2 Basic Concepts and Applications of Structural Equation Models

2.1 Introduction

Structural equation models (SEMs) are a flexible class of models that allow complex modeling of correlated multivariate data for assessing interrelationships among observed and latent variables. It is well known in the fields of social and psychological sciences that this class of models subsumes many widely used statistical models, such as regression, factor analysis, canonical correlations, and analysis of variance and covariance. Traditional methods for analyzing SEMs were mainly developed in psychometrics, and have been extensively applied in behavioral, educational, social, and psychological researches in the past twenty years. Recently, SEMs start to draw a great deal of attention in public health, biological, and medical sciences. Today, due to strong demands in various disciplines, there are more than a dozen SEM software packages, such as AMOS, EQS6, LISREL, and Mplus. Among the various ways to specify SEMs in these software packages, we choose the key idea of the LISREL's (Jöreskog and Sörbom, 1996) formulation in defining the basic model through a measurement equation and a structural equation. The main reasons of this choice are: (i) The measurement and structural equations are very similar to the familiar regression models, hence more direct interpretation can be achieved and the common techniques in regression such as outlier and residual analyses can be conducted. (ii) It gives clear distinction between observed and latent variables. (iii) It directly models raw individual observations with latent variables, hence it can be naturally generalized to handle complex situations, and results in a direct estimation of

latent variables. (iv) The developments of statistical methodologies for subtle SEMs are more natural and comparatively easier.

In practical applications of SEMs, from the objective of a substantive study researchers usually know which are the outcome latent variables, and which are the explanatory latent variables of interest. In certain cases, a latent variable can be naturally defined by some observed variables; for example, the latent variable 'blood pressure' is naturally formed by combining systolic and diastolic blood pressures. In other cases, researchers may need to carefully select observed variables (or even though designing reliable and validated instruments) to measure latent variables of their interest. Substantive knowledge of the related study is important in making the decision. Hence, in practice, applied researchers of SEMs usually have a prior information on which observed variables should be used to define a specific latent variable. This kind of information is useful in formulating the measurement equation which relates latent variables to their corresponding observed variables. From the above discussion, it is clear that the measurement equation in SEMs is a confirmatory tool rather than an exploratory tool. Basically, the measurement equation can be regarded as a confirmatory factor analysis model. The effects of explanatory latent variables on outcome latent variables are assessed through the structural equation in the model. The use of this equation is very similar to the application of the regression model, except here latent variables are involved. At this stage, we may need to compare several structural equations, and select the most appropriate one. This is done by model comparison via reliable model comparison statistics.

The first objective of this chapter is to introduce the basic concepts of SEMs through models with a linear or a nonlinear structural equation. The second objective is to illustrate how to apply these models to substantive researches. Illustrative examples

with real medical studies will be presented.

2.2 Linear SEMs

Linear SEMs are formulated with a measurement equation and a structural equation with linear terms of explanatory latent variables. Under the assumption that the observed variables are continuous, and identically and independently distributed with a normal distribution, the linear SEM is the most basic SEM.

Let $\mathbf{y} = (y_1, \cdots, y_p)^T$ be a $p \times 1$ vector of observed variables that have been selected for the analysis, and let $\boldsymbol{\omega} = (\omega_1, \cdots, \omega_q)^T$ be a $q \times 1$ vector of latent variables that are expected to be formed from the observed variables in $\mathbf{y}$. The link between the observed variables and all the latent variables in $\boldsymbol{\omega}$ is defined by the following measurement equation: For $j = 1, \cdots, p$,

$$y_j = \mu_j + \lambda_{j1}\omega_1 + \cdots + \lambda_{jq}\omega_q + \epsilon_j, \tag{2.1}$$

where μ_j is an intercept, λ_{jk}'s are unknown coefficients that relate y_j and ω_k, and ϵ_j is the residual error. In the factor analysis terminology, λ_{jk}'s are called factor loadings. Now, according to the objective of the underlying substantive research, we have the following partition of $\boldsymbol{\omega} = (\boldsymbol{\eta}^T, \boldsymbol{\xi}^T)^T$, where $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ are $q_1 \times 1$ and $q_2 (= q - q_1) \times 1$ random vectors which respectively contain the outcome and explanatory latent variables in $\boldsymbol{\omega}$. The effects of $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_{q_2})^T$ on $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_{q_1})^T$ are assessed by the following structural equation: For $j = 1, \cdots, q_1$,

$$\eta_j = \gamma_{j1}\xi_1 + \cdots + \gamma_{jq_2}\xi_{q_2} + \delta_j, \tag{2.2}$$

where γ_{jk}'s are unknown coefficients that represent the effects of ξ_k on η_j, and δ_j is the residual error. Equations (2.1) and (2.2) define the most basic linear SEM. These two

equations can be rewritten in matrix notation:

$$\text{Measurement Equation : } \mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\boldsymbol{\omega} + \boldsymbol{\epsilon}, \tag{2.3}$$

$$\text{Structural Equation : } \boldsymbol{\eta} = \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\delta}, \tag{2.4}$$

where $\mathbf{y}$ is a $p \times 1$ random vector of observed variables, $\boldsymbol{\mu}$ is a $p \times 1$ vector of intercepts, $\boldsymbol{\Lambda}$ is a $p \times q$ unknown matrix of factor loadings, $\boldsymbol{\Gamma}$ is a $q_1 \times q_2$ unknown matrix of regression coefficients, and $\boldsymbol{\epsilon}$ and $\boldsymbol{\delta}$ are $p \times 1$ and $q_1 \times 1$ random vectors of measurement (residual) errors, respectively.

2.2.1 Measurement Equation

Most applications of SEMs are related to the study of interrelationships among latent variables. In particular, they are useful for examining the effects of explanatory latent variables on outcome latent variables of interest. For such situations, researchers usually have in mind what observed variables should be selected from the whole data set for the analysis, and how these observed variables are grouped to form latent variables. The purpose of the measurement equation in an SEM is to relate the latent variables in $\boldsymbol{\omega}$ to the observed variables in $\mathbf{y}$. It represents the link between observed and latent variables, through the specified factor loading matrix $\boldsymbol{\Lambda}$. The vector of measurement error, $\boldsymbol{\epsilon}$, is used for taking the residual errors into account.

The most important issue in formulating the measurement equation is to specify the structure of the factor loading matrix, $\boldsymbol{\Lambda}$, based on the knowledge of the observed variables in the study. Any element of $\boldsymbol{\Lambda}$ can be a free parameter or a fixed parameter with a preassigned value. The positions and the preassigned values of fixed parameters are decided on the basis of the prior knowledge of the observed variables and latent variables, and they are also related to the interpretations of latent variables. We give

here a simple example to illustrate. Consider a study concerning the effects of blood pressure and obesity on kidney disease of type 2 diabetic patients. From its objective, we are interested in three latent variables, namely one outcome latent variable about kidney disease, and two explanatory latent variables about blood pressure and obesity. Based on the related medical knowledge, plasma creatine (PCr) and urinary albumin creatinine ratio (ACR) are measured to obtain the observed variables for forming the latent variable 'kidney disease (KD)'; systolic blood pressure (SBP) and diastolic blood pressure (DBP) are measured for achieving the latent variable 'blood pressure (BP)'; and body mass index (BMI) and waist hip ratio (WHR) are measured for achieving the latent variable 'obesity (OB)'. From clear interpretation of BP, and the meaning of the observed variables, BP should only relate to SBP and DBP, but not to other observed variables. This rationale also applies to latent variables KD and OB. Thus, the system of measurement equations is defined as:

$$\begin{aligned}
\text{PCr} &= \mu_1 + \lambda_{11}\text{KD} + \epsilon_1, \\
\text{ACR} &= \mu_2 + \lambda_{21}\text{KD} + \epsilon_2, \\
\text{SBP} &= \mu_3 + \lambda_{32}\text{BP} + \epsilon_3, \\
\text{DBP} &= \mu_4 + \lambda_{42}\text{BP} + \epsilon_4, \\
\text{BMI} &= \mu_5 + \lambda_{53}\text{OB} + \epsilon_5, \\
\text{WHR} &= \mu_6 + \lambda_{63}\text{OB} + \epsilon_6,
\end{aligned} \tag{2.5}$$

or in matrix notation:

$$\begin{bmatrix} \text{PCr} \\ \text{ACR} \\ \text{SBP} \\ \text{DBP} \\ \text{BMI} \\ \text{WHR} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \\ \mu_6 \end{bmatrix} + \begin{bmatrix} \lambda_{11} & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ 0 & \lambda_{32} & 0 \\ 0 & \lambda_{42} & 0 \\ 0 & 0 & \lambda_{53} \\ 0 & 0 & \lambda_{63} \end{bmatrix} \begin{bmatrix} \text{KD} \\ \text{BP} \\ \text{OB} \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \end{bmatrix}, \tag{2.6}$$

or

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\boldsymbol{\omega} + \boldsymbol{\epsilon},$$

