>
<td>.3</td>
<td>.32</td>
</tr>
<tr>
<td><strong>χ²(df=23)</strong></td>
<td>22.20</td>
<td>21.97</td>
</tr>
<tr>
<td><strong>RMR</strong></td>
<td>.03</td>
<td>.03</td>
</tr>
</tbody>
</table>

Large $χ^2$ values correspond to bad fit and small $χ^2$ values to good fit (Jöreskog & Sörbom, 1989). RMR measures the square root of the average of the squared residuals between the obtained estimate of the population covariance.
matrix ($\Sigma$) and the sample covariance matrix ($S$). Small RMR values indicate good fit (Jöreskog & Sörbom, 1989).

Overall, when the first-order MV's were normal, ML and GLS methods provided accurate estimates although the normality assumption was violated as a result of including the product and quadratic terms. The estimates in Table 9 for the normal data were very close to the true parameters for all methods. Some estimates by ULS were less accurate although the model fitted slightly better with ULS ($\chi^2 = 16.72$, df = 23 and RMR = .02). All three methods were able to provide the estimates efficiently. The numbers of iterations that took ML, GLS, and ULS were 10, 47, and 28 respectively.

When the first-order MV's were slightly non-normal, different methods were affected to different degrees. It took 256 iterations for the ML method to converge. The largest discrepancy between the ML estimates and the true values was .35. The GLS method used 49 iterations to converge. The largest difference between the GLS estimates and the true values was .23. Lastly, the number of iterations it took ULS to converge was 33. The largest discrepancy between the ULS estimates and the true values was .60. In all cases, the fit was poor with large $\chi^2$ values. LISREL 8 printed a message, *The fit is lousy!* , and did not provide any additional fit indices.
However, ADF had difficulties in analyzing both the normal and slightly non-normal data sets. In both cases, after 500 iterations, ADF still did not converge. The intermediate results showed very large negative $\chi^2$ values and near zero estimates. This could be a result of the fact the sample covariance matrix and the asymptotic covariance matrix ($\text{ACC}$) being nearly singular. Eigenvalues of the sample covariance matrix ($S$) with the created variables and the ACC were examined (Belsley, Kuh, & Welsch, 1980).

Results showed that ACC for both normal and non-normal data were extremely ill conditioned. For the normal data, among 36 eigenvalues of ACC, the largest eigenvalue was 249.96 and the smallest one was near zero (-0.00001, due to rounding error). For the non-normal data, the largest eigenvalue was 3232.10 and the smallest one was also near zero (-0.000113, due to rounding error). These near-zero eigenvalues indicated that ACC's for both normal data and non-normal data were near singular. In addition, the large differences between the largest eigenvalues and the smallest eigenvalues indicated that these two ACC’s were clearly ill-conditioned. This problem could be caused by the constructed indicators that were included in the model.

Because the eigenvalues of $S$ from the both normal and non-normal distributions were not near zero, condition numbers were used to measure the degree of ill condition (Belsley, Kuh, & Welsch, 1980). The condition number is the
ratio of the maximal to minimal eigenvalues, which provides summary information on the potential difficulties to be encountered in various calculations based on the matrix. The larger the condition number is, the more ill conditioned the given matrix is. For the data with the quadratic effects, the condition numbers of $S$ were 44.18 and 83.29 for the normal and non-normal data respectively.

**Moderator Effects.** LISREL 8 was also used to estimate the moderator effects of $\xi_1$ and $\xi_2$ with the $\Lambda_x$ loading matrix specified in Table 2 and constraints defined by relations among LV's in Equations (63) - (71) for both normal and non-normal data in Table 5 and Table 7. Results are presented in Table 10.

Similar to the estimation of the quadratic model, estimates provided by ML and ULS methods were quite accurate for the normal data, although the normality assumption was violated as a result of including the product terms. The goodness of fit index showed that estimates by ULS, ($\chi^2 = 41.09$, df = 32 and RMR = .03), fitted the data slightly better than those by ML ($\chi^2 = 48.67$, df = 32 and RMR = .04) and by GLS ($\chi^2 = 47.89$, df = 32 and RMR = .04). Also, the ML, GLS, and ULS methods provided estimates efficiently. The numbers of iterations that took ML, GLS, and ULS to converge were 10, 27, and 11 respectively.
When the first-order MV's were slightly non-normal, all estimates were less accurate. Comparing to the true values, the largest discrepancy was .18 for the ML estimates, .13 for the GLS estimates, and .22 for the ULS estimates. These three methods did not have much difficulties to converge either. The numbers of iterations that took ML, GLS, and ULS to converge were 20, 36, and 20 respectively.

ADF also had difficulties in estimating the moderator effects. The eigenvalues were also examined to see whether the $S$ and $ACC$ matrices were ill conditioned. For the data with the moderator effects, among 45 eigenvalues of $ACC$, the largest one was 124.84 and the smallest one was near zero (-0.000001, due to rounding error) for the normal. For the non-normal data, the largest eigenvalue of $ACC$ was 506.41 and the smallest one was 0.00002. In another words, the $ACC$ matrices were ill conditioned and were almost singular for the data with the moderator effects. This feature of $ACC$ matrices used in the present study explained again the reason that ADF had computational difficulties to estimate the moderator effect.

The condition numbers of $S$, which were used for estimating the moderator model were 53.75 and 75.30 for the normal and non-normal data respectively.
Table 10

True Values and Estimates by Different Methods -- Moderator Data and Moderator Model for Normal and Non-Normal Data

<table>
<thead>
<tr>
<th>True Value</th>
<th>Normal Data</th>
<th>Non-Normal Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ML</td>
<td>GLS</td>
</tr>
<tr>
<td>f</td>
<td>.6</td>
<td>.61</td>
</tr>
<tr>
<td>g</td>
<td>.6</td>
<td>.61</td>
</tr>
<tr>
<td>a</td>
<td>.8</td>
<td>.81</td>
</tr>
<tr>
<td>b</td>
<td>.8</td>
<td>.79</td>
</tr>
<tr>
<td>c</td>
<td>.5</td>
<td>.50</td>
</tr>
<tr>
<td>Cov(ξ₁,ξ₂)</td>
<td>.5</td>
<td>.51</td>
</tr>
<tr>
<td>Var(ξ₁)</td>
<td>1.0</td>
<td>.98</td>
</tr>
<tr>
<td>Var(ξ₂)</td>
<td>1.0</td>
<td>1.10</td>
</tr>
<tr>
<td>Var(δ₁)</td>
<td>.3</td>
<td>.31</td>
</tr>
<tr>
<td>Var(δ₂)</td>
<td>.3</td>
<td>.29</td>
</tr>
<tr>
<td>Var(δ₃)</td>
<td>.3</td>
<td>.27</td>
</tr>
<tr>
<td>Var(δ₄)</td>
<td>.3</td>
<td>.30</td>
</tr>
<tr>
<td>Var(ε)</td>
<td>.3</td>
<td>.30</td>
</tr>
<tr>
<td>$\chi^2$(df=32)</td>
<td>48.67</td>
<td>47.89</td>
</tr>
<tr>
<td>RMR</td>
<td>.04</td>
<td>.04</td>
</tr>
</tbody>
</table>
Discussion

Study One clearly demonstrates that violation of the normality assumption is inevitable in estimating quadratic and moderator effects of LV's with Kenny-Judd procedure, which requires the inclusion of squared and product terms of MV's. When the first-order MV's were not normally distributed, this violation were more severe. Different methods were affected to different degrees by the violation. When the first-order MV's were normally distributed, ML, GLS, and ULS methods were able to provide accurate estimates. When the first-order MV's were slightly not normal, estimates by different methods were much less accurate.

Therefore, before the Kenny-Judd procedures are applied in practice, one has to make sure that the first-order MV's involved meet the normality assumption and zero mean assumption.

This feature limits the application of the Kenny-Judd procedures. First of all, when the number of MV's involved is increased, more squared or product terms and more constraints have to be included to study the nonlinear effect. This characteristic makes the estimation process unnecessarily complicated. Secondly, the assumption of zero means for the MV's may not be practical. When empirical data are studied, mean structures often constitute a valuable part of a model and provide useful information.
about variables. The use of mean deviates for mere computational convenience can only prevent researchers from fully representing relations among MV's and LV's in a model. Thus, Study Two will be devoted to suggest and demonstrate a simpler and more general procedure to estimate nonlinear effects of LV's.
CHAPTER IV
STUDY TWO
ESTIMATING NON-LINEAR MODELS WITH ALTERNATIVE PROCEDURES IN CSM

Introduction

Study One showed that the Kenny-Judd procedures of estimating non-linear models had limited applications. The purpose of the present study was to suggest simple and general procedures for estimating quadratic and moderator models in CSM. This study would demonstrate that nonlinear effects could be estimated directly from the first-order MV's with proper constraints on certain elements of the covariance matrix of LV's and the mean structure of LV's. For example, the data matrix in Study One for estimating the quadratic effect involves $X_1, X_2, X_3, X_4, X_1X_3, X_1X_4, X_2X_3, X_2X_4,$ and $Y_q$, as shown in Figure 1. However, with the new approach, only a subset of variables, $X_1, X_2, X_3, X_4,$ and $Y_q$ are needed. In other words, the data matrix used in the new approach is a sub-matrix of the data matrix used in Kenny and Judd's method. The use of Kenny and Judd's procedures requires the full matrix, while the new procedures can be fit to the submatrix. In addition, when the means of LV's
are zero, the new model is just a submodel of the Kenny and Judd model.

\[
\begin{array}{cccccccc}
Y_q & X_1 & X_2 & X_3 & X_4 & X_1X_3 & X_1X_4 & X_2X_3 & X_2X_4 \\
\hline
Y_q & & & & & & & & \\
X_1 & & & & & & & & \\
X_2 & & & & & & & & \\
X_3 & & & & & & & & \\
X_4 & & & & & & & & \\
X_1X_3 & & & & & & & & \\
X_1X_4 & & & & & & & & \\
X_2X_3 & & & & & & & & \\
X_2X_4 & & & & & & & & \\
\end{array}
\]

Figure 1.
Comparison of Data Matrices Used in Kenny-Judd’s Procedures and the New Procedures.

The ML, GLS, ULS, and ADF estimation methods were used for the present study. In addition, the effect of distributional assumptions on the models was investigated.

**Estimating the Quadratic Model**

Unlike the Kenny-Judd procedure, where additional indicators are created for the quadratic term \( \xi_1^2 \), no additional indicators are needed for the alternative procedure. Considering a general case where there are two indicators for each independent LV, \( \xi_1 \) and \( \xi_2 \) and two
indicators for a dependent LV \( \eta \), and relations between MV's and LV's are the following:

\[
\begin{align*}
Q_1 &= \cdot \eta + \epsilon_1, \\
Q_2 &= h \eta + \epsilon_2, \\
X_1 &= \xi_1 + \delta_1, \\
X_2 &= f \xi_1 + \delta_2, \\
X_3 &= \xi_2 + \delta_3, \\
X_4 &= g \xi_2 + \delta_4, \\
\eta &= a \xi_1 + b \xi_2 + c \xi_1^2 + \zeta,
\end{align*}
\]

where \( E(\xi_1) = \mu_1, \ Var(\xi_1) = \sigma_1^2, \ E(\xi_2) = \mu_2, \ Var(\xi_2) = \sigma_2^2, \ Cov(\xi_1\xi_2) = \sigma_{12}, \ E(\xi) = E(\xi_1) = E(\xi_2) = E(\delta_1) = E(\delta_2) = E(\delta_3) = E(\delta_4) = 0, \ Var(\xi) = \psi, \ Var(\epsilon_1) = \theta_{\epsilon_1}, \ Var(\epsilon_2) = \theta_{\epsilon_2}, \ Var(\delta_1) = \theta_{\delta_1}, \ Var(\delta_2) = \theta_{\delta_2}, \ Var(\delta_3) = \theta_{\delta_3}, \ and \ Var(\delta_4) = \theta_{\delta_4}. \) The relations among LV's, \( \xi_1, \xi_2, \) and \( \xi_1^2, \) can be expressed by the following equations:

\[
\begin{align*}
E(\xi_1^2) &= \mu_1^2 + \sigma_1^2, \\
\Var(\xi_1^2) &= 2\sigma_1^4 + 4\mu_1 \sigma_1^2, \\
\Cov(\xi_1, \xi_1^2) &= 2\mu_1 \sigma_1^2,
\end{align*}
\]
\[ \text{Cov}(\xi_2, \xi_3^2) = 2\mu_1 \sigma_{12}. \] 

With constraints specified in Equations (79) - (82), parameters can be estimated based on the covariance matrix for the six MV's. Thus, the amount of computation is greatly reduced.

The most important feature of the present procedure is the relations among LV's. The inclusion of the mean structure of the LV's not only provides a general model but also avoids an identification problem related to the presence of the squared term \( \xi_1^2 \). In Equations (72) - (78), parameter \( c \) for the \( \xi_1^2 \) only appears once in Equation (78). Thus, the parameter \( c \) can be indeterminate. One way to solve this identification problem is to involve the mean structure and to apply Equations (79) - (82).

