Appendix A

Notation

Word count:

The structural modeling notation system I use for this book follows a traditional system that uses Greek symbols for each of the parameters based on a matrix organizing principle, usually referred as LISREL notation, short for "LInear Structural RELations” (Frisch & Waugh, 1933). The general SEM model and notation system is perhaps more accurately called the JKW model, after the authors credited with synthesizing and expanding decades of prior work on path analysis and factor analysis into a highly generalizable structure equation framework (Jöreskog, 1973; Keesling, 1972; Wiley, 1973). The LISREL term has become primarily associated with the software developed by Jöreskog and Sörbom (1974), but the notation system has become widely applied regardless of the SEM software package used.[1] I use LISREL notation throughout the book for one very important reason. A large majority of statistical articles about SEM use this notation. Many introductory textbooks now avoid LISREL notation in order to increase accessibility, which is indeed an objective I sympathize with. For those who wish to learn more about SEM after an initial introduction, however, unfamiliarity with the LISREL notation system leaves readers with what I believe to be a serious literacy gap.

Although LISREL notation is tied to matrix algebra, it is really not necessary to know matrix algebra to read and understand this book. Matrix algebra, a kind of short hand system that can be used for manipulating many simultaneous equations, is convenient for describing the linear regression equations used in SEM. Learning the Greek symbols associated with the LISREL notation is a separate matter from understanding matrix algebra and is, at least, an initial

[1] The LISREL notation system is neither universal nor necessary. There are many minor variations and several major alternative notation systems. Most notably, the Bentler-Weeks (Bentler & Weeks, 1980) system, which is associated with EQS software, is another matrix-based notation alternative.

step. I do encourage the reader to learn matrix algebra to increase the understanding of this topic and to better understand some of the mathematical underpinnings of SEM. There are a fairly limited number of definitions and simple algebra rules that can be absorbed with a small investment in effort. I do not provide an introduction to matrix algebra with this text, because there are many excellent introductions (e.g., Bollen, 1989; Hayduk, 1987; Mulaik, 2009; Namboodiri, 1984).

"All-y" LISREL Notation

Most of the formulas in this book use an abbreviated version of the full LISREL notation, which is commonly used by authors and easier to learn. The full LISREL notation system (described later in this appendix) distinguishes between exogenous variables and endogenous variables. Exogenous variable are not caused by other variables in the model and endogenous variables are those caused by other variables in the model.

Measurement Model Parameters

Table A.1 is a summary of all of the Greek symbols used in the LISREL model. Each latent variable is designated by η ("eta"). In this text, I will index latent variables with subscript *k*, up to a total of *K* latent variables in the model. Observed variables will be indexed with *j*, and there are *J* total observed variables in model. Loadings, λ_{jk} ("lambda"), represent a regression of a measured variable y_j on factor η_k, using subscript *jk* to indicate the *j*th observed variable is predicted by the *k*th latent variable. The "effect" always precedes the "cause" in the order of subscripts for loadings (and structural paths). The intercept in this regression is ν ("nu") with subscript *j*. We can then write an equation for a simple regression that represents the relation of the observed variable to the factor.

$$y_j = \nu_j + \lambda_{jk}\eta_k + \varepsilon_j$$

We could add an index *i* representing an individual case in the data set for the observed variable y_{ji}, the latent variable η_{ki}, and the measurement residual (or error term) ε_{ij}, but I omit this in most instances to simplify the notation as much as possible.

The individual parameters are organized into matrices or vectors, matrices with one column (or, if transposed, a row). Capital letters (bolded in this text) are used for to represent each matrix. Loadings are organized into a matrix with *J* rows and *K* columns, said to be of dimension $J \times K$. Each entry in the loading matrix represents the intersection of an observed variable and a factor. For example, a model with two latent variables with three indicators loading on each factor would be a 6×2 $\mathbf{\Lambda}$ matrix, with rows corresponding to observed variables y_1, y_2, y_3, y_4, y_5, and y_6, and columns corresponding to latent variables η_1 and η_2,

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \end{bmatrix}$$

The 0s show that the indicator the loading is set equal to 0, or, in other words, does not load on that factor

The measurement residuals are organized into a square matrix $\mathbf{\Theta}$ ("theta") with the diagonal elements representing the variances $Var(\varepsilon_{jj}) = \theta_{jj}$, and the off-diagonal elements representing covariances among measurement residuals [e.g., $Cov(\varepsilon_1,\varepsilon_2) = \theta_{12}$]. The two-factor example mentioned above might have a 6×6 $\mathbf{\Theta}$ matrix that looks like the following, if each of the measurement residuals were freely estimated and one covariance between y_1 and y_2 was estimated.