where $\mathbf{y}$, $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}$, $\boldsymbol{\omega}$, and $\boldsymbol{\epsilon}$ are defined as in (2.3). The interpretations of μ_j and λ_{jk} are the same as the interpretations of the intercept and regression coefficient in a regression model. It is clear from (2.5) or (2.6) that λ_{jk} is the coefficient linking the observed variable y_j with the latent variable ω_k. For instance, PCr and KD are linked up via λ_{11}. From the structure of $\boldsymbol{\Lambda}$, we know that KD is only linked with PCr and ACR, BP is only linked with SBP and DBP, and OB is only linked with BMI and WHR. As a result, the interpretation of the latent variables, KD, BP, and OB, is clear. This specific structure of $\boldsymbol{\Lambda}$ is called a non-overlapping structure. In applications of SEMs, factor loading matrices with non-overlapping structures are frequently used. In most situations, it is not necessary, or even not advisable, to use a more general structure of $\boldsymbol{\Lambda}$. For example, if λ_{12} in the above defined factor loading matrix $\boldsymbol{\Lambda}$ (see Equation (2.6)) is not zero, then BP is also related to PCr. Hence, BP cannot be interpreted as blood pressure, and the effect of blood pressure on kidney disease cannot be clearly assessed. To achieve better interpretation, we use $\boldsymbol{\Lambda}$ with a non-overlapping structure in all real applications presented in this book. In other applications with more observed and latent variables,

we need to appropriately define a specific loading matrix to formulate their relationships. From the above discussion, we recognize that the measurement equation is a confirmatory tool with a specifically defined loading matrix.

2.2.2 Structural Equation and One Extension

Recall that the latent variables identified through the measurement equation are distinguished to a $q_1 \times 1$ vector of outcome latent variables, and a $q_2(= q - q_1) \times 1$ vector of explanatory latent variables. The choices of the outcome and explanatory latent variables are based on the objective of the substantive study. Subsequently, q_1 and q_2 are defined. In the kidney disease study, it is clear from its objective that KD is the outcome latent variable, and BP and OB are the explanatory latent variables; hence, $q_1 = 1$ and $q_2 = 2$. The structural equation (2.2) or (2.4) is essentially a regression type model which regresses $\boldsymbol{\eta}$ on $\boldsymbol{\xi}$. For example, the structural equation of the kidney disease study can be:

$$\text{KD} = \gamma_1 \text{BP} + \gamma_2 \text{OB} + \delta. \tag{2.7}$$

This equation is linear in the variables and linear in the parameters. The interpretations of γ_1 and γ_2 are the same as in a regression model. Hence, they represent the magnitude of the expected changes in KD for one unit change in BP and OB, respectively. The outcome latent variables are only partially explained by the explanatory latent variables, the unexplained part is taken into account by the residual error δ.

Sometimes SEMs are called 'causal' models, and they have been used for achieving causality. We wish to emphasize that the structural equation is just a regression model with latent variables. Great caution should be taken in using this regression type model for achieving causality. See Bollen (1989) for more discussion on this issue.

A slight extension of the structural equation (2.4) that is particularly useful in business

and social-psychological research is defined by

$$\boldsymbol{\eta} = \boldsymbol{\Pi}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\delta}, \tag{2.8}$$

where $\boldsymbol{\Pi}$ is a $q_1 \times q_1$ matrix of unknown coefficients, such that $\mathbf{I} - \boldsymbol{\Pi}$ is nonsingular and the diagonal elements of $\boldsymbol{\Pi}$ are zero; and the definitions of $\boldsymbol{\Gamma}$, $\boldsymbol{\xi}$, and $\boldsymbol{\delta}$ are the same as before. According to specific applications, elements in $\boldsymbol{\Pi}$ and $\boldsymbol{\Gamma}$ can be fixed to preassigned values. This structural equation allows some outcome latent variables depend on the other outcome latent variables through an appropriately defined $\boldsymbol{\Pi}$. For example, we wish to study the effects of BP and OB on KD, as well as a disease A that is represented by an outcome latent variable η_A. Moreover, suppose that it is also interesting to examine the possible effect of KD on disease A. To tackle this problem, the following structural equation can be used:

$$\begin{pmatrix} \text{KD} \\ \eta_A \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \pi & 0 \end{pmatrix} \begin{pmatrix} \text{KD} \\ \eta_A \end{pmatrix} + \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_3 & \gamma_4 \end{pmatrix} \begin{pmatrix} \text{BP} \\ \text{OB} \end{pmatrix} + \begin{pmatrix} \delta \\ \delta_A \end{pmatrix}. \tag{2.9}$$

Here, $\boldsymbol{\eta} = (\text{KD}, \eta_A)^T$, π is the unknown coefficient that represents the effect of KD on disease A, and $\boldsymbol{\Gamma}$ is the parameter matrix with elements γ_i. The relationship of KD with BP and OB is again given by Equation (2.7), while the relationship of disease A with KD, BP, and OB is given by

$$\eta_A = \pi\text{KD} + \gamma_3\text{BP} + \gamma_4\text{OB} + \delta_A. \tag{2.10}$$

By allowing elements in $\boldsymbol{\Pi}$ and $\boldsymbol{\Gamma}$ to be fixed at any preassigned values, structural equation (2.8) achieves considerable flexibility in handling rather complex relationships among latent variables. More general structural equations with fixed covariates and nonlinear terms of latent variables will be discussed in Sections 2.3 and 2.4 of this chapter.

2.2.3 Assumptions of Linear SEMs

Like most statistical models, the standard linear SEMs have some assumptions. In practical applications of SEMs, it is important to make sure that these assumptions are valid. Let $\{\mathbf{y}_i,\ i = 1, \cdots, n\}$ be the observed data set with a sample size n, where $\mathbf{y}_i$ can be modeled via the linear SEM with latent variables $\boldsymbol{\eta}_i$ and $\boldsymbol{\xi}_i$, and measurement errors $\boldsymbol{\epsilon}_i$ and $\boldsymbol{\delta}_i$, as defined by the measurement equation (2.3) and the structural equation (2.8). The assumptions of the model are given as follows: For $i = 1, \cdots, n$,

Assumption A1: The random vectors of residual errors $\boldsymbol{\epsilon}_i$ are independently and identically distributed (i.i.d.) according to $N[\mathbf{0}, \boldsymbol{\Psi}_\epsilon]$, where $\boldsymbol{\Psi}_\epsilon$ is a diagonal covariance matrix.

Assumption A2: The random vectors of explanatory latent variables $\boldsymbol{\xi}_i$ are i.i.d. according to $N[\mathbf{0}, \boldsymbol{\Phi}]$, where $\boldsymbol{\Phi}$ is a general covariance matrix.

Assumption A3: The random vectors of residual errors $\boldsymbol{\delta}_i$ are i.i.d. according to $N[\mathbf{0}, \boldsymbol{\Psi}_\delta]$, where $\boldsymbol{\Psi}_\delta$ is a diagonal covariance matrix.

Assumption A4: $\boldsymbol{\delta}_i$ is independent of $\boldsymbol{\xi}_i$, and $\boldsymbol{\epsilon}_i$ is independent of $\boldsymbol{\omega}_i$ and $\boldsymbol{\delta}_i$.

These assumptions introduce three unknown parameter matrices, namely $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}_\epsilon$, and $\boldsymbol{\Psi}_\delta$. As $\boldsymbol{\eta}_i$ is a linear combination of $\boldsymbol{\xi}_i$ and $\boldsymbol{\delta}_i$, the observations of $\boldsymbol{\eta}_i$ are also i.i.d. according to a normal distribution with a covariance matrix that depends on $\boldsymbol{\Pi}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\Phi}$, and $\boldsymbol{\Psi}_\delta$. As a result, based on Assumptions A2 and A3 and the linear structural equation, $\boldsymbol{\omega}_i$ are i.i.d. according to a normal distribution. Moreover, these assumptions also implicitly restrict $\mathbf{y}_i$ to be i.i.d. according to a normal distribution.