In addition, this procedure can easily be applied to other cases when more indicators are involved. As long as the number of LV's remains the same, no changes are needed for the constraints between LV's. The only changes needed will be in the \( \Lambda_\gamma \) or \( \Lambda_\xi \) matrix based on the model. However, when the number of LV's is changed, some algebraic derivations may be required to obtain proper relations among LV's of interest. For example, if \( \eta \) in Equation (78) is a function of \( \xi_1, \xi_2, \xi_3, \) and \( \xi_1^2 \), the constraints in Equations (79) - (82) will still remain the same. What is needed will be to derive the relations between \( \xi_1, \xi_2, \) or \( \xi_1^2 \) and \( \xi_3 \).
Estimating the Moderator Model

The alternative procedure for estimating the moderator effects is similar to the procedure for estimating the quadratic effects. It also involves the use of first-order MV's and constraints on variances and covariances of independent LV's. Again, considering a case where there are six indicators, \( M_1, M_2, X_1, X_2, X_3, \) and \( X_4 \), as defined in Equations (83) - (88) and the dependent latent variable \( \eta \) is a function of \( \xi_1, \xi_2, \) and \( \xi_1 \xi_2 \):

\[
M_1 = \eta_m + \epsilon_1 ,
\]

\[
M_2 = h \eta_m + \epsilon_2 ,
\]

\[
X_1 = \xi_1 + \delta_1 ,
\]

\[
X_2 = f \xi_1 + \delta_2 ,
\]

\[
X_3 = \xi_2 + \delta_3 ,
\]

\[
X_4 = g \xi_2 + \delta_4 ,
\]

\[
\eta_m = a \xi_1 + b \xi_2 + c \xi_1 \xi_2 + \zeta .
\]

With the assumptions, \( E(\xi_1) = \mu_1, \ Var(\xi_1) = \sigma_1^2, \ E(\xi_2) = \mu_2, \ Var(\xi_2) = \sigma_2^2, \ Cov(\xi_1 \xi_2) = \sigma_{12}, \ E(\xi) = E(\epsilon_1) = E(\epsilon_2) = E(\delta_1) = E(\delta_2) = E(\delta_3) = E(\delta_4) = 0, \ Var(\delta) = \psi, \ Var(\epsilon_1) = \theta_{\epsilon 1}, \ Var(\epsilon_2) = \theta_{\epsilon 2}, \ Var(\delta_1) = \theta_{\delta 1}, \ Var(\delta_2) = \theta_{\delta 2}, \ Var(\delta_3) = \theta_{\delta 3}, \) and \( \ Var(\delta_4) = \)
\( \theta_m \), the relations among LV's, \( \xi_1 \), \( \xi_2 \), and \( \xi_1 \xi_2 \), can be expressed by the following equations:

\[
E(\xi_1 \xi_2) = \mu_1 \mu_2 + \sigma_{12},
\]

\[
\text{Var}(\xi_1 \xi_2) = \mu_1^2 \sigma_2^2 + \mu_2^2 \sigma_1^2 + \sigma_1^2 \sigma_2^2 + 2\mu_1 \mu_2 \sigma_{12} + \sigma_{12}^2,
\]

\[
\text{Cov}(\xi_1, \xi_1 \xi_2) = \mu_1 \sigma_{12} + \mu_2 \sigma_1^2,
\]

\[
\text{Cov}(\xi_2, \xi_1 \xi_2) = \mu_2 \sigma_{12} + \mu_1 \sigma_2^2.
\]

With constraints specified in Equations (90) - (93), the parameters can be estimated based on the covariance matrix for the six MV's. Thus, there is no need to include additional product terms of these MV's as indicators of \( \xi_1 \xi_2 \).

Similar to the procedure of estimating the quadratic model, the most important feature of the present procedure is the relations among LV's. The mean structure of the LV's will be included not only for a general representation but also for avoiding an identification problem related to the presence of the product term \( \xi_1 \xi_2 \). The parameter \( c \) for the \( \xi_1 \xi_2 \) appears only once in Equation (89). In addition, the variance of \( \xi_1 \xi_2 \) is a function of the variance of \( \xi_1 \) and \( \xi_2 \), and \( \text{Cov}(\xi_1, \xi_2) \), when the mean of \( \xi_1 \) is fixed at zero, i.e. \( \text{Var}(\xi_1 \xi_2) = \sigma_1^2 \sigma_2^2 + \sigma_{12}^2 \). Thus, the parameter \( c \) can be indeterminate. One way to solve this identification problem is to include the mean structure and to apply Equations (90) - (93).
In addition, this procedure can be easily applied to other cases when more indicators are involved. As long as the number of LV's remains the same, no changes are needed for the constraints between LV's. The only changes needed will be in the $\Lambda_y$ or $\Lambda_z$ matrix based on the model. However, when the number of LV's is changed, some algebraic derivations may be required to obtain proper relations among LV's of interest. For example, if $\eta$ in Equation (89) is a function of $\xi_1$, $\xi_2$, $\xi_3$, and $\xi_1 \xi_3$, one has to derive the constraints for the variance of $\xi_1 \xi_3$ and the covariances of $\xi_1$, $\xi_2$, and $\xi_3$ with $\xi_1 \xi_3$.

Furthermore, the new approach allows both the moderator model and the quadratic model to be fitted to the same data, which makes possible comparison of the two models. Examination of Equations (72) - (82) and (83) - (93) shows that the only difference between fitting the moderator and quadratic models involves the constraints imposed, as expressed in Equations (79) - (82) for the quadratic model and in Equations (90) - (93) for the moderator model. However, with the Kenny-Judd procedures, both the moderator and the quadratic models can not be fitted to the same sample covariance matrix. Different models involve different constructed indicators in the Kenny and Judd approaches.
In summary, the procedures proposed in this chapter provide simple and general alternatives to the Kenny-Judd procedures and they were tested in a Monte Carlo study. Data were generated and analyzed with the LISREL 8 program. The estimates provided by ML, GLS, ULS, and ADF methods were compared with the true parameters used for generating data for accuracy and goodness of fit.

Method

**Data Generation.** The data generation procedure used in Study One was also used to generate data for the present study with minor changes. A computer program written in Microsoft FORTRAN (Appendix F) was used to generate $\xi_1$, $\xi_2$, $\xi$, $\epsilon_1$, $\epsilon_2$, $\delta_1$, $\delta_2$, $\delta_3$, and $\delta_4$ with $E(\xi_1) = 9.0$, $E(\xi_2) = 8.0$, $\text{Var}(\xi_1) = \text{Var}(\xi_2) = 1.0$, $\text{Cov}(\xi_1, \xi_2) = 0.5$, $E(\xi) = E(\epsilon_1) = E(\epsilon_2) = E(\delta_1) = E(\delta_2) = E(\delta_3) = E(\delta_4) = 0$, $\text{Var}(\xi) = 50.0$, $\text{Var}(\epsilon_1) = \text{Var}(\epsilon_2) = \text{Var}(\delta_1) = \text{Var}(\delta_2) = \text{Var}(\delta_3) = \text{Var}(\delta_4) = 0.3$. The latent variable $\eta$ was derived from Equation (78) or Equation (89) with $a = .6$, $b = 0.8$, and $c = 0.5$. Lastly $Q_1$, $Q_2$, $X_1$, $X_2$, $X_3$, $X_4$, $M_1$, and $M_2$ were derived from Equations (72) - (77) and Equations (83) - (84) with $h = 0.7$, $f = 0.5$, and $g = 0.6$. A sample size $N = 100,000$ was used to reduce the influence of sampling error.

A large residual variance was chosen ($\text{Var}(\xi) = 50.0$) for computational purposes. In the present study, the variance of the dependent latent variable $\eta_q$ or $\eta_m$, as
defined in Equation (78) or (89), is a function of mean, variances, and covariance of $\xi_1$ and $\xi_2$ and variance of $\zeta$, or

\[
\text{Var}(\eta_q) = a^2\text{Var}(\xi_1) + b^2\text{Var}(\xi_2) + 2ab\text{Cov}(\xi_1, \xi_2) + c^2\text{Var}(\xi_2^2) + 
2ac\text{Cov}(\xi_1, \xi_2^2) + 2bc\text{Cov}(\xi_2, \xi_2^2) + \text{Var}(\zeta)
\]

\[
= a^2\text{Var}(\xi_1) + b^2\text{Var}(\xi_2) + 2ab\text{Cov}(\xi_1, \xi_2) + c^2\{2[\text{Var}(\xi_1)]^2 
+ 4[E(\xi_1)]^2\text{Var}(\xi_1)\} + 2ac[2E(\xi_1)\text{Var}(\xi_1)]
+ 2bc\{2[E(\xi_1)\text{Cov}(\xi_1, \xi_2)]\} + \text{Var}(\zeta),
\]

\[
\text{Var}(\eta_m) = a^2\text{Var}(\xi_1) + b^2\text{Var}(\xi_2) + 2ab\text{Cov}(\xi_1, \xi_2) + c^2\text{Var}(\xi_1\xi_2) + 
2ac\text{Cov}(\xi_1, \xi_1\xi_2) + 2bc\text{Cov}(\xi_2, \xi_1\xi_2) + \text{Var}(\zeta)
\]

\[
= a^2\text{Var}(\xi_1) + b^2\text{Var}(\xi_2) + 2ab\text{Cov}(\xi_1, \xi_2) + c^2\{[E(\xi_1)]^2 
\text{Var}(\xi_2) + [E(\xi_2)]^2\text{Var}(\xi_1)\} + \text{Var}(\xi_1)\text{Var}(\xi_2) + 2E(\xi_1)E(\xi_2)
\text{Cov}(\xi_1, \xi_2) + [\text{Cov}(\xi_1, \xi_2)]^2\} + 2ac[E(\xi_1)\text{Cov}(\xi_1, \xi_2) + E(\xi_2)
\text{Var}(\xi_1)] + 2bc[E(\xi_2)\text{Cov}(\xi_1, \xi_2) + E(\xi_1)\text{Var}(\xi_2)] + \text{Var}(\zeta).
\]

Clearly, the variance of $\eta_q$ will be increased greatly when the means of $\xi_1$ and $\xi_2$ become larger. Thus, the proportion of $\text{Var}(\zeta)$ in $\text{Var}(\eta_q)$ becomes much smaller. For example, when $E(\xi_1) = 9.0$ and $E(\xi_2) = 8.0$, $\text{Var}(\eta_q) = 101.28$ if $\text{Var}(\zeta) = .3$. In this case, the variance of $\zeta$ is only about .3% of $\text{Var}(\eta_q)$ and it has very little effect in the estimation process. Therefore, all estimation methods have computational difficulties in providing proper solutions. Setting $\text{Var}(\zeta) = 50.0$, on the other hand, increases the
percentage of $\text{Var}(\xi)$ in $\text{Var}(\eta_4)$ to 33.12\% and avoids the empirical indeterminacy to some degrees.

**LISREL 8.** The computer generated data with variables $Q_1, Q_2, M_1, M_2, X_1, X_2, X_3,$ and $X_4$ for $N = 100,000$ were first analyzed by PRELIS 2 to obtain descriptive statistics. The MV's were selected for estimating the corresponding models. That is, $Q_1, Q_2, X_1, X_2, X_3,$ and $X_4$ were used for estimating the quadratic model and $M_1, M_2, X_1, X_2, X_3,$ and $X_4$ were selected to estimate the moderator model. In input files for LISREL 8, the following matrices were defined and used for both the quadratic model and the moderator model:

$$
\Lambda_y = \begin{bmatrix} 1 & 0 & 0 \\ h & f & 0 \\ 0 & 0 & g \\ 0 & 0 & 0 \end{bmatrix}, \quad \Lambda_x = \begin{bmatrix} 1 & 0 & 0 \\ f & 0 & 0 \\ 0 & 1 & 0 \\ 0 & g & 0 \end{bmatrix}, \quad \Gamma = [a \ b \ c],
$$

$$
\Psi = [\psi_{11}], \quad \Theta_\varepsilon = \begin{bmatrix} \theta_{\varepsilon_1} \\ \theta_{\varepsilon_2} \end{bmatrix}, \quad \Theta_\delta = \begin{bmatrix} \theta_{\delta_1} \\ \theta_{\delta_2} \\ \theta_{\delta_3} \\ \theta_{\delta_4} \end{bmatrix}.
$$

In addition, the $\Phi$ matrix was a symmetric matrix with free parameters and constraints on $\text{Var}(\xi_1^2)$, $\text{Cov}(\xi_1, \xi_1^2)$, and $\text{Cov}(\xi_2, \xi_1^2)$ specified in Equations (80) - (82) for the quadratic model or with constraints on $\text{Var}(\xi_1 \xi_2)$, $\text{Cov}(\xi_1, \xi_1 \xi_2)$, and $\text{Cov}(\xi_2, \xi_1 \xi_2)$ specified in Equations (91) - (93) for the moderator model. The $\kappa$ matrix, which represents the mean structure of $\xi_1, \xi_2,$ and $\xi_1^2$ or $\xi_1 \xi_2$ was a
full vector with free parameters and a constraint on $E(\xi_1^2)$ specified in Equation (79) for the quadratic model or a constraint on $E(\xi_1\xi_2)$ specified in Equation (90) for the moderator model. The ML, GLS, ULS, and ADF methods in LISREL 8 were used to estimate the nonlinear effects.

Results

Descriptive statistics such as mean, standard deviation, skewness and kurtosis showed that, for the data generated ($N = 100,000$), $\xi_1$ and $\xi_2$ were normally distributed with means of 9.0 and 8.0 and unit variances and with covariance of .5 (see Table 11). The error terms $\zeta$, $\epsilon_1$, $\epsilon_2$, $\delta_1$, $\delta_2$, $\delta_3$, and $\delta_4$ were also normally distributed with zero means and variances .3 except 50.0 for $\zeta$ as generated, and covariances among them were near zero. The covariance matrix for the independent LV's is presented in Table 12.