$$\boldsymbol{\Theta} = \begin{bmatrix} \theta_{11} & \theta_{12} & 0 & 0 & 0 & 0 \\ \theta_{21} & \theta_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \theta_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & \theta_{66} \end{bmatrix}$$

The observed variables and the latent variables are vectors (single-column matrix) with J and K rows, respectively. The factor variances and covariances have the symbol ψ ("psi") and are appear in the $\boldsymbol{\Psi}$ matrix, which is square with dimension $K \times K$. The diagonal elements are the variances and the off-diagonal elements are the covariances. If the two-factor example estimated both factor variances and the covariance, the $\boldsymbol{\Psi}$ matrix would be

$$\boldsymbol{\Psi} = \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{bmatrix}$$

The measurement model states the covariance matrix in terms of these matrices.

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Lambda\Psi\Lambda}' + \boldsymbol{\Theta}$$

The prime symbol $'$ indicates the $\boldsymbol{\Lambda}$ matrix is transposed (rows and columns are switched).

Factor means, α_k ("alpha") and measurement intercepts, ν_j, are not included in the measurement equation above, but they can be added to the model. Each is a vector of the same name in the matrix system, $\boldsymbol{\alpha}$ and $\boldsymbol{\nu}$, respectively.

Structural Model Parameters

The structural portion of the model involves paths between latent variables, represented by β ("beta"). Although β is used for standardized coefficients in some regression texts, they represent unstandardized coefficients here. I will use β^* for standardized coefficients. The order of the subscripts is such that dependent or "effect" variable precedes the predictor or "cause" variable. For example, a path for η_2 predicted by η_1 would be labeled β_{21}. The path coefficients are organized into the **B** matrix, with the dimensions $K \times K$. Naturally, many of the elements will be 0 in practice, because usually only one direction can be estimated in practice. Disturbances (residuals, errors) in the structural model are represented by ζ ("zeta"). Disturbances may appear as a vector of individual parameters, $\boldsymbol{\zeta}$. Because dependent variables have only conditional variances, the variance of the disturbances diagonal elements in the $\boldsymbol{\Psi}$ matrix, where $Var(\zeta) = \psi$. Covariances of disturbances, $Cov(\zeta,\zeta)$, are off-diagonal elements in the $\boldsymbol{\Psi}$ matrix.

The formal LISREL notation system assumes only structural relations among latent variables not between observed variables or between latent variables and observed variables. Each observed variable must be an indicator of a latent variable, even if there is only one indicator per latent variable (identified by setting the loading equal to 1 and the measurement residual equal to 0). Most SEM software packages, however, allow structural paths between measured variables and measured and latent variables. As consequence many articles and texts allow structural paths directly between observed variables and latent variables. This convenience has no impact on the underlying mathematics, however. I therefore use x and y in equations when there are structural relations among them (predictive paths or correlations) and show them within squares with direct relations to latent or other observed variables in Figures.

Path Diagrams

Figure A.1 summarizes the notation in the depiction of one possible model. Notice that when variable numbers have two digits, a comma is used to separate subscript number pairs (e.g., $\lambda_{15,4}$). I follow most of the usual path diagram conventions for structural models. One exception is that I do not represent means and intercepts as triangles in the diagrams as in the RAM diagram approach (McArdle & MacDonald, 1984), primarily to simplify the diagrams of some of the rather complex models in some chapters. Instead, when means or intercepts are estimated in the model, I depict this by placing the symbol next to the ellipse (latent means and intercepts) or rectangle (measurement intercept).[2] For variances, ψ_{kk} appears next to the circle or rectangle in a similar fashion.

When a parameter is to be set to a specific value, such as 0 or 1, the number appears in the diagram in the location the parameter normally appears. Mean or intercept values are in square brackets to distinguish them. Figure A.1 illustrates the use of specific values, where the first loading is set equal to 1 and the measurement intercept set equal to 0, shown as [0]. These are commonly used values for the referent or marker method of identifying the factor variance and mean.

Full Matrix Notation

The "all-y" notation is commonly used by authors, but its use is not universal. The original and more formal LISREL system involves separate matrices for exogenous and endogenous parameters. *Exogenous* variables are those not predicted by other variables in the model and *endogenous* variables are those predicted by other variables. This distinction usually can be dropped without loss of generality, but the full notation is needed for clarity in some instances. To simplify as much as possible, I use the "all-y" system whenever possible.