2.2.4 Model Identification

Model identification is an issue relevant to all structural equation models. In general, consider an SEM with a measurement equation and a structural equation which are formulated with an unknown parameter vector $\boldsymbol{\theta}$. The traditional definition of identification is based on $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, the population covariance matrix of the observed variables in $\mathbf{y}$. The model is said to be identified if for any $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, $\boldsymbol{\Sigma}(\boldsymbol{\theta}_1) = \boldsymbol{\Sigma}(\boldsymbol{\theta}_2)$ implies $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$ (see Bollen, 1989). This definition is difficult to apply to a complex SEM whose $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is very complicated or even impossible to derive. For almost all existing SEMs, the parameters involved in the measurement equation are different from those involved in the structural equation, which are respectively defined by distinctive $\boldsymbol{\theta}$ and $\boldsymbol{\theta}^*$ that have no common elements. Hence, we consider a definition of identification on the basis of the fundamental measurement equation $m(\boldsymbol{\theta})$ and structural equation $s(\boldsymbol{\theta}^*)$. We regard the measurement equation as identified if for any $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, $m(\boldsymbol{\theta}_1) = m(\boldsymbol{\theta}_2)$ implies $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$. Similarly, we regard the structural equation as identified if for any $\boldsymbol{\theta}_1^*$ and $\boldsymbol{\theta}_2^*$, $s(\boldsymbol{\theta}_1^*) = s(\boldsymbol{\theta}_2^*)$ implies $\boldsymbol{\theta}_1^* = \boldsymbol{\theta}_2^*$. Moreover, we regard the SEM as identified if both of its measurement equation and structural equation are identified. General necessary and sufficient conditions to guarantee the identifiability of an SEM are difficult to find. Hence, in practical applications of SEMs, we mainly concern the sufficient conditions for achieving an identified model. For most SEMs, such sufficient conditions are usually available, and the issue is approached on a problem-by-problem basis.

Consider linear SEMs. The measurement equation is not identified without imposing some identification condition. This is because for any nonsingular matrix $\mathbf{M}$, we have

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\boldsymbol{\omega} + \boldsymbol{\epsilon} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{M}\mathbf{M}^{-1}\boldsymbol{\omega} + \boldsymbol{\epsilon} \tag{2.11}$$

$$= \boldsymbol{\mu} + \boldsymbol{\Lambda}^*\boldsymbol{\omega}^* + \boldsymbol{\epsilon}, \tag{2.12}$$

where $\mathbf{\Lambda}^* = \mathbf{\Lambda}\mathbf{M}$, and $\boldsymbol{\omega}^* = \mathbf{M}^{-1}\boldsymbol{\omega}$, which is a random vector of latent variables with distribution $N[\mathbf{0}, \mathbf{M}^{-1}\mathbf{\Phi}^+(\mathbf{M}^{-1})^T]$, where $\mathbf{\Phi}^+$ is the covariance matrix of $\boldsymbol{\omega}$. To identify the measurement equation, we have to impose restrictions on $\mathbf{\Lambda}$ and/or $\mathbf{\Phi}^+$, such that the only nonsingular matrix $\mathbf{M}$ that satisfies the imposed conditions is the identity matrix. A simple and common method is using a $\mathbf{\Lambda}$ with the non-overlapping structure. Consider an illustrative example with $p = 10$ observed variables and $q = 3$ latent variables, in which the first four observed variables are related to ω_1, the next and the last three observed variables are related to ω_2 and ω_3, respectively. A non-overlapping structure of $\mathbf{\Lambda}$ is given as follows:

$$\mathbf{\Lambda}^T = \begin{bmatrix} 1 & \lambda_{21} & \lambda_{31} & \lambda_{41} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \lambda_{62} & \lambda_{72} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \lambda_{93} & \lambda_{10,3} \end{bmatrix},$$

where elements 1's and 0's in $\mathbf{\Lambda}$ are known parameters with the fixed values, and the other λ_{jk}'s are unknown parameters. The fixed value 1 is used to introduce a scale to the corresponding latent variable. In the above $\mathbf{\Lambda}$, λ_{11} is fixed as 1 to introduce a scale to ω_1. The choice of λ_{11} is only for convenience, one can fix λ_{21} as 1 and let λ_{11} be an unknown parameter. Similarly, we can fix λ_{62} (or λ_{72}) and λ_{93} (or $\lambda_{10,3}$) as 1, and let λ_{52} and λ_{83} be unknown parameters. Based on the objective of the underlying confirmatory study about the target latent variables, and the meaning of the observed variables, we can have a clear idea about the positions of the parameters fixed as 0; see Section 2.2.1.

There are other methods to identify the measurement equation. For instance, we may allow λ_{11}, λ_{52}, and/or λ_{83} in the above defined $\mathbf{\Lambda}$ to be unknown parameters, and fix the diagonal elements of $\mathbf{\Phi}^+$ as 1. This method restricts the variances of latent variables to be 1; hence $\mathbf{\Phi}^+$ is a correlation matrix. As this method is not convenient for identifying an SEM with a structural equation, and it induces complication in the Bayesian analysis (see

Chapter 3), we use the first identification method to identify the measurement equation throughout this book. After obtaining the estimates of $\mathbf{\Lambda}$ and $\mathbf{\Phi}^+$, say $\hat{\mathbf{\Lambda}}$ and $\hat{\mathbf{\Phi}}^+$, we can get another set of equivalent estimates $\hat{\mathbf{\Lambda}}^*(= \hat{\mathbf{\Lambda}}\mathbf{M})$ and $\hat{\mathbf{\Phi}}^*(= \mathbf{M}^{-1}\hat{\mathbf{\Phi}}^+(\mathbf{M}^{-1})^T)$ that satisfy the same measurement equation via a nonsingular matrix $\mathbf{M}$; see Equation (2.11).

For almost all applications of SEMs, the structural equation is identified with identified $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$. If necessary, the above simple method (via fixing appropriate parameters) for identifying the measurement equation can be used to identify the structural equation.

2.2.5 Path Diagram

A path diagram is a pictorial representation of the measurement and structural equations. It is useful for presenting and discussing the related SEM. In practical applications, it is worthwhile to draw the path diagram related to the hypothesized SEM for effective communication of the basic conceptual ideas under the real study. We first use the example discussed in Sections 2.2.1 and 2.2.2 to illustrate the relation between a path diagram and the measurement and structural equations. The following conventions (see Jöreskog and Sörbom, 1996) are assumed:

(i) Observed variables such as x- and y-variables are enclosed in rectangles or squares. Latent variables such as ξ- and η-variables are enclosed in ellipses or circles. Residual errors such as δ- and ϵ-variables are included in the path diagram but are not enclosed.

(ii) A one-way arrow between two variables indicates a postulated direct influence of one variable on another. A two-way arrow between two variables indicates that these variables may be correlated.

(iii) The coefficient associated with each arrow indicates the corresponding parameter.

(iv) All direct influences of one variable on another are included in the path diagram. Hence the non-existence of an arrow between two variables means that these two variables are assumed not related directly.

Sometimes, two-way arrows between two correlated variables and/or residual errors are not drawn for clarity. Moreover, the means (intercepts) may not be presented in the diagram.

Following the above conventions, the path diagram related to the SEM with measurement equation (2.6) and structural equation (2.7) is presented in Figure 2.1.

Figure 2.1 here

In applying SEMs to a complex study with non-trivial relationships among variables, it is worthwhile to first draw the path diagram that can clearly display the conceptual ideas under the real situation and then formulate the measurement and structural equations of the model on the basis of the path diagram.

2.3 SEMs with Fixed Covariates

In the basic linear SEMs, $\mathbf{y} = (y_1, \cdots, y_p)^T$ on the left-hand side of measurement equation (see Equation (2.3)) contains only observed variables in the model. These observed variables are related to the latent variables in $\boldsymbol{\omega}$ through the loading matrix $\boldsymbol{\Lambda}$. For developing better models, it is often desirable to incorporate explanatory observed variables on the right-hand sides of the measurement and structural equations. In the field of SEM, these explanatory observed variables are regarded as fixed covariates. Accommodation of fixed covariates in the measurement equation provides additional information about the latent exposure and thus reduces estimation uncertainty for the latent variables. For the structural equation, fixed covariates give more ingredients to account for

the outcome latent variables, in addition to the explanatory latent variables. Hence, the residual errors in both equations can be reduced by incorporating fixed covariates.

2.3.1 The Model

SEMs with fixed covariates are defined as follows. For an observed $p \times 1$ random vector $\mathbf{y}$, the measurement equation is given by:

$$\mathbf{y} = \mathbf{A}\mathbf{c} + \boldsymbol{\Lambda}\boldsymbol{\omega} + \boldsymbol{\epsilon}, \tag{2.13}$$

where $\mathbf{A}$ is a $p \times r_1$ matrix of unknown coefficients, $\mathbf{c}$ is an $r_1 \times 1$ vector of fixed covariates (with their values observed), and $\boldsymbol{\Lambda}$, $\boldsymbol{\omega}$, and $\boldsymbol{\epsilon}$ are exactly defined as in Section 2.2. A simple example with one intercept and one fixed covariate c_2 is:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ \vdots & \vdots \\ a_{p1} & a_{p2} \end{bmatrix} \begin{bmatrix} 1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \lambda_{11} & \cdots & \lambda_{1q} \\ \vdots & \ddots & \vdots \\ \lambda_{p1} & \cdots & \lambda_{pq} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_q \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_p \end{bmatrix}, \tag{2.14}$$

or

$$y_j = a_{j1} + a_{j2}c_2 + \lambda_{j1}\omega_1 + \cdots + \lambda_{jq}\omega_q + \epsilon_j, \quad j = 1, \cdots, p.$$

If $c_2 = 0$, and let $\mu_j = a_{j1}$, Equation (2.14) reduces to Equation (2.3). The structural equation is defined by

$$\boldsymbol{\eta} = \mathbf{B}\mathbf{d} + \boldsymbol{\Pi}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\delta}, \tag{2.15}$$

where $\mathbf{B}$ is a $q_1 \times r_2$ matrix of unknown coefficients, $\mathbf{d}$ is an $r_2 \times 1$ vector of fixed covariates, and $\boldsymbol{\Pi}$, $\boldsymbol{\Gamma}$, and $\boldsymbol{\delta}$ are exactly defined as in Section 2.2. Note that $\mathbf{c}$ and $\mathbf{d}$ may have common elements; and (2.15) reduces to (2.8) if $\mathbf{d} = \mathbf{0}$. A simple example with one outcome latent variable η, two covariates d_1 and d_2, three explanatory latent variables ξ_1, ξ_2, and ξ_3, and $\Pi = 0$ is given by:

$$\eta = b_1 d_1 + b_2 d_2 + \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_3 + \delta,$$

where $\mathbf{B} = (b_1, b_2)$, and $\mathbf{\Gamma} = (\gamma_1, \gamma_2, \gamma_3)$.