Comparisons between means and variances of $\xi_1$, $\xi_2$, $\xi_1^2$, and $\xi_1\xi_2$, covariance terms such as $\text{Cov}(\xi_1,\xi_1^2)$, $\text{Cov}(\xi_2,\xi_1^2)$, $\text{Cov}(\xi_1,\xi_1\xi_2)$, and $\text{Cov}(\xi_2,\xi_1\xi_2)$, in Table 12 and the corresponding true values in Table 16 and Table 17 not only indicated that the data were generated as expected but also confirmed the relations expressed in Equations (79) - (82) and (90) - (93).
Table 11
Descriptive Results for the LV’s (N=100,000)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Variances</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\xi_1$</td>
<td>9.003</td>
<td>.999</td>
<td>.008</td>
<td>.007</td>
</tr>
<tr>
<td>$\xi_2$</td>
<td>8.002</td>
<td>1.006</td>
<td>-.005</td>
<td>-.014</td>
</tr>
<tr>
<td>$\xi_1^2$</td>
<td>82.045</td>
<td>326.129</td>
<td>.340</td>
<td>1.159</td>
</tr>
<tr>
<td>$\xi_1\xi_2$</td>
<td>72.543</td>
<td>219.117</td>
<td>.306</td>
<td>.110</td>
</tr>
<tr>
<td>$\eta_q$</td>
<td>52.782</td>
<td>151.972</td>
<td>.171</td>
<td>.064</td>
</tr>
<tr>
<td>$\eta_m$</td>
<td>48.031</td>
<td>125.091</td>
<td>.123</td>
<td>.031</td>
</tr>
<tr>
<td>$\xi$</td>
<td>-.044</td>
<td>50.051</td>
<td>-.003</td>
<td>.007</td>
</tr>
<tr>
<td>$\epsilon_1$</td>
<td>.000</td>
<td>.301</td>
<td>.003</td>
<td>-.022</td>
</tr>
<tr>
<td>$\epsilon_2$</td>
<td>.001</td>
<td>.301</td>
<td>-.007</td>
<td>-.005</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>-.001</td>
<td>.300</td>
<td>-.003</td>
<td>.001</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>-.001</td>
<td>.299</td>
<td>-.009</td>
<td>-.031</td>
</tr>
<tr>
<td>$\delta_3$</td>
<td>.000</td>
<td>.300</td>
<td>.008</td>
<td>.005</td>
</tr>
<tr>
<td>$\delta_4$</td>
<td>.003</td>
<td>.301</td>
<td>.013</td>
<td>.014</td>
</tr>
</tbody>
</table>
Table 12
Covariance Matrix for the LV’s

<table>
<thead>
<tr>
<th></th>
<th>$\xi_1$</th>
<th>$\xi_2$</th>
<th>$\xi_1^2$</th>
<th>$\xi_1 \xi_2$</th>
<th>$\eta_q$</th>
<th>$\eta_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\xi_1$</td>
<td>.999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\xi_2$</td>
<td>.501</td>
<td>1.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\xi_1^2$</td>
<td>17.994</td>
<td>9.031</td>
<td>326.129</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\xi_1 \xi_2$</td>
<td>12.513</td>
<td>13.071</td>
<td>226.411</td>
<td>219.117</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\eta_q$</td>
<td>10.040</td>
<td>5.660</td>
<td>181.856</td>
<td>131.868</td>
<td>151.972</td>
<td></td>
</tr>
<tr>
<td>$\eta_m$</td>
<td>7.299</td>
<td>7.680</td>
<td>131.997</td>
<td>128.222</td>
<td>126.978</td>
<td>125.091</td>
</tr>
</tbody>
</table>

Descriptive results for the MV’s are presented in Table 13. All indicators of independent LV’s, $X_1$, $X_2$, $X_3$, and $X_4$ were normally distributed with near zero skewness and kurtosis. The variances of $X_1$, $X_2$, $X_3$, and $X_4$ also approximated values derived from variances of $\xi_1$ or $\xi_2$, $\delta_1$, $\delta_2$, $\delta_3$, or $\delta_4$ accordingly. Because indicators of the dependent latent variable, $Q_1$, $Q_2$, $M_1$, and $M_2$, were functions of $\eta_q$ or $\eta_m$, their skewness and kurtosis were close to the skewness and kurtosis of $\eta_q$ or $\eta_m$. Means and variances of $Q_1$, $Q_2$, $M_1$, and $M_2$ depended on $\eta_q$ or $\eta_m$, which were functions of the distribution of $\xi_1$, $\xi_2$, and $\xi_1^2$, or $\xi_1 \xi_2$. Their values in Table 13 approximated the values derived from variances of $\eta_q$ or $\eta_m$, $\zeta$, and $\epsilon_1$ or $\epsilon_2$ accordingly.
Table 13

Descriptive Results for the MV’s (N=100,000)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
<th>Skew</th>
<th>Kurt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q₁</td>
<td>52.782</td>
<td>12.339</td>
<td>.171</td>
<td>.063</td>
</tr>
<tr>
<td>Q₂</td>
<td>36.949</td>
<td>8.649</td>
<td>.170</td>
<td>.066</td>
</tr>
<tr>
<td>M₁</td>
<td>48.031</td>
<td>11.197</td>
<td>.123</td>
<td>.030</td>
</tr>
<tr>
<td>M₂</td>
<td>33.623</td>
<td>7.850</td>
<td>.121</td>
<td>.034</td>
</tr>
<tr>
<td>X₁</td>
<td>9.002</td>
<td>1.136</td>
<td>.012</td>
<td>-.009</td>
</tr>
<tr>
<td>X₂</td>
<td>4.501</td>
<td>.742</td>
<td>-.003</td>
<td>-.013</td>
</tr>
<tr>
<td>X₃</td>
<td>8.002</td>
<td>1.141</td>
<td>-.011</td>
<td>-.024</td>
</tr>
<tr>
<td>X₄</td>
<td>4.804</td>
<td>.815</td>
<td>.010</td>
<td>-.019</td>
</tr>
</tbody>
</table>

The covariance matrix for the MV’s is presented in Table 14. Because the covariance between Q₁ or Q₂ and M₁ or M₂ were not involved in estimation of either the quadratic or the moderator effects, they were excluded from the table. The values of the sample covariances among the MV’s were compared with the true values that were calculated from the known parameters in Table 15. The differences were again very small. See Appendix I for formulas used to compute the expected values.
Table 14
The Sample Covariance Matrix of MV's (N=100,000)

<table>
<thead>
<tr>
<th></th>
<th>Q₁</th>
<th>Q₂</th>
<th>M₁</th>
<th>M₂</th>
<th>X₁</th>
<th>X₂</th>
<th>X₃</th>
<th>X₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q₁</td>
<td>152.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q₂</td>
<td>106.41</td>
<td>74.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M₁</td>
<td></td>
<td></td>
<td>125.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M₂</td>
<td></td>
<td></td>
<td></td>
<td>87.58</td>
<td>61.63</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₁</td>
<td>9.98</td>
<td>6.99</td>
<td>7.25</td>
<td>5.08</td>
<td>1.29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₂</td>
<td>5.05</td>
<td>3.53</td>
<td>3.67</td>
<td>2.57</td>
<td>.50</td>
<td>.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₃</td>
<td>5.66</td>
<td>3.96</td>
<td>7.66</td>
<td>5.37</td>
<td>.50</td>
<td>.25</td>
<td>1.30</td>
<td></td>
</tr>
<tr>
<td>X₄</td>
<td>3.44</td>
<td>2.41</td>
<td>4.65</td>
<td>3.26</td>
<td>.30</td>
<td>.15</td>
<td>.60</td>
<td>.66</td>
</tr>
</tbody>
</table>

Table 15
The Expected Covariance Matrix of MV's in Population

<table>
<thead>
<tr>
<th></th>
<th>Q₁</th>
<th>Q₂</th>
<th>M₁</th>
<th>M₂</th>
<th>X₁</th>
<th>X₂</th>
<th>X₃</th>
<th>X₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q₁</td>
<td>151.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q₂</td>
<td>105.69</td>
<td>74.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M₁</td>
<td></td>
<td></td>
<td>124.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M₂</td>
<td></td>
<td></td>
<td></td>
<td>86.76</td>
<td>61.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₁</td>
<td>10.00</td>
<td>7.00</td>
<td>7.25</td>
<td>5.08</td>
<td>1.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₂</td>
<td>5.00</td>
<td>3.50</td>
<td>3.63</td>
<td>2.54</td>
<td>.50</td>
<td>.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₃</td>
<td>5.60</td>
<td>3.92</td>
<td>7.60</td>
<td>5.32</td>
<td>.50</td>
<td>.25</td>
<td>1.30</td>
<td></td>
</tr>
<tr>
<td>X₄</td>
<td>3.36</td>
<td>2.35</td>
<td>4.56</td>
<td>3.19</td>
<td>.30</td>
<td>.15</td>
<td>.60</td>
<td>.66</td>
</tr>
</tbody>
</table>
When the ACC matrix was requested in PRELIS for the present data set, $Q_1$ and $M_1$ were scaled with factor .01. Thus, the means of $Q_1$ and $M_1$ became .528 and .480 respectively and the variances of $Q_1$ and $M_1$ became .015 and .013 respectively when ADF was used as the estimation method. In addition, the covariance terms between $Q_1$ or $M_1$ and other MV's were scaled by .01. This rescaling should not affect the accuracy of estimates or the goodness of fit in any way. The relevant estimates associating with $Q_1$ or $M_1$ should be comparable with the corresponding estimates provided by other methods after rescaling with 100.

Variables were selected from the data to estimate either the quadratic model or the moderator model. That is, variables $Q_1$, $Q_2$, $X_1$, $X_2$, $X_3$, and $X_4$ were used for estimating the quadratic model, whereas variables $M_1$, $M_2$, $X_1$, $X_2$, $X_3$, and $X_4$ were used for estimating the moderator model. The ML, GLS, ULS and ADF methods were used to estimate the nonlinear effects (see Appendix G and H for the LISREL 8 input files).

During the process of estimation, a two-step method was used to provide starting values when needed. Because of the nature of the data and model used in the present study, the fitting functions may be somewhat flat in the area near the minimum value, and some estimation methods had difficulties in converging. Thus, the first step was to obtain estimates when the variance of $\xi$ was fixed at 0.0 in the LISREL input file, or PS(1,1)=0. The second step was to
use these estimates as starting values for estimating the nonlinear effects with the variance of $\zeta$ then defined as a free parameter. This method turned out to be very useful and effective.

In addition to the $\chi^2$ value and RMR, the squared multiple correlation for structural equations (SMCSE) in LISREL 8 was used as a measure of variance accounted for by the model. The SMC for the $i$:th structural equation is defined as

$$1 - \frac{\text{Var}(\zeta_i)}{\text{Var}(\eta_i)},$$

where $\text{Var}(\zeta_i)$ and $\text{Var}(\eta_i)$ are the estimated variances of $\zeta_i$ and $\eta_i$ respectively. The SMCSE measures the strength of a linear relationship. The range of its value is between zero and one, large values indicating good structural models.

Results of estimates by different methods are compared with the true parameters in Table 16 for the quadratic model and in Table 17 for the moderator model. For the quadratic model, fairly accurate estimates and good fit were obtained by the GLS method after 190 iterations. However, with the help of starting values, GLS provided the same solution with only 15 iterations. The largest discrepancy between the estimates by GLS and the true values was 1.72. The $\chi^2$ value and RMR also indicated that the model fit the data very well. The estimates by the ULS method were only slightly
less accurate with largest discrepancy being 3.52. The $\chi^2$ value was large and no other goodness of fit indices were printed by LISREL 8 because of the poor fit, and it took 465 iterations for ULS to converge with starting values. The SMCSE values showed that the structural model accounted for the same amount of variance for estimates by both GLS and ULS.

Both ML and ADF methods had computational difficulties in estimation of the quadratic model in the present study. After 1000 iterations, they still did not converge. For ML, the starting values obtained by the two-step method did not help convergency either. ADF encountered run-time error: floating-point overflow, when the starting values provided by the two-step method was used.

It was relatively easier for the GLS and ML methods to estimate the moderator model in the present study. Both GLS and ML provided fairly accurate solutions without the help of any starting values with 71 iterations for GLS and 32 for ML. The estimates were very accurate with the largest discrepancy of 1.13 for the GLS solutions and 1.16 for the ML solutions. The goodness of fit indices, $\chi^2$ and RMR, indicated that the model fit the data well for both methods. The SMCSE values showed that the structural model accounted for the same amount of variance for estimates by both GLS and ML.
The ULS and ADF methods had computational difficulties to estimate the moderator model in the present study. They failed to converge after 500 iterations. When the starting values obtained from the two-step method were used, they still did not converge after 500 iterations. The intermediate results showed that ULS found $b$ not identified and ADF found $a$ not identified and they also gave estimates that were out of bounds.

Discussion

The present study demonstrated that, with proper constraints on the covariance structure of LV's and mean structure, it is not necessary to include any additional indicators as in the Kenny-Judd procedures. These simpler procedures also avoided computational problems due to inclusion of additional created indicators. These characteristics may allow these procedures to be applied easily to the cases where more than two MV's are used as indicators of LV's or more than two LV's are involved. Although the alternative procedures could be stated mathematically, they resulted in variety of computational problems. Solutions were obtained fairly consistently using GLS. However, GLS is based on normal theory, which is not appropriate for nonlinear models.

Because the derivations of the constraints used in the alternative procedures were based on the normality
assumption, the present procedures can be applied only when the variables are normally distributed. However, these procedures may be less affected by this assumption than the Kenny-Judd procedures are, due to the fact that no additional product or squared terms are included to escalate the violation of normality assumption. Whether this speculation is true or not is subject to further study.

The reason that the estimation of nonlinear models was easy in some cases but difficult in others may be data related. One possible reason is the ratio between \( \text{Var}(Q_1), \text{Var}(Q_2), \text{Var}(M_1), \) or \( \text{Var}(M_2) \) and \( \text{Var}(\epsilon_1) \) or \( \text{Var}(\epsilon_2) \). Because both \( \text{Var}(\epsilon_1) \) and \( \text{Var}(\epsilon_2) \) were generated at .3 in the present study, they only accounted for very small portions of \( \text{Var}(Q_1), \text{Var}(Q_2), \text{Var}(M_1), \) or \( \text{Var}(M_2) \) accordingly. This characteristic of the data may make the estimation process more difficult for some methods.