The assumptions of SEMs with fixed covariates are the same as Assumptions A1, A2, A3, and A4 given in Section 2.2.3. As fixed covariates are observed values, the distributions of $\boldsymbol{\omega}_i$ and $\mathbf{y}_i$ are still normal. Similar to the basic linear SEMs, SEMs with fixed covariates can be identified by fixing appropriate parameters at given values.

2.3.2 An Artificial Example

The purpose of this artificial example is to illustrate the flexibility in incorporating fixed covariates in SEMs. After researchers have decided their main objective, it is desirable to use a path diagram to obtain a clear picture of the hypothesized model. For complicated model with many fixed covariates and latent variables, the path diagram representing the whole model will be rather involved. For clarity, it is worthwhile to use one path diagram to represent the measurement equation, and another path diagram to represent the structural equation. Moreover, if the context is clear, the residual errors can be ignored in the path diagrams.

Suppose that the main objective is on studying the complex diabetic kidney disease, with emphasis on assessing effects of blood pressure, obesity, lipid control as well as some covariates on that disease. Based on some known medical knowledge, data related to the following observed variables: {PCr, ACR, SBP, DBP, BMI, WHR} $= (y_1, \cdots, y_6)$ were collected for forming latent variables 'kidney disease (KD)', 'blood pressure (BP)' and 'obesity (OB)'; see also Section 2.2.1. Data from observed variables {non-high-density lipoprotein cholesterol (non-HDL-C), low-density lipoprotein cholesterol (LDL-C), plasma triglyceride (TG)} $= (y_7, y_8, y_9)$ were collected to measure the latent variable 'lipid control (LIP)'. The relationships of the aforementioned observed and latent variables can be assessed through the measurement equation (2.3) with an appropriate non-overlapping

loading matrix $\mathbf{\Lambda}$. Now, if we know from the real study that smoking and alcohol intake per day may be helpful in relating these observed and latent variables, then 'smoking (c_1)' and 'alcohol (c_2)' can be accommodated in the measurement equation as follows:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \\ a_{51} & a_{52} \\ a_{61} & a_{62} \\ a_{71} & a_{72} \\ a_{81} & a_{82} \\ a_{91} & a_{92} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \lambda_{11} & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ 0 & \lambda_{32} & 0 & 0 \\ 0 & \lambda_{42} & 0 & 0 \\ 0 & 0 & \lambda_{53} & 0 \\ 0 & 0 & \lambda_{63} & 0 \\ 0 & 0 & 0 & \lambda_{74} \\ 0 & 0 & 0 & \lambda_{84} \\ 0 & 0 & 0 & \lambda_{94} \end{bmatrix} \begin{bmatrix} \text{KD} \\ \text{BP} \\ \text{OB} \\ \text{LIP} \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \end{bmatrix}. \tag{2.16}$$

To study various explanatory effects on the key outcome latent variable KD, we can incorporate fixed covariates 'age (d_1)' and 'gender (d_2)' into the structural equation as follows:

$$\text{KD} = b_1\text{age} + b_2\text{gender} + \gamma_1\text{BP} + \gamma_2\text{OB} + \gamma_3\text{LIP} + \delta, \tag{2.17}$$

where b_1 and b_2, γ_1, γ_2, and γ_3 are unknown regression coefficients. The path diagram related to this model is presented in Figure 2.2. Note that the same fixed covariates c_1 and c_2 appear on both the left- and right-hand sides of the path diagram. Moreover, paths related to the residual errors and correlations among latent variables are not presented.

Figure 2.2 here

2.4 Nonlinear SEMs

Nonlinear SEMs are formulated with a measurement equation which is basically the same as in linear SEMs, and a structural equation with nonlinear terms of explanatory latent variables. The theoretical motivation of this generalization is natural; it is similar to the extension of simple regression with latent variables to multiple regression with latent variables. From a practical point of view, the development of nonlinear SEMs is motivated by the fact that nonlinear relations among latent variables are important for establishing more meaningful and correct models in some complicated situations; see Schumacker and Marcoulides (1998) and references therein on the importance of interaction and quadratic effects of latent variables in social and psychological researches. In the biomedical research, the importance of interaction effects has been increasingly recognized. In the study of pathogenesis of complex diseases, it is necessary to consider the gene-gene and gene-environment interactions (Chen *et al.*, 1999). In the case of diabetic kidney disease, there are interactions among glucose, lipid, and haemodynamic pathways in the activation of the renin angiotensin system (Fioretto *et al.*, 1998; Parving, Tarnow and Rossing, 1996).

2.4.1 The Basic Nonlinear SEMs

Let $\mathbf{y}$, $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}$, $\boldsymbol{\epsilon}$, $\boldsymbol{\omega}$, $\boldsymbol{\eta}$, and $\boldsymbol{\xi}$ denote the random vectors and parameters that have the same definitions as in Section 2.2. The measurement equation of nonlinear SEMs is defined as

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\boldsymbol{\omega} + \boldsymbol{\epsilon}, \tag{2.18}$$

which has exactly the same form as in (2.3). The structural equation is formulated as

$$\boldsymbol{\eta} = \boldsymbol{\Pi}\boldsymbol{\eta} + \boldsymbol{\Gamma}\mathbf{F}(\boldsymbol{\xi}) + \boldsymbol{\delta}, \tag{2.19}$$

where the definition of $\mathbf{\Pi}$, $\mathbf{\Gamma}$, and $\boldsymbol{\delta}$ are the same as before, and $\mathbf{F}(\boldsymbol{\xi}) = (f_1(\boldsymbol{\xi}), \cdots, f_t(\boldsymbol{\xi}))^T$ is a $t \times 1$ vector-valued function with nonzero, known, and linearly independent differentiable functions $f_1, \cdots, f_t$, and $t \geq q_2$.

Let $\{\mathbf{y}_i,\ i = 1, \cdots, n\}$ be the observed data set with a sample size n, where $\mathbf{y}_i$ are modeled via a nonlinear SEM with latent variables $\boldsymbol{\eta}_i$ and $\boldsymbol{\xi}_i$, and measurement errors $\boldsymbol{\epsilon}_i$ and $\boldsymbol{\delta}_i$. The assumptions of the model for developing related statistical methods are the same as Assumptions A1, A2, A3, and A4 given in Section 2.3.2. Note that, due to the presence of the nonlinear terms of $\boldsymbol{\xi}$ in $\mathbf{F}(\boldsymbol{\xi})$, the distributions of $\boldsymbol{\omega}_i$ and $\mathbf{y}_i$ are no longer normal. In other words, nonlinear SEMs do not assume that $\boldsymbol{\omega}_i$ and $\mathbf{y}_i$ are normal.

Similar to linear SEMs, the measurement equation of nonlinear SEMs can be identified by fixing appropriate parameters (particularly those in $\mathbf{\Lambda}$) at some given values. To achieve an identified structural equation, the choice of $\mathbf{F}(\boldsymbol{\xi})$ in the structural equation is not completely arbitrary. For example, the following obvious cases are not allowed: $\mathbf{F}_1(\boldsymbol{\xi}) = (\xi_1, \xi_2, \xi_1^2, \xi_1^2)^T$ and $\mathbf{F}_2(\boldsymbol{\xi}) = (\xi_1, \xi_2, \xi_1\xi_2, 0)^T$. They should be modified as $\mathbf{F}_1(\boldsymbol{\xi}) = (\xi_1, \xi_2, \xi_1^2)^T$ and $\mathbf{F}_2(\boldsymbol{\xi}) = (\xi_1, \xi_2, \xi_1\xi_2)^T$, respectively. An example of identified structural equations is :

$$\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} 0 & \pi \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} & 0 & 0 & 0 \\ \gamma_{21} & \gamma_{22} & \gamma_{23} & \gamma_{24} & \gamma_{25} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_1^2 \\ \xi_1\xi_2 \\ \xi_2^2 \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix}.$$

The interpretation of the parameter matrices $\mathbf{\Lambda}$, $\mathbf{\Pi}$, and $\mathbf{\Gamma}$ is basically the same as before; that is, they can be interpreted as regression coefficients in regression models. More care is needed to interpret the mean vector of $\mathbf{y}$. Let $\mathbf{\Lambda}_k^T$ be the kth row of $\mathbf{\Lambda}$. For $k = 1, \cdots, p$,

it follows from Equation (2.3) that $E(y_k) = \mu_k + \mathbf{\Lambda}_k^T E(\boldsymbol{\omega})$. Although $E(\boldsymbol{\xi}) = \mathbf{0}$, it follows from Equation (2.19) that $E(\boldsymbol{\eta}) \neq \mathbf{0}$ if $\mathbf{F}(\boldsymbol{\xi})$ is a vector-valued nonlinear function of $\boldsymbol{\xi}$ and $E(\mathbf{F}(\boldsymbol{\xi})) \neq 0$. Hence $E(\boldsymbol{\omega}) \neq \mathbf{0}$ and $E(y_k) \neq \mu_k$. Let $\mathbf{\Lambda}_k^T = (\mathbf{\Lambda}_{k\eta}^T, \mathbf{\Lambda}_{k\xi}^T)$ be a partition of $\mathbf{\Lambda}_k^T$ that corresponds to the partition of $\boldsymbol{\omega} = (\boldsymbol{\eta}^T, \boldsymbol{\xi}^T)^T$. Because $E(\boldsymbol{\xi}) = \mathbf{0}$ and $E(\boldsymbol{\eta}) = [(\mathbf{I} - \mathbf{\Pi})^{-1}\mathbf{\Gamma}]E(\mathbf{F}(\boldsymbol{\xi}))$, it follows from Equations (2.18) that

$$E(y_k) = \mu_k + \mathbf{\Lambda}_{k\eta}^T E(\boldsymbol{\eta}) + \mathbf{\Lambda}_{k\xi}^T E(\boldsymbol{\xi}) = \mu_k + \mathbf{\Lambda}_{k\eta}^T [(\mathbf{I} - \mathbf{\Pi})^{-1}\mathbf{\Gamma}]E(\mathbf{F}(\boldsymbol{\xi})). \tag{2.20}$$

As $\mathbf{F}(\boldsymbol{\xi})$ is usually not very complicated in most practical applications, $E(\mathbf{F}(\boldsymbol{\xi}))$ is not very complex and thus the computation of $E(y_k)$ is not difficult.

2.4.2 Nonlinear SEMs with Fixed Covariates

Linear SEMs with fixed covariates can be naturally generalized to nonlinear SEMs with fixed covariates through the following measurement and structural equations:

$$\mathbf{y} = \mathbf{A}\mathbf{c} + \mathbf{\Lambda}\boldsymbol{\omega} + \boldsymbol{\epsilon}, \tag{2.21}$$

$$\boldsymbol{\eta} = \mathbf{B}\mathbf{d} + \mathbf{\Pi}\boldsymbol{\eta} + \mathbf{\Gamma}\mathbf{F}(\boldsymbol{\xi}) + \boldsymbol{\delta}, \tag{2.22}$$

where the definitions of the random vectors and the parameter matrices are the same as in Sections 2.2 and 2.3. In this model, the measurement equation is the same as in (2.13), while the structural equation can be regarded as a natural extension of equations (2.15) and (2.19). As a simple example, we consider a continuation of the artificial example as described in Section 2.3.2. Suppose that we wish to study various interactive effects of the explanatory latent variables BP, OB, and LIP on KD. To achieve our goal, we consider a model with its measurement equation given in (2.16), while the structural equation is formulated as

$$\text{KD} = b_1 d_1 + b_2 d_2 + \gamma_1 \text{BP} + \gamma_2 \text{OB} + \gamma_3 \text{LIP} + \gamma_4 (\text{BP} \times \text{OB}) + \gamma_5 (\text{BP} \times \text{LIP}) + \gamma_6 (\text{OB} \times \text{LIP}) + \delta. \tag{2.23}$$

In this formulation, $\mathbf{B} = (b_1, b_2)$, $\mathbf{d} = (d_1, d_2)^T$, $\mathbf{\Gamma} = (\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6)$, and $\mathbf{F}(\boldsymbol{\xi}) = (\text{BP}, \text{OB}, \text{LIP}, \text{BP} \times \text{OB}, \text{BP} \times \text{LIP}, \text{OB} \times \text{LIP})^T$.

While the structural equation (2.22) can include nonlinear terms of $\boldsymbol{\xi}$ in predicting $\boldsymbol{\eta}$, it does not accommodate nonlinear terms of $\boldsymbol{\xi}$ and $\mathbf{d}$ simultaneously. A simple extension of the structural equation to cope with the above consideration is:

$$\boldsymbol{\eta} = \mathbf{\Pi}\boldsymbol{\eta} + \mathbf{\Lambda}_\omega \mathbf{G}(\mathbf{d}, \boldsymbol{\xi}) + \boldsymbol{\delta}. \tag{2.24}$$

where $\mathbf{G}(\mathbf{d}, \boldsymbol{\xi}) = (g_1(\mathbf{d}, \boldsymbol{\xi}), \cdots, g_t(\mathbf{d}, \boldsymbol{\xi}))^T$ is a vector-valued function with nonzero, known, and linearly independent differentiable functions. A special case of this general structural equation is the one defined by (2.22) with $\mathbf{\Lambda}_\omega = (\mathbf{B}, \mathbf{\Gamma})$ and $\mathbf{G}(\mathbf{d}, \boldsymbol{\xi}) = (\mathbf{d}^T, \mathbf{F}(\boldsymbol{\xi})^T)^T$. The assumptions of this nonlinear SEM with nonlinear terms of covariates are the same as the Assumptions A1, A2, A3, and A4 given in Section 2.2.3. Moreover, identification of the model can be achieved via the similar method as described in Section 2.2.4. Again, care should be taken to interpret the mean of $\mathbf{y}$. Using the same notation as in Section 2.4.1, we have

$$E(y_k) = \mathbf{A}_k^T \mathbf{c} + \mathbf{\Lambda}_{k\eta}^T [(\mathbf{I} - \mathbf{\Pi})^{-1} \mathbf{\Lambda}_\omega] E[\mathbf{G}(\mathbf{d}, \boldsymbol{\xi})], \tag{2.25}$$

where $\mathbf{A}_k^T$ is the kth row of $\mathbf{A}$. The artificial example presented above is used again to illustrate the key idea of incorporating nonlinear terms of fixed covariates and explanatory latent variables in the structural equation. While the measurement equation is again defined by (2.16), and the structural equation can be formulated as follows:

$$\begin{aligned} \text{KD} = {} & b_1 d_1 + b_2 d_2 + \gamma_1 \text{BP} + \gamma_2 \text{OB} + \gamma_3 \text{LIP} + \gamma_4 (\text{BP} \times \text{OB}) + \gamma_5 (\text{BP} \times \text{LIP}) \\ & + \gamma_6 (\text{OB} \times \text{LIP}) + \gamma_7 (d_1 \times \text{BP}) + \gamma_8 (d_1 \times \text{OB}) + \gamma_9 (d_2 \times \text{OB}) + \gamma_{10} (d_2 \times \text{LIP}) + \delta. \end{aligned} \tag{2.26}$$

Note that more complex product terms of d_1, d_2, BP, OB, and LIP can be assessed via other appropriately defined structural equations. The path diagram corresponding to

the structural equation (2.26) is presented in Figure 2.3. For clarity, and because the measurement equation is the same as before, Figure 2.3 does not include the observed variables in $\mathbf{y}$, and fixed covariates c_1 and c_2.

Figure 2.3 here

In general, as the fixed covariates in $\mathbf{d}$ may come from arbitrary distributions that result in continuous or discrete data, and the functions f_j in $\mathbf{G}(\mathbf{d}, \boldsymbol{\xi})$ are differentiable functions which include any product terms as special cases, the nonlinear SEM defined by (2.21) and (2.24) (or (2.22)) can handle a wide range of situations.

2.4.3 Remarks

Due to the non-normal distributions associated with the nonlinear terms of latent variables, development of statistical methods for inference based on the traditional approach of covariance structure analysis encounters certain difficulties. In the past years, several approaches have been proposed; examples are the product-indicator method (Jaccard and Wan, 1995; among others), the moment-based method (Wall and Amemiya, 2000), and the latent moderated structural equations (Klein and Moosbrugger, 2000) approaches. Most of these approaches used some unnatural or technically involved techniques to handle the non-normality problem. Hence, nonlinear SEMs have been regarded as complicated models. In the next chapter, we will discuss the application of the Bayesian approach to the analysis of SEMs. Using the Bayesian approach, the development of statistical methods for analyzing nonlinear SEMs is essentially the same as linear SEMs. Hence, we regard nonlinear SEMs as standard models rather than complicated models.