Table 16

<table>
<thead>
<tr>
<th>True Value</th>
<th>GLS</th>
<th>ULS</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h )</td>
<td>.7</td>
<td>.70</td>
</tr>
<tr>
<td>( f )</td>
<td>.5</td>
<td>.50</td>
</tr>
<tr>
<td>( g )</td>
<td>.6</td>
<td>.60</td>
</tr>
<tr>
<td>( a )</td>
<td>.6</td>
<td>.43</td>
</tr>
<tr>
<td>( b )</td>
<td>.8</td>
<td>.88</td>
</tr>
</tbody>
</table>
Table 16 (continued)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.5</td>
<td>.51</td>
<td>.51</td>
</tr>
<tr>
<td>Cov($\xi_1\xi_2$)</td>
<td>.5</td>
<td>.50</td>
<td>.50</td>
</tr>
<tr>
<td>Var($\xi_1$)</td>
<td>1.0</td>
<td>.99</td>
<td>.99</td>
</tr>
<tr>
<td>Var($\xi_2$)</td>
<td>1.0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Var($\xi_2^2$)</td>
<td>326.0</td>
<td>324.28</td>
<td>322.48</td>
</tr>
<tr>
<td>Cov($\xi_1, \xi_2^2$)</td>
<td>18.0</td>
<td>17.90</td>
<td>17.83</td>
</tr>
<tr>
<td>Cov($\xi_2, \xi_1^2$)</td>
<td>9.0</td>
<td>8.98</td>
<td>8.94</td>
</tr>
<tr>
<td>Var($\eta_\delta$)</td>
<td>150.98</td>
<td>151.96</td>
<td>152.00</td>
</tr>
<tr>
<td>Var($\delta_1$)</td>
<td>.3</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var($\delta_2$)</td>
<td>.3</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var($\delta_3$)</td>
<td>.3</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var($\delta_4$)</td>
<td>.3</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var($\epsilon_1$)</td>
<td>.3</td>
<td>.30</td>
<td>.26</td>
</tr>
<tr>
<td>Var($\epsilon_2$)</td>
<td>.3</td>
<td>.30</td>
<td>.32</td>
</tr>
<tr>
<td>Var($\zeta$)</td>
<td>50.0</td>
<td>50.17</td>
<td>50.18</td>
</tr>
<tr>
<td>E($\xi_1$)</td>
<td>9.0</td>
<td>9.00</td>
<td>8.99</td>
</tr>
<tr>
<td>E($\xi_2$)</td>
<td>8.0</td>
<td>8.00</td>
<td>7.99</td>
</tr>
<tr>
<td>E($\xi_2^2$)</td>
<td>82.0</td>
<td>82.02</td>
<td>81.80</td>
</tr>
<tr>
<td>SMCSE</td>
<td>.67</td>
<td></td>
<td>.67</td>
</tr>
<tr>
<td>$\chi^2$(df=9)</td>
<td></td>
<td>17.68</td>
<td>255.20</td>
</tr>
<tr>
<td>RMR</td>
<td>.02</td>
<td></td>
<td>--</td>
</tr>
</tbody>
</table>
Table 17

True Values and Estimates -- Moderator Data and Moderator Model

<table>
<thead>
<tr>
<th></th>
<th>True Value</th>
<th>GLS</th>
<th>ML</th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>.7</td>
<td>.70</td>
<td>.70</td>
</tr>
<tr>
<td>f</td>
<td>.5</td>
<td>.50</td>
<td>.50</td>
</tr>
<tr>
<td>g</td>
<td>.6</td>
<td>.60</td>
<td>.60</td>
</tr>
<tr>
<td>a</td>
<td>.6</td>
<td>.53</td>
<td>.53</td>
</tr>
<tr>
<td>b</td>
<td>.8</td>
<td>.79</td>
<td>.79</td>
</tr>
<tr>
<td>c</td>
<td>.5</td>
<td>.51</td>
<td>.51</td>
</tr>
<tr>
<td>Cov(ξ₁,ξ₂)</td>
<td>.5</td>
<td>.50</td>
<td>.50</td>
</tr>
<tr>
<td>Var(ξ₁)</td>
<td>1.0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Var(ξ₂)</td>
<td>1.0</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Var(ξ₁,ξ₂)</td>
<td>218.25</td>
<td>218.02</td>
<td>218.03</td>
</tr>
<tr>
<td>Cov(ξ₁,ξ₁,ξ₂)</td>
<td>12.5</td>
<td>12.46</td>
<td>12.46</td>
</tr>
<tr>
<td>Cov(ξ₂,ξ₁,ξ₂)</td>
<td>13.0</td>
<td>13.01</td>
<td>13.01</td>
</tr>
<tr>
<td>Var(ηₘ)</td>
<td>123.94</td>
<td>125.07</td>
<td>125.10</td>
</tr>
<tr>
<td>Var(δ₁)</td>
<td>.3</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(δ₂)</td>
<td>.3</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(δ₃)</td>
<td>.3</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(δ₄)</td>
<td>.3</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(ε₁)</td>
<td>.3</td>
<td>.31</td>
<td>.31</td>
</tr>
<tr>
<td>Var(ε₂)</td>
<td>.3</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(ξ)</td>
<td>50.0</td>
<td>49.94</td>
<td>49.95</td>
</tr>
<tr>
<td>E(ξ₁)</td>
<td>9.0</td>
<td>9.00</td>
<td>9.00</td>
</tr>
<tr>
<td>E(ξ₂)</td>
<td>8.0</td>
<td>8.00</td>
<td>8.00</td>
</tr>
</tbody>
</table>
Table 17 (continued)

<table>
<thead>
<tr>
<th></th>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(\xi_1^2)$</td>
<td>72.5</td>
<td>72.53</td>
<td>72.53</td>
</tr>
<tr>
<td>SMCSE</td>
<td>.60</td>
<td>.60</td>
<td></td>
</tr>
<tr>
<td>$\chi^2(df=9)$</td>
<td>16.75</td>
<td>16.74</td>
<td></td>
</tr>
<tr>
<td>RMR</td>
<td>.02</td>
<td>.02</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER V
STUDY THREE
DISTINGUISHING BETWEEN THE NONLINEAR MODELS

Introduction

In applications of covariance structure modeling, moderator and quadratic effects have often been neglected because of the complexity involved in estimation. Once nonlinear models are studied, evaluating and distinguishing between them become crucial. Given that the results of Study Two showed that, in CSM, moderator effects and quadratic effects of latent variables could be estimated by the simpler alternative procedures, this study was devoted to distinguishing between moderator and quadratic models in CSM, with those procedures. The main objective of the present study was to investigate whether, as in multiple regression analysis (Lubinski & Humphreys, 1990), distinction of moderator effect from quadratic effect needed careful consideration in covariance structure modeling as well.
Distinguishing Between Quadratic and Moderator Models

Because estimating moderator and quadratic models in CSM has not been an easy task, the issue of distinguishing between moderator and quadratic models in CSM has received little attention. However, this issue was studied in multiple regression analysis (Lubinski & Humphreys, 1990).

Lubinski and Humphreys (1990) showed that interpretations of results could be misleading if a "wrong" model was tested. A statistically significant linear × linear trend (e.g. $\xi_1 \times \xi_2$) may result when, in fact, a different higher-order trend (e.g. $\xi_1^2$) better describes the covariation between the predictor set and criterion because of the correlation between $\xi_1 \times \xi_2$ and $\xi_1^2$. That is, $\xi_1 \times \xi_2$ and $\xi_1^2$ will account for much of the same variance in the dependent variable, especially as the correlation between $\xi_1$ and $\xi_2$ increases. Therefore, in practice, merely testing for the presence of a moderator effect only may not provide a full understanding of data, and Lubinski and Humphreys (1990) recommended inspection of squared terms concurrently with analytic treatments aimed at assessing moderator effects.

This rationale might also be applied to research in the context of CSM. A statistically significant moderator effect might also result when actually, a quadratic trend meaningfully better describes the structural relations, because of the high correlation between the product term of
two LV's and squared term of one LV. However, it is not clear what would happen to parameter estimates and variance accounted for in dependent LV's if a "wrong" model is fitted to data. Is it also necessary to fit both moderator and quadratic models to data in order to reveal the relations among LV's? There are no available answers to these questions in the CSM literature yet (Austin & Wolfle, 1991).

The data and procedures developed in Study Two, instead of Kenny-Judd's procedures, were used for the present study. One reason was that the procedures proposed in Study Two were easy to use. Another reason was that the use of Kenny-Judd's procedures requires inclusion of created variables. Furthermore, what type of variables are included depends on the type of models that are fitted to the data. Therefore, distinguishing between the moderator and the quadratic models with the Kenny-Judd procedures will involve two different sample covariance matrices and results will not be comparable.

Estimates of parameters, variance accounted for, and goodness of fit were evaluated when a "wrong" model was fitted to data. Again, SMCSE was used as a measure of variance accounted for by the model. It was predicted that SMCSE might be decreased when a "wrong" model was fitted to the data.
Method

The artificial normal data generated in Study Two were used again. However, in the present study, variables $Q_1, Q_2, M_1, M_2, X_1, X_2, X_3$, and $X_4$ were selected so that a "wrong" model was fitted to data. That is, a quadratic model was fitted to the data with moderator effects or $M_1, M_2, X_1, X_2, X_3$, and $X_4$, and a moderator model was fitted to the data with quadratic effects, or $Q_1, Q_2, X_1, X_2, X_3$, and $X_4$. Results were compared to the estimates obtained in Study Two, where a "correct" model was fitted to data. In particularly, the GLS method was used for the present study because, in Study Two, no other methods were able to estimate both the quadratic and the moderator models. Thus, estimates, goodness of fit of models to data, and amount of variance accounted for in structural equation provided by the GLS method were compared for distinguishing moderator effects from quadratic effects using the same set of data.

Results

Comparisons in terms of parameter estimation, goodness of fit, and variance accounted for were obvious and straightforward in the present study. The estimates by the "correct" model were closer to the parameters used to generate the data than the "wrong" model.

The data with the quadratic effects. For the data with quadratic effects, the "wrong" model provided estimates
approximating the values that were used to generate the data for most parameters except for $a$, $b$, $c$ and $\text{Var}(\xi)$ (see Table 18). The t-test for these parameters showed that they were all statistically significant.

The goodness of fit indices were not informative for the purpose of distinguishing models because they were identical for both models. Apparently, these two were equivalent models. This issue will be addressed further in the discussion section. The variance accounted for, SMCSE, was not informative either. When the "wrong" model was tested, SMCSE decreased slightly, from .6698 for the correct model to .6691 for the "wrong" model.

The data with the moderator effects. However, the same pattern was not observed when the "wrong" model was fitted to the data with moderator effects. First of all, GLS had computational difficulty in applying the quadratic model on the moderator data. Even the starting values obtained by the two-step method did not help GLS to converge. Finally, with the values that were used to generate the data as starting values, GLS was able to find a solution for the "wrong" model. Again, comparisons between estimates by the "correct" model and the "wrong" model showed that two models were equivalent with identical goodness of fit indices. There were large differences on the estimates of $a$, $b$, and $c$ provided by the two models.
Table 18
Estimates by GLS — Comparison of Moderator and Quadratic Model on the Quadratic Data (N=100,000)

<table>
<thead>
<tr>
<th></th>
<th>Estimates Correct Model</th>
<th>Estimates Wrong Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>.70</td>
<td>.70</td>
</tr>
<tr>
<td>f</td>
<td>.50</td>
<td>.50</td>
</tr>
<tr>
<td>g</td>
<td>.60</td>
<td>.60</td>
</tr>
<tr>
<td>a</td>
<td>.43</td>
<td>5.05</td>
</tr>
<tr>
<td>b</td>
<td>.88</td>
<td>-4.27</td>
</tr>
<tr>
<td>c</td>
<td>.51</td>
<td>.57</td>
</tr>
<tr>
<td>Cov(ξ₁,ξ₂)</td>
<td>.50</td>
<td>.50</td>
</tr>
<tr>
<td>Var(ξ₁)</td>
<td>.99</td>
<td>.99</td>
</tr>
<tr>
<td>Var(ξ₂)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Var(δ₁)</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(δ₂)</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(δ₃)</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(δ₄)</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(ε₁)</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(ε₂)</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(ζ)</td>
<td>50.17</td>
<td>50.28</td>
</tr>
<tr>
<td>E(ξ₁)</td>
<td>9.00</td>
<td>9.00</td>
</tr>
<tr>
<td>E(ξ₂)</td>
<td>8.00</td>
<td>8.00</td>
</tr>
<tr>
<td>SMCSE</td>
<td>.67</td>
<td>.67</td>
</tr>
<tr>
<td>χ²(df=9)</td>
<td>17.68</td>
<td>17.68</td>
</tr>
<tr>
<td>RMR</td>
<td>.02</td>
<td>.02</td>
</tr>
</tbody>
</table>
Table 19

Estimates by GLS -- Comparison of Moderator and Quadratic Model on the Moderator Data (N=100000)

<table>
<thead>
<tr>
<th></th>
<th>Estimates by Correct Model</th>
<th>Estimates by Wrong Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>.70</td>
<td>.70</td>
</tr>
<tr>
<td>f</td>
<td>.50</td>
<td>.50</td>
</tr>
<tr>
<td>g</td>
<td>.60</td>
<td>.60</td>
</tr>
<tr>
<td>a</td>
<td>.53</td>
<td>-3.60</td>
</tr>
<tr>
<td>b</td>
<td>.79</td>
<td>5.38</td>
</tr>
<tr>
<td>c</td>
<td>.51</td>
<td>.46</td>
</tr>
<tr>
<td>Cov(ξ₁,ξ₂)</td>
<td>.50</td>
<td>.50</td>
</tr>
<tr>
<td>Var(ξ₁)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Var(ξ₂)</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Var(δ₁)</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(δ₂)</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(δ₃)</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(δ₄)</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(ε₁)</td>
<td>.31</td>
<td>.31</td>
</tr>
<tr>
<td>Var(ε₂)</td>
<td>.30</td>
<td>.30</td>
</tr>
<tr>
<td>Var(ζ)</td>
<td>49.94</td>
<td>49.86</td>
</tr>
<tr>
<td>E(ξ₁)</td>
<td>9.00</td>
<td>9.00</td>
</tr>
<tr>
<td>E(ξ₂)</td>
<td>8.00</td>
<td>8.00</td>
</tr>
<tr>
<td>SMCSE</td>
<td>.60</td>
<td>.60</td>
</tr>
<tr>
<td>χ²(df=9)</td>
<td>16.75</td>
<td>16.75</td>
</tr>
<tr>
<td>RMR</td>
<td>.02</td>
<td>.02</td>
</tr>
</tbody>
</table>
Discussion

The present study demonstrated that when a data set were fitted by both the moderator model and the quadratic model, the "correct" model provided estimates that are closer to the parameters used to generate data than the "wrong" model. With the present procedures, the goodness of fit indices and the squared multiple correlation for structural equations (SMCSE) should not be used to identify a correct model in practice because the "wrong" model fitted the data equally well.