2.5 Discussion and Conclusions

Real analyses of most complicated studies (for example, complex diseases) usually involve a large number of observed and latent variables of interest. Although the emphasis is on assessing the effects of explanatory latent variables on the key outcome latent variables, some particular explanatory latent variables may be significantly related to other explanatory latent variables and/or fixed covariates. For instance, in the artificial example discussed in Section 2.3.2, although KD is the key outcome latent variable, the explanatory latent variables BP and OB, as well as the fixed covariates, age (d_1) and gender (d_2), are also expected to have effects on the latent variable LIP. To provide a more concrete illustration, suppose that we are interested in assessing the SEM with the corresponding path diagram presented in Figure 2.4. Compared to the SEM presented in Figure 2.2, we further have: (i) two additional observed genetic variables GV1 and GV2 which correspond to a latent variable LGV, (ii) a path from LGV to LIP, (iii) a path from OB to LIP, and (iv) a path from age (d_1) to LIP.

Figure 2.4 here

The measurement equation is defined by:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ GV_1 \\ GV_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \\ a_{51} & a_{52} \\ a_{61} & a_{62} \\ a_{71} & a_{72} \\ a_{81} & a_{82} \\ a_{91} & a_{92} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \lambda_{11} & 0 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 & 0 \\ 0 & \lambda_{32} & 0 & 0 & 0 \\ 0 & \lambda_{42} & 0 & 0 & 0 \\ 0 & 0 & \lambda_{53} & 0 & 0 \\ 0 & 0 & \lambda_{63} & 0 & 0 \\ 0 & 0 & 0 & \lambda_{74} & 0 \\ 0 & 0 & 0 & \lambda_{84} & 0 \\ 0 & 0 & 0 & \lambda_{94} & 0 \\ 0 & 0 & 0 & 0 & \lambda_{10,5} \\ 0 & 0 & 0 & 0 & \lambda_{11,5} \end{bmatrix} \begin{bmatrix} \text{KD} \\ \text{BP} \\ \text{OB} \\ \text{LIP} \\ \text{LGV} \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \\ \epsilon_{11} \end{bmatrix}. \quad (2.27)$$

To formulate the structural equation associated with this path diagram, LIP is treated as an outcome latent variable. More specifically, the structural equation is defined as:

$$\begin{pmatrix} \text{KD} \\ \text{LIP} \end{pmatrix} = \begin{pmatrix} b_1 & b_2 \\ b_3 & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} + \begin{pmatrix} 0 & \pi_1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \text{KD} \\ \text{LIP} \end{pmatrix} + \begin{pmatrix} \gamma_1 & \gamma_2 & 0 \\ 0 & \gamma_3 & \gamma_4 \end{pmatrix} \begin{pmatrix} \text{BP} \\ \text{OB} \\ \text{LGV} \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix}. \quad (2.28)$$

Hence, through an appropriate formulation of the structural equation, SEMs provide considerable flexibility in assessing various kinds of relationships among latent variables and covariates. Here, as OB has a direct effect on LIP, which itself also has a direct effect on KD, OB has an indirect effect ($\pi_1 \times \gamma_3$) on KD. Hence, the total effect of OB on KD is $\gamma_2 + \pi_1\gamma_3$. Given the flexibility in formulating different structural equations, the above modeling concerning LIP can be considered for explanatory latent variables BP and OB.

In general, latent variables involved in the studies of complex situations commonly have interrelationships with other observed and latent variables similar to those discussed above. There is a temptation to develop a comprehensive SEM that takes into account all the interrelationships of the observed and latent variables. From a practical point of view, the following issues have to be carefully considered in the development of a comprehensive model:

(i) Due to the large amount of observed variables and latent variables, the size of the comprehensive SEM and the number of its unknown parameters are large. It is important to make sure the sample size of the available data set is large enough to achieve accurate statistical results.

(ii) The data structures associated with complex situations are usually complicated. Researchers may encounter mixed types of continuous, ordered and unordered categorical data with missing entries, and hierarchical or heterogeneous structures. Although the Bayesian methods (will be introduced in next chapter) are efficient and can be applied to handle those complex data, if the size of the proposed SEM and the number of parameters are large, we may encounter difficulties in achieving convergence of the related computing algorithm for obtaining statistical results.

(iii) Researchers need to make sure that the hypothesized model can be analyzed under the proposed SEM framework. In this chapter, the most general SEM is the nonlinear SEM with fixed covariates defined by Equations (2.21) and (2.24). This model has limitations. Note that the vector-valued function $\mathbf{G}(\mathbf{d}, \boldsymbol{\xi})$ for assessing nonlinear terms of $\mathbf{d}$ and $\boldsymbol{\xi}$ does not involve any outcome latent variables in $\boldsymbol{\eta}$. Hence, once a latent variable is treated as an outcome latent variable, nonlinear

terms of this latent variable and interactive effects between this latent variable and other explanatory latent variables (or fixed covariates) cannot be used for predicting the other outcome latent variables. For instance, consider the previous artificial example related to the SEM with the structural equation (2.28). In this model, as LIP is treated as an outcome variable, it is an element in $\boldsymbol{\eta}$. Hence, it cannot be accommodated in $\mathbf{G}(\mathbf{d}, \boldsymbol{\xi})$, and nonlinear effects of LIP on the key outcome variable KD cannot be assessed.

Ideally, establishing a comprehensive model that simultaneously takes into account all the interrelationships among all observed variables, latent variables, as well as fixed covariates is desirable. Practically, however, if the comprehensive model involves a large number of variables and unknown parameters, and the underlying data structure is complex, one may encounter serious difficulties in developing such a comprehensive model. Under these situations, it is advantageous to separate the comprehensive model into submodels, and then conduct the SEM analysis for each submodel. Of course, among the submodels, the one that involves the key outcome variables of main interest is the most important. For example, in the analysis of the above artificial example, the comprehensive SEM that treats LIP as an outcome latent variable and accommodates its nonlinear effects in the structural equation associated with the path diagram in Figure 2.4 can be separated into two submodels. One is the SEM represented by the measurement equation (2.16), and the structural equation (2.26) with main focus on the outcome latent variable KD. In this submodel, LIP is only treated as an explanatory latent variable (no paths from d_1, BP, and OB to LIP, see Figure 2.3). Hence, its nonlinear effects BP $\times$ LIP, OB $\times$ LIP, and $d_2\times$ LIP on KD can be assessed. To further assess the relationships among LIP and the other covariates and/or latent variables, the following submodel is used.

The measurement equation is defined by:

$$\begin{bmatrix} y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ \text{GV}_1 \\ \text{GV}_2 \end{bmatrix} = \begin{bmatrix} a_{51} & a_{52} \\ a_{61} & a_{62} \\ a_{71} & a_{72} \\ a_{81} & a_{82} \\ a_{91} & a_{92} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \lambda_{53} & 0 & 0 \\ \lambda_{63} & 0 & 0 \\ 0 & \lambda_{74} & 0 \\ 0 & \lambda_{84} & 0 \\ 0 & \lambda_{94} & 0 \\ 0 & 0 & \lambda_{10,5} \\ 0 & 0 & \lambda_{11,5} \end{bmatrix} \begin{bmatrix} \text{OB} \\ \text{LIP} \\ \text{LGV} \end{bmatrix} + \begin{bmatrix} \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \\ \epsilon_{11} \end{bmatrix}. \tag{2.29}$$

The structural equation is given by

$$\text{LIP} = b_3 d_1 + \gamma_4 \text{OB} + \gamma_5 \text{LGV} + \delta. \tag{2.30}$$

Figure 2.5 here

A path diagram representing the submodel defined by (2.29) and (2.30) is given in Figure 2.5. Here, LIP is treated as an outcome latent variable. Based on the estimates of the path coefficients γ_3 in the submodel associated with Figure 2.3, and γ_4 in the submodel defined by Equations (2.29) and (2.30) (see also Figure 2.5), we can obtain an idea about the indirect effect of OB on KD via $\hat{\gamma}_3 \times \hat{\gamma}_4$. However, as $\hat{\gamma}_3$ and $\hat{\gamma}_4$ are not simultaneously estimable through a single model, the estimate of this indirect effect is not optimal and should be interpreted with care. It should also be noted that there are two sets of estimates for the common parameters in these two submodels. One set of estimates is obtained through the analysis of the submodel defined by (2.16) and (2.26), while the other set of estimates is obtained through analysis of the submodel defined by (2.29) and (2.30). However, the differences between these two sets of estimates are very small; and in practice, they would not result in different interpretations of the results. Hence, this

issue is not important. The trade-off of these disadvantages in using submodels is the possibility of assessing various nonlinear effects in relation to LIP through the submodel associated with Figure 2.3. In practice, the choice between the different approaches in applying SEMs heavily depends on the objective of the real study. For instance, if assessing the nonlinear effect of LIP on KD is more important, it may be worthwhile to use submodels in the analysis.

In the analysis of SEMs, sound statistical methods that seriously take into consideration of the structures of the hypothesized model and data should be used. Parameter estimates should be obtained via valid statistical procedures. It is necessary to develop rigorous goodness-of-fit and model comparison statistics for assessing goodness-of-fit of the hypothesized model to the sample data and for comparing competing models. In the next chapter, we will introduce the Bayesian approach with optimal statistical properties for estimation, goodness-of-fit analysis, and model comparison.

In this chapter, we discuss the basic SEMs, namely the linear SEMs and nonlinear SEMs with fixed covariates, and their applications to the analysis of practical data. While these SEMs are helpful for analyzing real data sets related to biomedical research, they have limitations which are mainly induced by their underlying assumptions. There is a need to develop more subtle generalizations of these models to overcome the limitations. Based on the demand from substantive research, certain generalizations will be discussed in subsequent chapters. Given the developments of the basic models and their generalizations, SEMs provide efficient tools with great flexibility for analyzing multivariate data in behavioral, educational, medical, social, and psychological sciences.

References

Bollen, K. A. (1989) *Structural Equation Models with Latent Variables*. NJ: John Wiley & Sons, Inc.

Chen, C. C., Lu, R. B., Chen, Y. C., Wang, M. F., Chang, Y. C., Li, T. K. and Yin, S. J. (1999) Interaction between the functional polymorphisms of the alcohol-metabolism genes in protection against alcoholism. *The American Journal of Human Genetics*, **65**, 795-807.

Fioretto, P., Steffes, M. W., Sutherland, D. E. R., Goetz, F. C. and Mauer, M. (1998) Reversal of lesions of diabetic nephropathy after pancreas transplantation. *The New England Journal of Medicine*, **339**, 69-75.

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

Jöreskog, K. G. and Sörbom, D. (1996) LISREL 8: *Structural Equation Modeling with the SIMPLIS Command Language*. Scientific Software International.

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

Parving, H. H., Tarnow, L. and Rossing, P. (1996) Genetics of diabetic nephropathy. *Journal of the American Society of Nephrology*, **7**, 2509-2517.

Schumacker, R. E. and Marcoulides, G. A. (1998) *Interaction and Nonlinear Effects in Structural Equation Modeling*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Wall, M. M. and Amemiya, Y. (2000) Estimation for polynomial structural equation modeling. *Journal of the American Statistical Association*, **95**, 929-940.

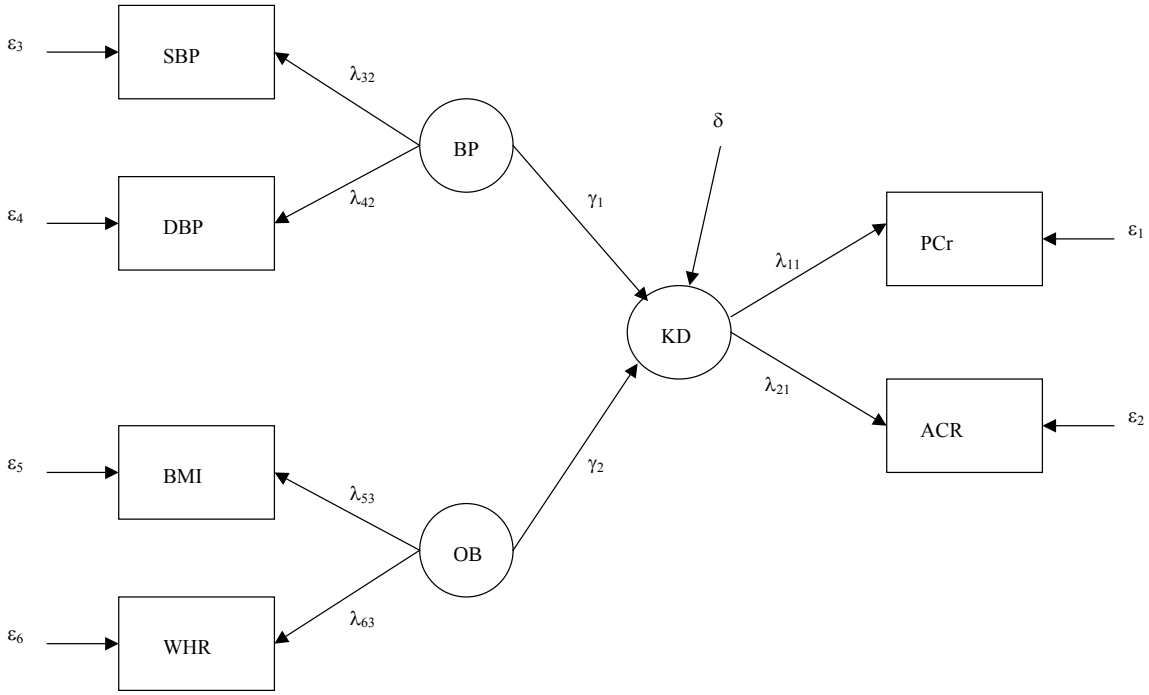

Figure 2.1: Path diagram representing the model defined by (2.6) and (2.7).

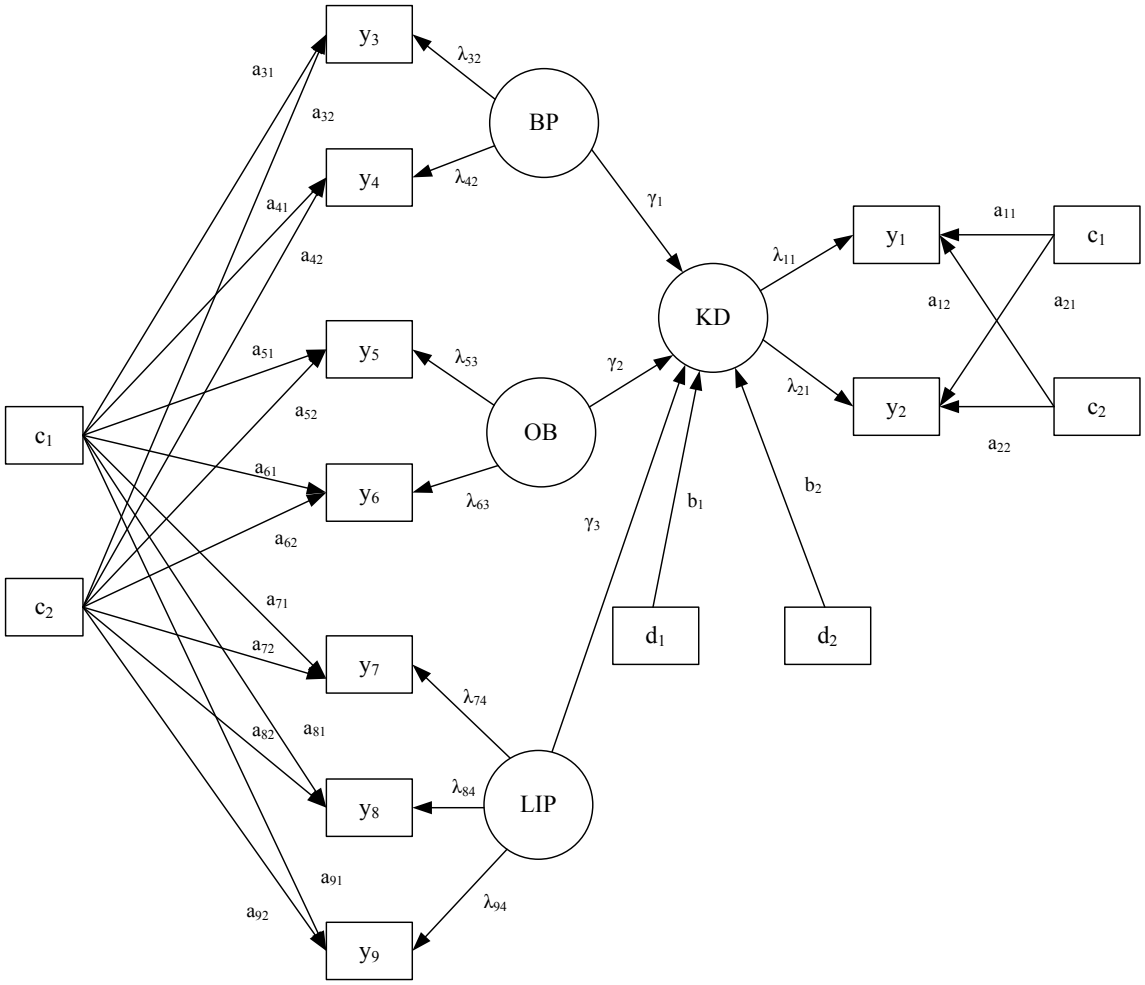

Figure 2.2: Path diagram representing the model defined by (2.16) and (2.17). Here, the residual errors are not drawn.