The reason that GLS had difficulty to estimate the moderator data with the "wrong" model may be data related. When the data are generated with moderator effects in the present study, variance of the product term $\xi_1 \xi_2$ were expected to be 218.25. However, when the quadratic model was fitted to this set of data, the constraints on $\text{Var}(\xi_1^2)$ expected a value of 326.00. This difference of 107.75 implied that the model was trying to account for more variances than there exist in the data. In addition, during the estimation process, the residual variance $\psi_{II}$ became smaller. Thus, the SMCSE was increased when the "wrong" model was fitted to the moderator data. However, if the data were generated by a different set of parameters, this problem may be avoided. But at least the difficulties in estimation can be an indication of lack of fit between the model and data.
The results also raised an interesting issue of equivalent models. In covariance structure modeling, two models are called equivalent if they generate identical estimates of population covariance matrices and consequently produce the same values for all goodness of fit indices (Stelzl, 1986; Lee & Hershberger, 1990). Most equivalent models that are recognized have been models with distinct path diagrams. In the present study, however, two equivalent models did not result from changing direction of one path. Rather it associated with replacing one latent variable $\xi_1^2$ with $\xi_1\xi_2$ and, in turn, replacing one set of constraints, e.g. Equations (79) - (82) with another set, e.g. Equations (90) - (93).

With the present approach of estimating nonlinear models in CSM, both the moderator model and the quadratic model estimate the same number of parameters. Both models involve two independent LV’s $\xi_1$ and $\xi_2$ and one dependent LV $\eta_q$ or $\eta_m$, each of which had two indicators. In addition, in both models, there is one latent variable without any indicators, or a phantom variable. It is the squared term $\xi_1^2$ for the quadratic model or the product term $\xi_1\xi_2$ for the moderator model. The only difference between the quadratic model and the moderator model was the constraints specified on the variance and mean of the latent variable $\xi_1^2$ or $\xi_1\xi_2$ and their covariance with $\xi_1$ and $\xi_2$. In another words, they
were different functions of $E(\xi_1)$, $E(\xi_2)$, $\text{Var}(\xi_1)$, and $\text{Var}(\xi_2)$.

It is a complicated mathematical problem to identify the source of the equivalency, which involves solving an equation system containing twenty one equations. However, let $P$ be the phantom variable and it can be demonstrated that there are relations among estimates $a$, $b$, $c$, $\text{Var}(P)$, $E(P)$, $\text{Cov}(\xi_1,P)$, $\text{Cov}(\xi_2,P)$, and $\text{Var}(\xi)$. Table 20 shows that when $\text{Var}(P)$, $E(P)$, $\text{Cov}(\xi_1,P)$, and $\text{Cov}(\xi_2,P)$ were fixed at different values the changes in $a$, $b$, $c$, and $\text{Var}(\xi)$, while other estimates remain the same.

### Table 20

The Summary of Changes in Estimates

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Var}(P)$</td>
<td>50</td>
<td>1*</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>$E(P)$</td>
<td>10</td>
<td>10</td>
<td>50*</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>$\text{Cov}(\xi_1,P)$</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>50*</td>
<td>10</td>
</tr>
<tr>
<td>$\text{Cov}(\xi_2,P)$</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>50*</td>
</tr>
<tr>
<td>$a$</td>
<td>9.60</td>
<td>9.60</td>
<td>9.59</td>
<td>4.39</td>
<td>12.44</td>
</tr>
<tr>
<td>$b$</td>
<td>-4.55</td>
<td>-4.55</td>
<td>-6.52</td>
<td>1.53</td>
<td>-7.57</td>
</tr>
<tr>
<td>$c$</td>
<td>.27</td>
<td>.27</td>
<td>.37</td>
<td>.10</td>
<td>.14</td>
</tr>
<tr>
<td>$\text{Var}(\xi)$</td>
<td>76.58</td>
<td>80.22</td>
<td>98.85</td>
<td>74.43</td>
<td>105.43</td>
</tr>
</tbody>
</table>

The five columns represent five demonstrations. The numbers with asterisks are the only ones whose values were changed in each demonstration with respect to demonstration 1. Results from demonstrations 2 - 5 were compared to the
results from demonstration 1 to examine the relations among these terms. When \( \text{Var}(P) \) was decreased, the estimates \( a, b, \) and \( c \) remained unchanged and \( \text{Var}(\xi) \) was increased in demonstration 2. When \( E(P) \) was increased, \( a \) and \( c \) were less affected than \( b \) and \( \text{Var}(\xi) \) in demonstration 3. In demonstration 4, when \( \text{Cov}(\xi_1, P) \) was increased, \( a \) and \( c \) became smaller, and \( b \) was increased. However, in demonstration 5, when \( \text{Cov}(\xi_2, P) \) was increased, \( a \) was increased, \( b \) and \( c \) became smaller.

These demonstrations indicated a complex indeterminacy among the following parameters: \( a, b, c, \text{Var}(P), E(P), \text{Cov}(\xi_1, P), \text{Cov}(\xi_2, P), \) and \( \text{Var}(\xi) \). When the variance of a phantom variable and its covariances with other latent variable were changed, the goodness of fit indices stayed the same and a few estimates were also changed. In the present case, there exit as many models as the phantom variable is defined.

In summary, results of the present study demonstrated that distinguishing between nonlinear models may not be easy for the approach proposed and evaluated in Study Two because these two models were equivalent. However, these two models may be distinguishable when higher order moments such as skewness and kurtosis are included. Further research are needed on this issue.
Introduction

The procedures for assessing moderator and nonlinear effects were applied to a real data set. The purpose of the present study is to demonstrate applications of the procedures proposed in Study Two.

The data set used were from the Project TALENT Data Bank (Wise, McLaughlin & Steel, 1979), where the latent variables of interest were: Advanced Mathematical ability (C), Mathematical ability (M), and Spatial ability (S). Three models, a linear model, a moderator model, and a quadratic model were fitted to the data. The linear model shows that the advanced mathematical abilities can be explained by a linear function of M and S. The moderator model suggests that advanced mathematical abilities would be influenced by students' performances on the mathematical ability tests, spatial ability tests, and their interaction (i.e. M, S, and M x S). The quadratic model indicates that
advanced mathematics abilities would be functions of the first- and second-order trends of mathematical abilities (M, S, and \(M^2\)). Results on estimates were compared. In addition, the issue of distinction among models was also evaluated.

Project TALENT Data

A tape containing Project TALENT Public Use File was obtained through the Inter-University Consortium for Political and Social Research (ICPSR). The data for Project TALENT Public Use File, 1960-1976 (ICPSR 7823) were originally collected by Flanagan, Tiedeman, Clemans, and Wise of the American Institutes for Research (AIR), Palo Alto, California. It is a subset of the entire data set and contains various measures for 4000 school children. Among all measures, the following variables were chosen to be included in the present study because they measure students' performance on either spatial tests or mathematical tests (Wise, McLaughlin, & Steel, 1979):

1. **Visualization in Three Dimensions** measures the ability to visualize how a figure would look after manipulation in three-dimensional space;

2. **Visualization in Two Dimensions** measures the ability to visualize how diagrams would look after being turned around on a flat surface, in contrast with the way they would look after being turned over;
(3) Mechanical Reasoning measures the ability to visualize the effects of the operation of everyday physical forces and basic kinds of mechanisms;

(4) Math Information contains items about definitions, the vocabulary of mathematics, mathematical notation, other kinds of factual information, and the understanding of mathematical concepts;

(5) Introductory High School Math measures achievement in all kinds of mathematics generally taught up to and including 9th grade, with the exception of the areas covered in the Arithmetic Reasoning. The primary emphasis of this test is on elementary algebra; other topics include fractions, decimals, percents, square roots, intuitive geometry, and elementary measurement formulas;

(6) Arithmetic Reasoning measures the ability to do the kind of reasoning required to solve arithmetic problems;

(7) Advanced Math covers topics normally taught in Grades 10-12 in college-preparatory courses. This test is intended primarily to test understanding and application of basic concepts and methods, not rote memory. A wide range of subjects is included: plane geometry, solid geometry, algebra, trigonometry, elements of analytic geometry, and introductory calculus.

Thus, in the present study, there were two latent variables, S and M. Three indicators for S are, $S_1$ - Visualization in Three Dimensions, $S_2$ - Visualization in Two
Dimensions, and \( S_3 \) - Mechanical Reasoning. The latent variable \( M \) also had three indicators, \( M_1 \) - Mathematics Information, \( M_2 \) - Arithmetic Reasoning, and \( M_3 \) - Introductory High School Mathematics. The dependent latent variable \( C \) had one indicators - Advanced Math.

**Estimating the Linear and the Nonlinear Models**

The procedures proposed in Study Two were applied to the Project TALENT Data with \( S \) and \( M \) being \( \xi_1 \) and \( \xi_2 \) respectively, and \( C \) being \( \eta \). The following matrices were used for the input file of LISREL 8:

\[
A_y = \begin{bmatrix}
1 & 0 & 0 \\
\lambda_{x_{11}} & 0 & 0 \\
\lambda_{x_{21}} & 0 & 0 \\
0 & 1 & 0 \\
0 & \lambda_{x_{32}} & 0 \\
0 & 0 & \lambda_{x_{62}}
\end{bmatrix},
\]

\[
A_x = \begin{bmatrix}
0 & 0 \\
0 & 1 \\
0 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0 & \gamma_{11} \\
0 & 0 & 0 & \gamma_{12} & \gamma_{13}
\end{bmatrix},
\]

\[
\Gamma = \begin{bmatrix}
\gamma_{11} & \gamma_{12} & \gamma_{13}
\end{bmatrix},
\]

\[
\Psi = \begin{bmatrix}
\psi_{11}
\end{bmatrix},
\]

\[
\Theta_\epsilon = \begin{bmatrix}
0
\end{bmatrix},
\]

\[
\Theta_\delta = \begin{bmatrix}
\theta_{\delta_1} & \theta_{\delta_2} \\
\theta_{\delta_3} & \theta_{\delta_4} & \theta_{\delta_5} \\
\theta_{\delta_6}
\end{bmatrix},
\]

The \( \Phi \) matrix was a symmetric matrix with free parameters and constraints on \( \text{Var}(M^2) \), \( \text{Cov}(M,M^2) \), and \( \text{Cov}(S,M^2) \) specified in
Equations (80) - (82) for the quadratic model or with constraints on $\text{Var}(MS)$, $\text{Cov}(M,MS)$, and $\text{Cov}(S,MS)$ specified in Equations (91) - (93) for the moderator model. For the linear model, these elements of the $\Phi$ matrix were fixed at zero except $\text{Var}(M^2)$ and $\text{Var}(MS)$ were fixed at one. The $\kappa$ matrix, which represents the mean structure of $M$, $S$, and $M^2$, or $MS$ was a full vector with free parameters and a constraint on $E(M^2)$ specified in Equation (79) for the quadratic model or a constraint on $E(MS)$ specified in Equation (90) for the moderator model. For the linear model, this element of the $\kappa$ matrix was fixed at zero. The GLS and ADF methods were used to estimate the linear and the nonlinear effects.

Distinguishing Nonlinear Models

The issue of distinguishing between moderator and quadratic models was also studied in the present study. Both the quadratic effect model (e.g. $M$, $S$, $M^2$) and the moderator effect model (e.g. $M$, $S$, $M \times S$) were fitted to the Project TALENT data as in Study Three. Because latent variables, $M \times S$ and $M^2$, are correlated with each other, this analysis may show that the model with $M^2$ also better characterizes relationships among LV's than the model with $M \times S$, as in Lubinski and Humphreys (1990).
Results

Before LISREL 8 was used to estimate the non-linear effects of the latent variables, the data were analyzed by PRELIS 2. Means, skewnesses, kurtoses, and covariance matrix were computed for $S_1$, $S_2$, $S_3$, $S_4$, $M_1$, $M_2$, $M_3$ and $C$. Table 21 shows the descriptive results of the Project TALENT Data.

Table 21
Descriptive Results of the Project TALENT Data (N=3323)

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
<th>Skew</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>3.39</td>
<td>2.19</td>
<td>1.42</td>
<td>2.48</td>
</tr>
<tr>
<td>$S_1$</td>
<td>8.65</td>
<td>3.11</td>
<td>.05</td>
<td>-.64</td>
</tr>
<tr>
<td>$S_2$</td>
<td>13.00</td>
<td>5.64</td>
<td>-.20</td>
<td>-.71</td>
</tr>
<tr>
<td>$S_3$</td>
<td>10.68</td>
<td>4.22</td>
<td>.07</td>
<td>-.81</td>
</tr>
<tr>
<td>$M_1$</td>
<td>8.67</td>
<td>5.12</td>
<td>.74</td>
<td>-.17</td>
</tr>
<tr>
<td>$M_2$</td>
<td>8.39</td>
<td>3.49</td>
<td>.11</td>
<td>-.81</td>
</tr>
<tr>
<td>$M_3$</td>
<td>10.66</td>
<td>4.76</td>
<td>.56</td>
<td>-.35</td>
</tr>
</tbody>
</table>

Because 210 students did not complete all the tests and 467 students had scores of zero on the tests, a total of 3323 cases was included in the present study. All MV’s, $S_1$, $S_2$, $S_3$, $M_1$, $M_2$, and $M_3$ appeared to be normally distributed with small skew and kurtosis. The variable $C$ was treated as
an ordinal variable because there were only 12 categories of scores. The covariance matrix for the MV's were presented in Table 22. Because scores on these MV's were not standardized, there were large differences on variances of variables and covariance between variables.

Table 22
Covariance Matrix for the MV's of the Project TALENT Data

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>S₁</th>
<th>S₂</th>
<th>S₃</th>
<th>M₁</th>
<th>M₂</th>
<th>M₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>4.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S₁</td>
<td></td>
<td>2.30</td>
<td>9.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S₂</td>
<td></td>
<td>2.80</td>
<td>8.27</td>
<td>31.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S₃</td>
<td></td>
<td>3.16</td>
<td>7.45</td>
<td>11.40</td>
<td>17.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M₂</td>
<td></td>
<td>3.28</td>
<td>5.10</td>
<td>6.38</td>
<td>7.23</td>
<td>11.33</td>
<td>12.18</td>
</tr>
<tr>
<td>M₃</td>
<td></td>
<td>5.93</td>
<td>6.75</td>
<td>8.94</td>
<td>9.54</td>
<td>19.02</td>
<td>11.06</td>
</tr>
</tbody>
</table>

This covariance matrix was fitted by the linear, moderator, and quadratic models. Results on estimates, their t-values, $\chi^2$, and SMCSE by GLS and ADF for the moderator model and the quadratic model are showed in Table 23. A t-value is defined as the ratio between the parameter estimate and its standard error which estimates the precision of each parameter estimate. T-values are used to
examine whether the true parameter is zero and they are independent of the units of measurement. Parameters whose t-values are larger than two in magnitude are normally judged to be different from zero (Jöreskog & Sörbom, 1989).

Results on the linear model showed that both GLS and ADF were able to estimate the linear model without the help of the starting values. It took GLS 41 iterations to converge and ADF 48 iterations. Most estimates by both methods were quite similar (see Table 23). In both cases, the fit was poor and no additional goodness of fit indices were given by LISREL. In another words, the advanced mathematical abilities were not explained well by the linear combination of mathematical ability and spatial ability.

Table 23
Results for the linear model on the Project TALENT Data (T-values are in the parentheses)

<table>
<thead>
<tr>
<th></th>
<th>GLS</th>
<th>ADF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{x21}$</td>
<td>1.50</td>
<td>1.50</td>
</tr>
<tr>
<td></td>
<td>(142.70)</td>
<td>(1045.79)</td>
</tr>
<tr>
<td>$\lambda_{x31}$</td>
<td>1.24</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>(167.55)</td>
<td>(889.52)</td>
</tr>
<tr>
<td>$\lambda_{x52}$</td>
<td>.93</td>
<td>.97</td>
</tr>
<tr>
<td></td>
<td>(133.95)</td>
<td>(845.14)</td>
</tr>
<tr>
<td>$\lambda_{x62}$</td>
<td>1.20</td>
<td>1.23</td>
</tr>
<tr>
<td></td>
<td>(167.40)</td>
<td>(897.77)</td>
</tr>
</tbody>
</table>
Table 23 (continued)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_{11}$</td>
<td>.06</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>(2.95)</td>
<td>(9.50)</td>
</tr>
<tr>
<td>$\gamma_{12}$</td>
<td>.33</td>
<td>.19</td>
</tr>
<tr>
<td></td>
<td>(17.80)</td>
<td>(8.67)</td>
</tr>
<tr>
<td>$\text{Var}(S)$</td>
<td>5.28</td>
<td>5.05</td>
</tr>
<tr>
<td></td>
<td>(28.41)</td>
<td>(32.04)</td>
</tr>
<tr>
<td>$\text{Var}(M)$</td>
<td>8.89</td>
<td>7.97</td>
</tr>
<tr>
<td></td>
<td>(27.49)</td>
<td>(33.79)</td>
</tr>
<tr>
<td>$\text{Cov}(M,S)$</td>
<td>4.72</td>
<td>4.41</td>
</tr>
<tr>
<td></td>
<td>(24.66)</td>
<td>(26.60)</td>
</tr>
<tr>
<td>$\psi_{11}$</td>
<td>2.67</td>
<td>2.14</td>
</tr>
<tr>
<td></td>
<td>(37.04)</td>
<td>(27.32)</td>
</tr>
<tr>
<td>$\text{Var}(\delta_1)$</td>
<td>4.03</td>
<td>4.09</td>
</tr>
<tr>
<td></td>
<td>(28.21)</td>
<td>(30.26)</td>
</tr>
<tr>
<td>$\text{Var}(\delta_2)$</td>
<td>19.45</td>
<td>20.05</td>
</tr>
<tr>
<td></td>
<td>(34.82)</td>
<td>(38.57)</td>
</tr>
<tr>
<td>$\text{Var}(\delta_3)$</td>
<td>7.23</td>
<td>7.83</td>
</tr>
<tr>
<td></td>
<td>(29.65)</td>
<td>(33.77)</td>
</tr>
<tr>
<td>$\text{Var}(\delta_4)$</td>
<td>5.72</td>
<td>6.52</td>
</tr>
<tr>
<td></td>
<td>(26.52)</td>
<td>(30.88)</td>
</tr>
<tr>
<td>$\text{Var}(\delta_5)$</td>
<td>4.23</td>
<td>4.31</td>
</tr>
<tr>
<td></td>
<td>(27.44)</td>
<td>(28.96)</td>
</tr>
<tr>
<td>$\text{Var}(\delta_6)$</td>
<td>4.82</td>
<td>4.92</td>
</tr>
<tr>
<td></td>
<td>(25.12)</td>
<td>(26.26)</td>
</tr>
<tr>
<td>$\text{E}(S)$</td>
<td>8.68</td>
<td>8.65</td>
</tr>
<tr>
<td></td>
<td>(161.61)</td>
<td>(1180.36)</td>
</tr>
<tr>
<td>$\text{E}(M)$</td>
<td>8.98</td>
<td>8.68</td>
</tr>
<tr>
<td></td>
<td>(103.04)</td>
<td>(1456.14)</td>
</tr>
<tr>
<td>SMCSE</td>
<td>.30</td>
<td>.28</td>
</tr>
<tr>
<td>$\chi^2(df=17)$</td>
<td>685.015</td>
<td>758.805</td>
</tr>
</tbody>
</table>
Results also showed that the nonlinear models did not fit the data well either. Actually, the variances accounted for decreased when the nonlinear models were fitted to the data (Table 24). In another words, for the present data set, the advanced mathematical ability was not explained well either by the model with interaction effect of mathematical ability and spatial ability or by the model with quadratic effect of mathematical ability as in Lubinski and Humphreys's study (1990). However, results on nonlinear models are discussed in this chapter as a demonstration of the application of the procedures proposed in Study Two.

For estimating the nonlinear models, the GLS method had less difficulties to estimate the moderator effect than the quadratic effect. It took 83 iterations for the GLS method to provide proper solutions for the moderator model. However, for the quadratic model, the GLS method could not find solutions without the starting values. Results indicated that all estimates were highly significant, except for $\gamma_{ll}$ of the quadratic model, which indicated that the effect of spatial ability was not significant on the advanced mathematical ability in the quadratic model. Although the value of the coefficient for the $M \times S$ and $M^2$ were small, -.01 and -.01 respectively, their t-values showed that both the moderator effect and the quadratic effect were statistically significant. The $\chi^2$ values were identical because they were two equivalent models. The
large value indicating a poor fit. In another words, both model did not explain the data well. In addition, there were very small differences on the SMCSE between both models, SMCSE = .264 for the moderator model and SMCSE = .266 for the quadratic model.

As for the ADF method, the pattern was just the opposite. ADF was able to estimate the quadratic model with the help of starting values, but not the moderator model. For the quadratic model, the ADF estimates were quite close to the GLS estimates. The largest difference was the estimate for the variance of the quadratic term, \( \text{Var}(M^2) = 3049.90 \) for GLS and \( \text{Var}(M^2) = 2410.78 \) for ADF. However, for the moderator model, the ADF method did not converge after 500 iterations. When the two-step method was used to obtain starting values, the ADF estimation method encountered mathematical error of floating-point overflow.

Table 24

Results for the Project TALENT Data
(T-values are in the parentheses)

<table>
<thead>
<tr>
<th></th>
<th>GLS</th>
<th>ADF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Moderator Model</td>
<td>Quadratic Model</td>
</tr>
<tr>
<td></td>
<td>( M, S, M \times S )</td>
<td>( S, M, M^2 )</td>
</tr>
<tr>
<td>( \lambda_{21} )</td>
<td>1.50 (142.69)</td>
<td>1.50 (142.69)</td>
</tr>
<tr>
<td>( \lambda_{31} )</td>
<td>1.24 (167.53)</td>
<td>1.24 (167.53)</td>
</tr>
</tbody>
</table>
Table 24 (continued)

| \( \lambda_{x52} \) | .93 | .93 | .97 |
| (133.92) | (133.92) | (844.87) |
| \( \lambda_{z62} \) | 1.20 | 1.20 | 1.23 |
| (167.30) | (167.30) | (897.45) |
| \( \gamma_{11} \) | .07 | .00 | .04 |
| ( 3.37) | ( -0.01) | ( 1.83) |
| \( \gamma_{12} \) | .39 | .46 | .56 |
| (16.16) | (12.11) | (14.32) |
| \( \gamma_{13} \) | -.01 | -.01 | -.02 |
| ( -4.00) | ( -3.99) | (-10.51) |
| \( \text{Var}(S) \) | 5.34 | 5.34 | 5.27 |
| (28.64) | (28.64) | (33.16) |
| \( \text{Var}(M) \) | 8.97 | 8.97 | 7.62 |
| (27.65) | (27.65) | (32.05) |
| \( \text{Cov}(M,S) \) | 4.79 | 4.79 | 4.31 |
| (24.90) | (24.90) | (25.97) |
| \( \text{Var}(MS) \) | 1921.54 | -- | -- |
| (26.58) | | |
| \( \text{Cov}(S,MS) \) | 89.46 | -- | -- |
| (28.04) | | |
| \( \text{Cov}(M,MS) \) | 120.78 | -- | -- |
| (27.42) | | |
| \( \text{Var}(M^2) \) | -- | 3049.90 | 2410.78 |
| | (21.50) | (30.48) |
| \( \text{Cov}(S,M^2) \) | -- | 85.93 | 74.76 |
| | (23.39) | (25.95) |
| \( \text{Cov}(M,M^2) \) | -- | 160.97 | 132.25 |
| | (25.08) | (30.48) |
| \( \psi_{11} \) | 2.68 | 2.67 | 1.93 |
| (37.03) | (36.86) | (22.00) |
| \( \text{Var}(\delta_1) \) | 4.01 | 4.01 | 3.95 |
| (28.05) | (28.05) | (29.09) |
Table 24 (continued)

<table>
<thead>
<tr>
<th>Var(δ₂)</th>
<th>19.44</th>
<th>19.44</th>
<th>19.91</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(34.82)</td>
<td>(34.82)</td>
<td>(38.30)</td>
</tr>
<tr>
<td>Var(δ₃)</td>
<td>7.23</td>
<td>7.23</td>
<td>7.70</td>
</tr>
<tr>
<td></td>
<td>(29.65)</td>
<td>(29.65)</td>
<td>(33.17)</td>
</tr>
<tr>
<td>Var(δ₄)</td>
<td>5.80</td>
<td>5.80</td>
<td>6.73</td>
</tr>
<tr>
<td></td>
<td>(26.76)</td>
<td>(26.76)</td>
<td>(31.73)</td>
</tr>
<tr>
<td>Var(δ₅)</td>
<td>4.23</td>
<td>4.23</td>
<td>4.35</td>
</tr>
<tr>
<td></td>
<td>(27.41)</td>
<td>(27.41)</td>
<td>(29.26)</td>
</tr>
<tr>
<td>Var(δ₆)</td>
<td>4.78</td>
<td>4.78</td>
<td>4.83</td>
</tr>
<tr>
<td></td>
<td>(24.81)</td>
<td>(24.81)</td>
<td>(25.73)</td>
</tr>
<tr>
<td>E(S)</td>
<td>8.68</td>
<td>8.68</td>
<td>8.65</td>
</tr>
<tr>
<td></td>
<td>(161.53)</td>
<td>(161.53)</td>
<td>(1180.18)</td>
</tr>
<tr>
<td>E(M)</td>
<td>8.97</td>
<td>8.97</td>
<td>8.68</td>
</tr>
<tr>
<td></td>
<td>(102.97)</td>
<td>(102.97)</td>
<td>(1455.00)</td>
</tr>
<tr>
<td>E(MS)</td>
<td>82.64</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(74.84)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E(M²)</td>
<td>--</td>
<td>89.49</td>
<td>82.88</td>
</tr>
<tr>
<td></td>
<td>(53.27)</td>
<td></td>
<td>(312.66)</td>
</tr>
<tr>
<td>SMCSE</td>
<td>.26</td>
<td>.26</td>
<td>.16</td>
</tr>
<tr>
<td>χ²(df=5)</td>
<td>669.13</td>
<td>669.13</td>
<td>643.998</td>
</tr>
</tbody>
</table>

Discussion

Results of modeling on the Project TALENT Data showed that the data were not explained well by the linear model. In addition, the moderator or quadratic model did not explain more variance in the dependent LV's than the linear model. That is, the advanced mathematical ability was not predicted well by the linear function of general mathematical ability (M) and spatial ability (S), or the
moderator model with $M$, $S$, and $M \times S$, or the quadratic model with $M$, $S$, and $M^2$. The reason that the present study did not replicate the study by Lubinski and Humphreys (1990) may be related to the samples used. Lubinski and Humphreys' study was based on 400,000 students, while the public usage file used in present study contained information for only 4,000 students. Another reason may be that Lubinski and Humphreys tested these models for different sex and grade groups, while all students were pooled together in the present study.

Although the nonlinear models did not fit the data well, this study demonstrated that the procedures proposed in Study Two were easy to be applied in practice. However, these procedures are not very useful for distinction of nonlinear models because they are equivalent models. This feature may prevent users from relying on statistical results only and overlook the theoretical meaning of a nonlinear model.
CHAPTER VII
GENERAL DISCUSSION

This dissertation studied the effect of distribution assumptions of normality and non-zero means on the modeling of nonlinear effects of LV’s. Also, simple alternative procedures were proposed and demonstrated with both computer generated data and real data for estimating moderator effects and quadratic effects of LV’s. In addition, the issue of distinguishing between a moderator model and a quadratic model was also discussed.