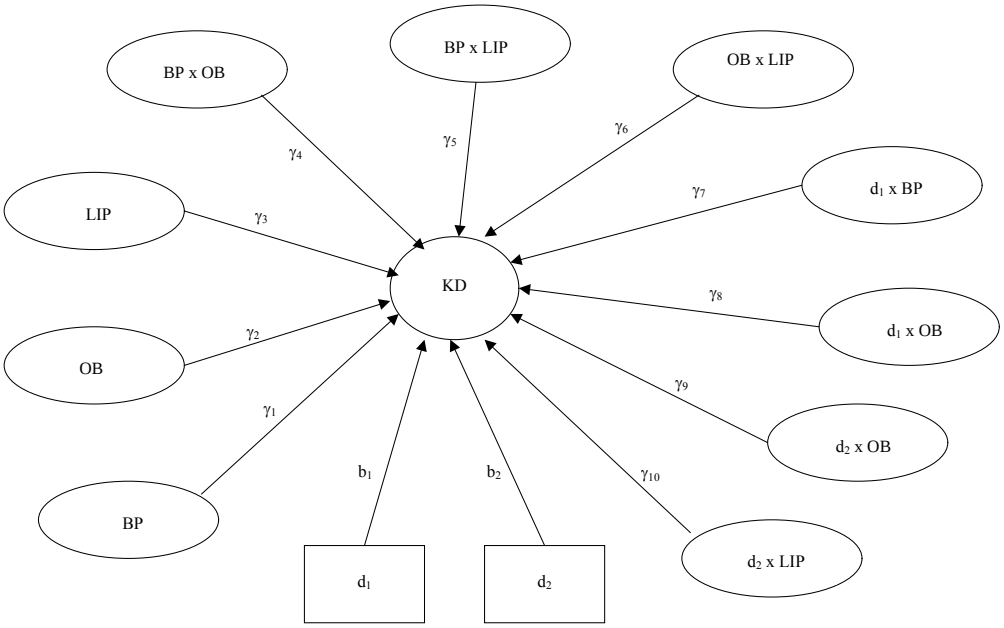

Figure 2.3: Path diagram representing the structural equation (2.26). Here, the residual error is not drawn.

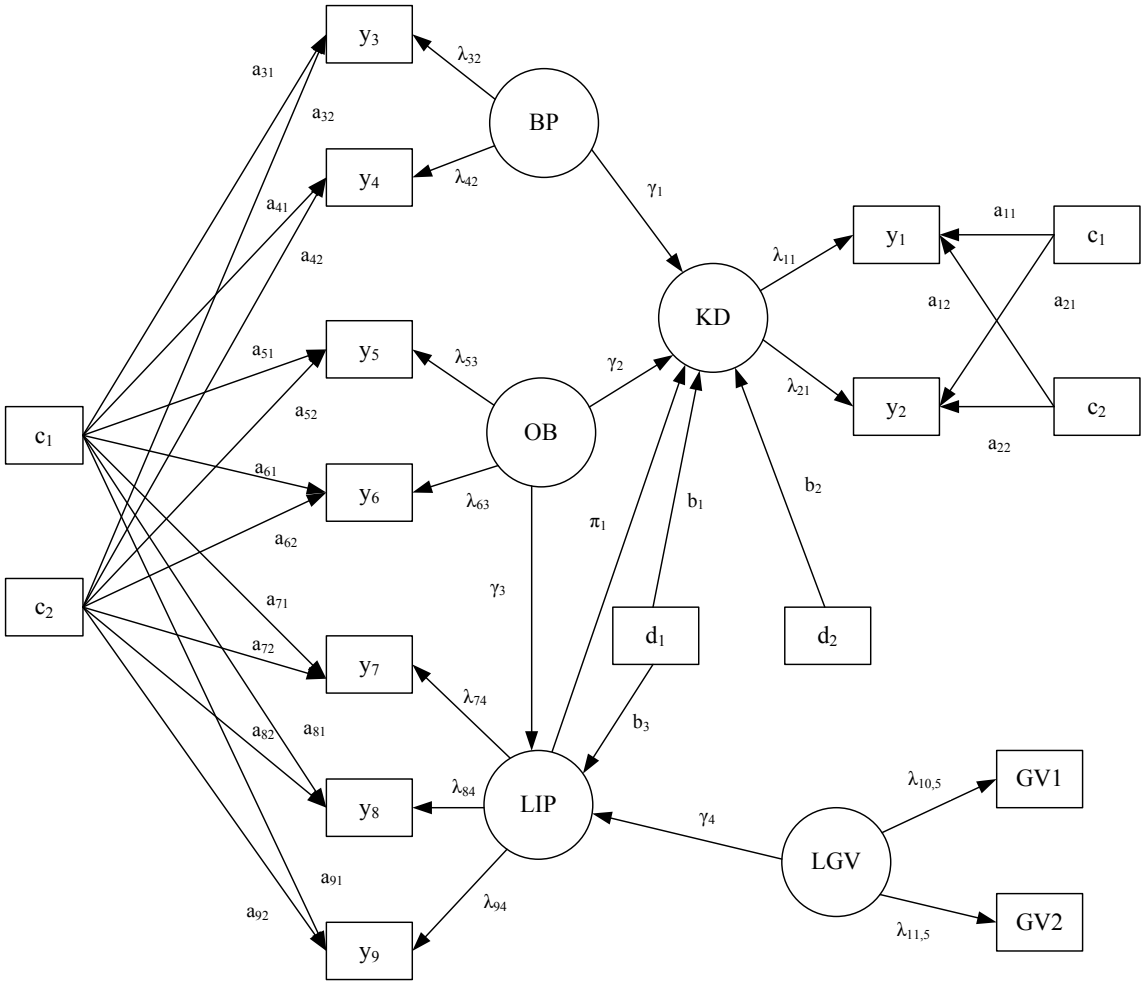

Figure 2.4: Path diagram representing the model defined by (2.27) and (2.28). Here, the residual errors are not drawn.

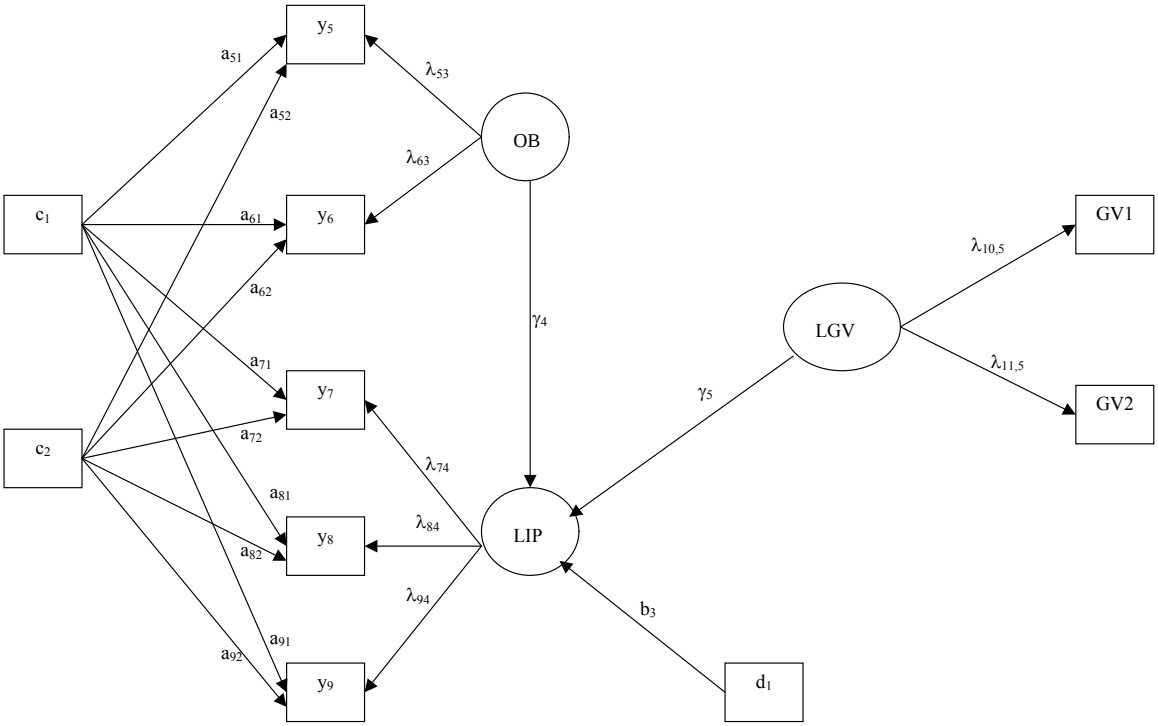

Figure 2.5: Path diagram representing submodel defined by (2.29) and (2.30). Here, the residual errors are not drawn.