Study One showed that Kenny-Judd procedures (Kenny & Judd, 1984) for estimating nonlinear models have limited applications because the inclusion of squared and product terms violates the normality assumption inevitably. Although ML, GLS, and ULS are able to provide accurate estimates when the first-order MV’s are normally distributed with zero means, the procedures are too complicated to be applied in practice, especially when the number of variables, MV’s or LV’s, is increased.

Study Two and Study Four demonstrated that, with constraints on relevant elements of the covariance matrix and mean vector of LV’s and without additional created MV’s
needed, the simple alternative procedures provide much more general and simpler ways to estimate the nonlinear effects. They are more general because nonzero means of LV's are permitted. They are simpler because no additional MV's have to be included and because they are less affected by the number of MV's involved.

Study Three illustrated one of the disadvantage of the alternative procedures. Although the alternative procedures can be applied easily in practice, they are not useful for the purpose of distinguishing between the moderator and quadratic models. These two models are equivalent. However, they may be distinguishable with higher order statistics. Further research is needed on this issue.
APPENDIX A

THE FORTRAN PROGRAM FOR GENERATING NORMAL DATA
Program main
C For the project on nonlinear models in CSM 4/20/1992
C
C This program will generate data with interactive
effects and quadratic effects
C
implicit real*8 (a-h, o-z)
real*8 f,g,a,b,c
dimension rmt(2,2), evec(2,2), eval(1,2), vsqrt(2,2)
dimension sqteval(1,2), fp(2,2)
dimension x(1000,22), z(1000,2)
common /b/il
open(5, file='disser.dat')
open(6, file='disser.out')
open(7, file='check.out')

C read parameters and correlation matrix

read(5, 900) n
read(5, 900) nrep
900 format(i4)
read(5, 905) seed2
905 format(f10.0)
read(5, 902) rby
902 format(f5.3)

   do 10 i=l,2
       read(5, *) (rmt(i,j),j=l,2)
10 continue
read(5,901) f,g,a,b,c
901 format(5f5.2)

c compute eigen vectors and eigen values
   call deigen(rmt,evec,eval,2,2)

c compute F=F*sqrt(V)
   do 30 i=1,2
      sqteval(1,i)=0.0
      sqteval(1,i)=sqrt(eval(1,i))
30 continue
   do 80 k=1,2
      do 80 j=1,2
         if (k .ne. j) then
            vsqrt(k,j)=0.0d0
         end if
80 continue

do 90 i=1,2
vsqrt(i,i)=0.0
vsqrt(i,i)=sqteval(1,i)
90 continue

    do 50 i=1,2
    do 50 j=1,2
      fp(i,j)=0.0
    do 50 k=1,2
      fp(i,j)=fp(i,j)+evec(i,k)*vsqrt(k,j)
70 continue

C begin loop for replications
C
    call ranset(seed2)

    do 600 krep=1,nrep

C generate random numbers from N(0,1) for X,Y and U1,U2,U3,U4,U5

    do 70 i=1,n
      do 70 j=1,2
        z(i,j)=0.0
    70 z(i,j)=ggnml(v)

DO 880 I=1,N
DO 860 J=1,2
  X(I,J)=0.0
DO 860 K=1,2
  X(I,J)=X(I,J)+Z(I,K)*FP(J,K)
860 continue

DO 870 J=3,7
  X(i,j)=0.0
  X(I,J)=ggnml(v)
  X(I,J)=X(I,J)*sqrt(RBY)
870 continue
880 continue

C OBTAIN PRODUCT AND QUADRATIC LV'S,
    DO 400 I=1,N
    X(I,8)=X(I,1)*X(I,2)
    X(I,9)=X(I,1)**2

C obtain independent MV's -- X1, X2, Y1, Y2
    X(I,10)=X(I,1)+X(I,3)
    X(I,11)=F*X(I,1)+X(I,4)
    X(I,12)=X(I,2)+X(I,5)
    X(I,13)=G*X(I,2)+X(I,6)

C obtain dependent MV for Ym and Yq
    X(I,14)=A*X(I,1)+B*X(I,2)+C*X(I,8)+X(I,7)
    X(I,15)=A*X(I,1)+B*X(I,2)+C*X(I,9)+X(I,7)
400 continue

C
C Compute product and quadratic terms of MV's

```
do 490 i=1,n
  do 490 j=16,22
    X(I,16)=X(I,10)*X(I,11)
    X(I,17)=X(I,10)**2
    X(I,18)=X(I,11)**2
    X(I,19)=X(I,10)*X(I,12)
    X(I,20)=X(I,10)*X(I,13)
    X(I,21)=X(I,11)*X(I,12)
    X(I,22)=X(I,11)*X(I,13)
  490 CONTINUE

  do 720 i=1,n
    write(6,700) (x(i,j), j=10,22)
  700 format (7f10.4/6f10.4)

  do 760 i=1,n
    write(7,740) (x(i,j), j=1,9)
  740 format (9f8.4)

  600 CONTINUE
stop
end
```

```
real*8 function ranf(u)
  implicit real*8 (a-h, o-z)
  real*8 seed, mult, modulo
  common /random/ seed, mult, modulo
  seed = dmod(seed*mult,modulo)
  ranf = seed/modulo
  return
end

subroutine ranset(seed2)
  implicit real*8 (a-h, o-z)
  real*8 seed, mult, modulo, seed2
  common /random/ seed, mult, modulo
  mult = 16807
  modulo = 2147483647
  seed = seed2
  return
end

real*8 function ggnml(v)
  implicit real*8 (a-h, o-z)
  real*8 u1,u2
  u1=ranf(u)
  u2=ranf(u)
  ggnml=dsqrt(-2.0d0*dlog(u1))*dcos(6.2831853*u2)
  return
end
```
APPENDIX B

THE FORTRAN PROGRAM FOR GENERATING NON-NORMAL DATA
Program main

C For the project on nonlinear models in CSM 4/20/1992
C
C This program will generate data with interactive
effects and quadratic effects for non-normal data
with mean=0, var=1, skew=.25 and kurtosis=2.5

implicit real*8 (a-h, o-z)
real*8 f,g,a,b,c
dimension rmt(2,2), evec(2,2), eval(1,2), vsqrt(2,2)
dimension sqteval(1,2), fp(2,2)
dimension x(1000,22), z(1000,2)
common /b/ii
open(5, file='disser.dat')
open(6, file='dissern.out')
open(7, file='checknn.out')

C read parameters and correlation matrix

read(5, 900) n
read(5, 900) nrep
900 format(i4)
read(5, 905) seed2
905 format(f10.0)

read(5, 902) rby
902 format(f5.3)

    do 10 i=1,2
      read(5, *) (rmt(i,j),j=1,2)
10 continue
read(5,901) f,g,a,b,c
901 format(5f5.2)

C compute eigen vectors and eigen values
call deigen(rmt,evec,eval,2,2)

C compute FP=F*sqrt(V)
    do 30 i=1, 2
      sqteval(1,i)=0.0
      sqteval(1,i)=sqrt(eval(1,i))
30 continue
    do 80 k=1, 2
      do 80 j=1, 2
        if (k .ne. j) then
          vsqrt(k,j)=0.0d0
        end if
80 continue

    do 90 i=1,2
      vsqrt(i,i)=0.0
90 continue
vsqrt(i,i)=sqteval(1,i)

103 continue

DO 50 i=1,2
  do 50 j=1,2
    fp(i,j)=0.0
  do 50 k=1,2
    fp(i,j)=fp(i,j)+evec(i,k)*vsqrt(k,j)
50 continue

c begin loop for replications
c

call ranset(seed2)

do 600 krep=1,nrep
c
generate random numbers from N(0,1) for X,Y and U1,U2,U3,U4,U5

do 70 i=1,n
  do 70 j=1,2
    z(i,j)=0.0
    Z(i,j)=ggnml(v)
70 continue

DO 880 I=1,N
  DO 860 J=1,2
    X(I,J)=0.0
    DO 860 K=1,2
      X(I,J)=X(I,J)+Z(I,K)*FP(J,K)
860 continue

DO 870 J=3,7
  x(i,j)=0.0
  X(I,J)=ggnml(v)
  X(I,J)=X(I,J)*SQRT(RBY)
870 continue
880 continue

transformation to non-normal distribution
do 861 i=1,n
  do 861 j=1,2
    X(i,j)=-.02991+(.81075*X(i,j))+(.02991*
      *(X(i,j)**2))+(.05925*(X(i,j)**3))
861 continue

c obtian product and quadratic lv's,

dO 400 I=1,N
  X(I,8)=X(I,1)*X(I,2)
  X(I,9)=X(I,1)**2

c obtain independent mv's -- X1, X2, Y1, Y2
  X(I,10)=X(I,1)+X(I,3)
  X(I,11)=F*X(I,1)+X(I,4)
\[ X(I,12) = X(I,2) + X(I,5) \]
\[ X(I,13) = G \cdot X(I,2) + X(I,6) \]

C obtain dependent MV for interactive effect
\[ X(I,14) = A \cdot X(I,1) + B \cdot X(I,2) + C \cdot X(I,8) + X(I,7) \]
\[ X(I,15) = A \cdot X(I,1) + B \cdot X(I,2) + C \cdot X(I,9) + X(I,7) \]

400 CONTINUE

C Compute product and quadratic terms of MV's

do 490 i=1,n
  do 490 j=16,22
    \[ X(I,16) = X(I,10) \cdot X(I,11) \]
    \[ X(I,17) = X(I,10)^2 \]
    \[ X(I,18) = X(I,11)^2 \]
    \[ X(I,19) = X(I,10) \cdot X(I,12) \]
    \[ X(I,20) = X(I,10) \cdot X(I,13) \]
    \[ X(I,21) = X(I,11) \cdot X(I,12) \]
    \[ X(I,22) = X(I,11) \cdot X(I,13) \]
  490 CONTINUE

do 720 i=1,n
  write(6,700) (x(i,j), j=10,22)
700 format (7f10.4/6f10.4)
720 continue

do 865 i=1,n
  write(7,76) (x(i,j), j=1,9)
76 format (9f8.4)
865 continue

600 CONTINUE
STOP
END

real*8 function ranf(u)
implicit real*8 (a-h, o-z)
real*8 seed, mult, modulo
common /random/ seed, mult, modulo
seed = dmod(seed*mult,modulo)
ranf = seed/modulo
return
end

subroutine ranset(seed2)
implicit real*8 (a-h, o-z)
real*8 seed, mult, modulo, seed2
common /random/ seed, mult, modulo
mult = 16807
modulo = 2147483647
seed = seed2
return
end
real*8 function ggnml(v)
implicit real*8 (a-h, o-z)
real*8 u1,u2
u1=ranf(u)
u2=ranf(u)
ggnml=dsqrt(-2.0d0*dlog(u1))*dcos(6.2831853*u2)
return
end
APPENDIX C

PRELIS INPUT
computing cm for k-j generated data
da ni=22
ra fi=disser.out
la
v1 v2 v3 v4 v5 v6 v7 v8 v9 x1 x2 x3 x4 y v15 v16
v17 v18 x1x3 x1x4 x2x3 x2x4
sd 1-9 15-18
c0 all
ou ma=cm sm disser.cml
APPENDIX D

LISREL 8 INPUT FILE FOR ESTIMATING THE QUADRATIC EFFECTS
WITH KENNY-JUDD APPROACH
Nonlinear model with XSQ effect
DA NI=13 NO=1000 MA=CM
LA
X1 X2 X3 X4 YM YQ X1X2 X1SQ X2SQ X1X3 X1X4 X2X3 X2X4
CM FI=disser.CML
SE
X1 X2 X3 X4 X1SQ X2SQ X1X2 YQ/
MO NX=8 NK=12 LX=FU,FI PH=SY,FI TD=DI,FI
LK
X Y XSQ XU1 XU2 U1 U2 U3 U4 U1SQ U2SQ U1U2
FR LX 2 1 LX 4 2 LX 6 3 LX 6 5
FR LX 7 3 LX 7 4 LX 8 1 LX 8 2 LX 8 3 TD 8 8
EQ LX 2 1 LX 7 3 LX 7 4
FR PH 1 1 PH 2 2 PH 3 3 PH 4 4 PH 5 5 PH 6 6 PH 7 7
FR PH 8 8 PH 9 9 PH 10 10 PH 11 11 PH 12 12 PH 2 1
CO LX(6,3)=LX(2,1)**2
CO LX(6,5)=2*LX(2,1)
VA 1.0 LX 1 1 LX 1 6 LX 2 7 LX 3 2 LX 3 8
VA 1.0 LX 4 9 LX 5 3 LX 5 10 LX 6 11 LX 7 5 LX 7 12
VA 2.0 LX 5 4
CO PH(3,3)=2*PH(1,1)**2
CO PH(10,10)=2*PH(6,6)**2
CO PH(11,11)=2*PH(7,7)**2
CO PH(4,4)=PH(1,1)*PH(6,6)
CO PH(5,5)=PH(1,1)*PH(7,7)
CO PH(12,12)=PH(6,6)*PH(7,7)
OU AD=OFF IT=100 ME=GL
APPENDIX E

LISREL 8 INPUT FILE FOR ESTIMATING THE MODERATOR EFFECTS WITH KENNY-JUDD APPROACH
Model with INTERACTIVE effect -- Generated data
DA NI=13 NO=1000 MA=CM
LA
X1 X2 Y1 Y2 Cm Cq X1X2 X1SQ X1Y1 X1Y2 X2Y1 X2Y2
CM FI=DISSENN.CML
SE
X1 X2 Y1 Y2 X1Y1 X1Y2 X2Y1 X2Y2 Cm/
MO NX=9 NK=15 LX=FU,FI PH=SY,FI TD=DI,FI
LK
X Y XY U1 U2 U3 U4 U1U3 U1U4 U2U4 UXU3 UXU4 YU1 YU2
FR LX 2 1 LX 4 2 LX 6 3 LX 6 14 LX 9 1 LX 9 2 LX 9 3
FR LX 7 3 LX 7 12 LX 8 3 LX 8 13 LX 8 15 TD 9 9
FR PH 1 1 PH 2 2 PH 3 3 PH 4 4 PH 5 5 PH 6 6 PH 7 7
FR PH 8 8 PH 9 9 PH 10 10 PH 11 11 PH 12 12 PH 13 13
FR PH 14 14 PH 15 15
FR PH 2 1
EQ LX 2 1 LX 7 3 LX 7 12 LX 8 13
EQ LX 4 2 LX 6 3 LX 6 14 LX 8 15
CO LX(8,3)=LX(2,1)*LX(4,2)
VA 1.0 LX 1 1 LX 1 4 LX 2 5 LX 3 2 LX 3 6
VA 1.0 LX 4 7 LX 5 3 LX 5 8 LX 5 12 LX 5 14
VA 1.0 LX 6 9 LX 6 13 LX 7 10 LX 7 15 LX 8 11
CO PH(3,3)=PH(1,1)*PH(2,2)+PH(2,1)**2
CO PH(8,8)=PH(4,4)*PH(6,6)
CO PH(9,9)=PH(4,4)*PH(7,7)
CO PH(10,10)=PH(5,5)*PH(6,6)
CO PH(11,11)=PH(5,5)*PH(7,7)
CO PH(12,12)=PH(1,1)*PH(6,6)
CO PH(13,13)=PH(1,1)*PH(7,7)
CO PH(14,14)=PH(2,2)*PH(4,4)
CO PH(15,15)=PH(2,2)*PH(5,5)
OU NS AD=OFF IT=100 ME=GL
APPENDIX F

THE FORTRAN PROGRAM FOR GENERATING NON-ZERO MEAN DATA FOR STUDY TWO
Program main
C For the project on nonlinear models in CSM 4/18/1992
C
C This program will generate data with 8 indicators
for quadratic and moderator effects
C
implicit real*8 (a-h, o-z)
real*8 f,g,h,a,b,c,d,e
dimension rmt(2,2), evec(2,2), eval(1,2), vsqrt(2,2)
dimension sqteval(1,2), fp(2,2)
dimension x(1000,19), z(1000,2)
common /b/i1
open(5, file='alter.dat')
open(6, file='altermv.dat', access='append')
C
read parameters and correlation matrix
read(5, 900) n
read(5, 900) nrep
900 format(i4)
read(5, 905) seed2
905 format(f10.0)
read(5, 902) rby
902 format(f5.3)

   do 10 i=1,2
      read(5, *) (rmt(i,j),j=1,2)
10 continue
read(5,901) f,g,h,a,b,c,d,e
901 format(8f5.2)

C compute eigen vectors and eigen values
call deigen(rmt,evec,eval,2,2)

C compute FP=F*sqrt(V)
   do 30 i=1,2
      sqteval(1,i)=0.0
      sqteval(1,i)=sqrt(eval(1,i))
30 continue
   do 80 k=1,2
      do 80 j=1,2
         if (k.ne. j) then
            vsqrt(k,j)=0.0d0
         else
            vsqrt(k,j)=sqteval(1,i)
90 continue
do 50 i=1,2
  do 50 j=1,2
    fp(i,j)=0.0
    do 50 k=1,2
      fp(i,j)=fp(i,j)+evec(i,k)*vsqrt(k,j)
    continue
  continue

begin loop for replications

call ranseq(seed2)

do 600 krep=1,nrep

generate random numbers from N(0,1) for XI1 and XI2
  do 70 i=1,n
    do 70 j=1,2
      z(i,j)=0.0
    continue
  continue

DO 880 I=1,N
DO 860 J=1,2
X(I,J)=0.0
DO 860 K=1,2
  X(I,J)=X(I,J)+z(I,K)*FP(J,K)
continue

generate random numbers from N(0,.3) for el e2 u1-u4

generate random numbers from N(0,50) for zta

DO 870 J=4,9
  X(i,j)=0.0
  X(I,J)=ggnml(v)
  X(I,J)=X(I,J)*sqrt(RBY)
continue

add means to the xi1 and xi2 and OBTAIN eta

DO 400 I=1,N
  X(I,1)=X(I,1)+d
  X(I,2)=X(I,2)+e
  X(I,10)=a*X(I,1)+b*X(I,2)+c*X(I,1)*X(I,1)+X(I,3)
  X(I,11)=a*X(I,1)+b*X(I,2)+c*X(I,1)*X(I,2)+X(I,3)

obtain dependent MV's -- Q1, Q2, M1, M2

X(I,12)=X(I,10)+X(I,4)
X(I,13)=h*X(I,10)+X(I,5)
X(I,14)=X(I,11)+X(I,4)
X(I,15)=h*X(I,11)+X(I,5)

obtain independent MV's -- X1,X2,X3,X4
\[ X(I,16) = X(I,1) + X(I,6) \]
\[ X(I,17) = f \cdot X(I,1) + X(I,7) \]
\[ X(I,18) = X(I,2) + X(I,8) \]
\[ X(I,19) = g \cdot X(I,2) + X(I,9) \]

400 CONTINUE

\[ \text{do 720 } i=1,n \]
\[ \text{write}(6,700) \ (x(i,j),j=12,19) \]

700 format (8f10.4)

720 continue

600 CONTINUE
STOP
END

real*8 function ranf(u)
implicit real*8 (a-h, o-z)
real*8 seed, mult, modulo
common /random/ seed, mult, modulo
seed = dmod(seed*mult,modulo)
ranf = seed/modulo
return
end

subroutine ranset(seed2)
implicit real*8 (a-h, o-z)
real*8 seed, mult, modulo, seed2
common /random/ seed, mult, modulo
mult = 16807
modulo = 2147483647
seed = seed2
return
end

real*8 function ggnml(v)
implicit real*8 (a-h, o-z)
real*8 u1, u2
u1 = ranf(u)
u2 = ranf(u)
ggnml = dsqrt(-2.0d0*dlog(u1)) * dcos(6.2831853*u2)
return
end
APPENDIX G

LISREL 8 INPUT FOR ESTIMATING QUADRATIC EFFECTS
WITH THE ALTERNATIVE PROCEDURE
Nonlinear model with XSQ effect

DA NI=8 NO=100000 MA=CM
LA
Q1 Q2 M1 M2 X1 X2 X3 X4
CM FI=ALTERMV2.CML
ME FI=ALTERMV2.MEA
SE
Q1 Q2 X1 X2 X3 X4/
MO NX=4 NK=3 LX=FU,FI NY=2 NE=1 PH=SY,FR TD=DI,FR
GA=FU,FR PS=DI,FR LY=FU,FI KA=FR TE=DI,FR
LK
X Y XSQ
LE
QUAD
FR LX 2 1 LX 4 2 LY 2 1
VA 1.0 LX 1 1 LX 3 2 LY 1 1
CO KA(3)=PH(1,1)+KA(1)*KA(1)
CO PH(3,3)=2*PH(1,1)*PH(1,1)+4*PH(1,1)*KA(1)*KA(1)
CO PH(3,1)=2*KA(1)*PH(1,1)
CO PH(3,2)=2*KA(1)*PH(2,1)
OU AD=OFF IT=600 ME=UL
APPENDIX H

LISREL 8 INPUT FOR ESTIMATING MODERATOR EFFECTS
WITH THE ALTERNATIVE PROCEDURE
Nonlinear model with XY effect

DA NI=8 NO=100000 MA=CM
LA
Q1 Q2 M1 M2 X1 X2 X3 X4
CM FI=altermv2.CML
ME FI=ALTERMV2.MEA
SE
M1 M2 X1 X2 X3 X4/
MO NX=4 NK=3 LX=FU,FI NY=2 NE=1 PH=SY,FR TD=DI,FR
   GA=FU,FR PS=SY,FI TE=DI,FR LY=FU,FI KA=FR
LK
X Y XY
LE
MODER
FR LX 2 1 LX 4 2 LY 2 1
VA 1.0 LX 1 1 LX 3 2 LY 1 1
CO KA(3)=PH(2,1)+KA(1)*KA(2)
CO PH(3,3)=PH(2,2)*KA(1)**2+PH(1,1)*KA(2)**2+
   PH(1,1)*PH(2,2)+2*KA(1)*KA(2)*PH(2,1)+PH(2,1)**2
CO PH(3,1)=KA(1)*PH(2,1)+KA(2)*PH(1,1)
CO PH(3,2)=KA(2)*PH(2,1)+KA(1)*PH(2,2)
OU AD=OFF IT=200 ME=UL
APPENDIX I

FORMULAS USED TO COMPUTE THE EXPECTED VALUES IN STUDY TWO
\[ \text{Var}(Q_1) = a^2 \text{Var}(\xi_1) + b^2 \text{Var}(\xi_2) + 2ab \text{Cov}(\xi_1, \xi_2) + c^2 \text{Var}(\xi_1^2) + 2bc \text{Cov}(\xi_1, \xi_1^2) + 2bc \text{Cov}(\xi_2, \xi_1) + \text{Var}(\xi) + \text{Var}(\epsilon_1) \]

\[ \text{Var}(Q_2) = h^2 [a^2 \text{Var}(\xi_1) + b^2 \text{Var}(\xi_2) + 2ab \text{Cov}(\xi_1, \xi_2) + c^2 \text{Var}(\xi_1^2) + 2ac \text{Cov}(\xi_1, \xi_1^2) + \text{Var}(\xi)] + \text{Var}(\epsilon_2) \]

\[ \text{Var}(Q_3) = a^2 \text{Var}(\xi_1) + b^2 \text{Var}(\xi_2) + 2ab \text{Cov}(\xi_1, \xi_2) + c^2 \text{Var}(\xi_1^2) + 2ac \text{Cov}(\xi_1, \xi_1^2) + 2bc \text{Cov}(\xi_2, \xi_1) + \text{Var}(\xi_1) + \text{Var}(\epsilon_3) \]

\[ \text{Var}(Q_4) = h^2 [a^2 \text{Var}(\xi_1) + b^2 \text{Var}(\xi_2) + 2ab \text{Cov}(\xi_1, \xi_2) + c^2 \text{Var}(\xi_1^2) + 2ac \text{Cov}(\xi_1, \xi_1^2) + 2bc \text{Cov}(\xi_2, \xi_1) + \text{Var}(\xi)] + \text{Var}(\epsilon_4) \]

\[ \text{Var}(X_1) = \text{Var}(\xi_1) + \text{Var}(\delta_1) \]

\[ \text{Var}(X_2) = f^2 \text{Var}(\xi_1) + \text{Var}(\delta_2) \]

\[ \text{Var}(X_3) = \text{Var}(\xi_2) + \text{Var}(\delta_3) \]

\[ \text{Var}(X_4) = g^2 \text{Var}(\xi_2) + \text{Var}(\delta_4) \]

\[ \text{Cov}(Q_1, Q_2) = h [a^2 \text{Var}(\xi_1) + b^2 \text{Var}(\xi_2) + 2ab \text{Cov}(\xi_1, \xi_2) + c^2 \text{Var}(\xi_1^2) + 2ac \text{Cov}(\xi_1, \xi_1^2) + 2bc \text{Cov}(\xi_2, \xi_1) + \text{Var}(\xi)] \]

\[ \text{Cov}(X_1, X_2) = f \text{Var}(\xi_1) \]

\[ \text{Cov}(X_1, X_3) = \text{Cov}(\xi_1, \xi_2) \]

\[ \text{Cov}(X_1, X_4) = g \text{Cov}(\xi_1, \xi_2) \]

\[ \text{Cov}(X_2, X_3) = f \text{Cov}(\xi_1, \xi_2) \]

\[ \text{Cov}(X_2, X_4) = fg \text{Cov}(\xi_1, \xi_2) \]

\[ \text{Cov}(X_3, X_4) = g \text{Var}(\xi_2) \]

\[ \text{Cov}(X_1, Q_1) = a \text{Var}(\xi_1) + b \text{Cov}(\xi_1, \xi_2) + 2c \text{E}(\xi_1) \text{Var}(\xi_1) \]

\[ \text{Cov}(X_1, Q_2) = h \text{Cov}(X_1, Q_1) \]

\[ \text{Cov}(X_2, Q_1) = f \text{Cov}(X_1, Q_1) \]

\[ \text{Cov}(X_2, Q_2) = fh \text{Cov}(X_1, Q_1) \]

\[ \text{Cov}(X_3, Q_1) = b \text{Var}(\xi_2) + a \text{Cov}(\xi_1, \xi_2) + 2c \text{E}(\xi_1) \text{Cov}(\xi_1, \xi_2) \]

\[ \text{Cov}(X_3, Q_2) = h \text{Cov}(X_3, Q_1) \]
\[ \text{Cov}(X_4, Q_1) = g \text{Cov}(X_3, Q_1) \]
\[ \text{Cov}(X_4, Q_2) = g \, h \, \text{Cov}(X_3, Q_1) \]
\[ \text{Cov}(X_1, M_1) = a \text{Var}(\xi_1) + b \text{Cov}(\xi_1, \xi_2) + c \, E(\xi_1) \, \text{Cov}(\xi_1, \xi_2) + \]
\[ c \, E(\xi_2) \, \text{Var}(\xi_1) \]
\[ \text{Cov}(X_1, M_2) = h \, \text{Cov}(X_1, M_1) \]
\[ \text{Cov}(X_2, M_1) = f \, \text{Cov}(X_1, M_1) \]
\[ \text{Cov}(X_2, M_2) = f \, h \, \text{Cov}(X_1, M_1) \]
\[ \text{Cov}(X_3, M_1) = b \text{Var}(\xi_2) + a \text{Cov}(\xi_1, \xi_2) + c \, E(\xi_2) \, \text{Cov}(\xi_1, \xi_2) + \]
\[ c \, E(\xi_1) \, \text{Var}(\xi_2) \]
\[ \text{Cov}(X_3, M_2) = h \, \text{Cov}(X_3, M_1) \]
\[ \text{Cov}(X_4, M_1) = g \, \text{Cov}(X_3, M_1) \]
\[ \text{Cov}(X_4, M_2) = g \, h \, \text{Cov}(X_3, M_1) \]