[Table A.2 about here]

Table A.2 summarizes all the symbols and matrices used for the full LISREL notation. Each parameter has a separate notation for the parameter and accompanying matrix depending on the role of the variable as exogenous or endogenous in the model. Even when observed variables, *x*, are used as indicators of exogenous variables, and are therefore predicted by another variable, they are still considered exogenous in the model under the notation system. The

[2] I borrowed this convention from my colleague, Rich Jones.

symbols used for the endogenous parameters, sometimes with added y subscript, are the same as in the "all-y" system, but exogenous parameters and matrices use the following symbols: ϕ and $\mathbf{\Phi}$ for latent variable variances and covariances, λ_x and $\mathbf{\Lambda}_x$ for loadings, θ_δ and $\mathbf{\Theta}_\delta$ for measurement residual variances and covariances, κ and $\mathbf{K}$ for latent variable means, and ν_x and $\boldsymbol{\nu}_x$ for measurement intercepts. Figure A.2 depicts the full LISREL notation version of the same model shown in Figure A.1.

[Figure A.2 about here]

Other Notation Details Specific to this Text

My preference is to avoid subscripts wherever possible. When the indexing is obvious or not necessary, I omit subscripts. For instance, I refer to an observed variable as y instead of y_i, omitting the subscript because it can be assumed that a variable varies across individual cases in the data set unless otherwise indicated. Admittedly, there is imprecision in doing this and potential confusion, but I believe the benefits of simplicity outweigh the costs.

The following subscripts are used whenever indexing is needed: i for individual case, j for the jth observed variable, k for the kth factor, and t for the time point. For multiple groups, g is used to designate a group-specific value. Although this does not ever occur in the text, a full example would be to refer to an observed variable as y_{ijktg}, for an observed score y for case i, on observed variable j, loading on factor k, at time point t, in group g. Where any of these are understood or not necessary, they will be omitted. For example, where there is only one observed variable or one factor and the context is clear, I will omit the j subscript for a particular observed variable, and/or the k subscript referring to a particular factor. If the data are cross-sectional, I will omit the t subscript. Likewise, for most formulas, I will omit subscripts from Greek LISREL matrix symbols when the all-y format is used or when the references is easily understood from the context (e.g., Λ_y will be simply Λ, and λ_y will be simply λ). To refer to a parameter being held constant across time points, such as a survey question that has been repeated over time, I will use a subscript in parentheses, as in $_{(j)}$ or $_{(1)}$.

For any references to regression analysis or logistic regression analysis, I use β instead of *b* to refer to unstandardized coefficients. I do this in part to reduce the number of different symbols used over all but also to emphasize the equivalence of regression coefficients and path coefficients from a structural equation model. Primes $'$ are used to distinguish between estimates obtained with referent and factor identification (mainly in Chapter 1), and should not be confused with the transpose operation that is used in matrix algebra (appearing in a few places in the text as well, but only next to a matrix). To denote an indicator that is not equal to *j*, the notation j^{o} used occasionally.

References

Jöreskog, K. G. (1973). Analysis of covariance structures. *Multivariate analysis, 3*, 263-85.

Keesling, J.W. (1972). Maximum likelihood approaches to causal analysis. Ph.D. thesis, University of Chicago

Wiley, D.E. (1973). The identification problem for structural equation models with unmeasured variables. In: A.S. Goldberger & O.D. Duncan (Eds.), *Structural equation models in the social sciences*. New York: Seminar.

Table A.1

All-y LISREL notation

Individual Parameter	English Spelling	Parameter Matrix	Description
λ	lambda	$\mathbf{\Lambda}$	factor loadings
ψ	psi	$\mathbf{\Psi}$	variances and covariances of latent variables
β	beta	$\mathbf{B}$	causal paths
θ	theta	$\mathbf{\Theta}$	measurement residual variances
ε	epsilon	$\boldsymbol{\varepsilon}$	measurement residuals, variances are elements in the theta matrix, $Var(\varepsilon) = \theta$
η	eta	$\boldsymbol{\eta}$	latent variables
ζ	zeta	$\boldsymbol{\zeta}$	structural disturbances
α	alpha	$\boldsymbol{\alpha}$	latent variable means
ν	nu	$\boldsymbol{\nu}$	measurement intercepts

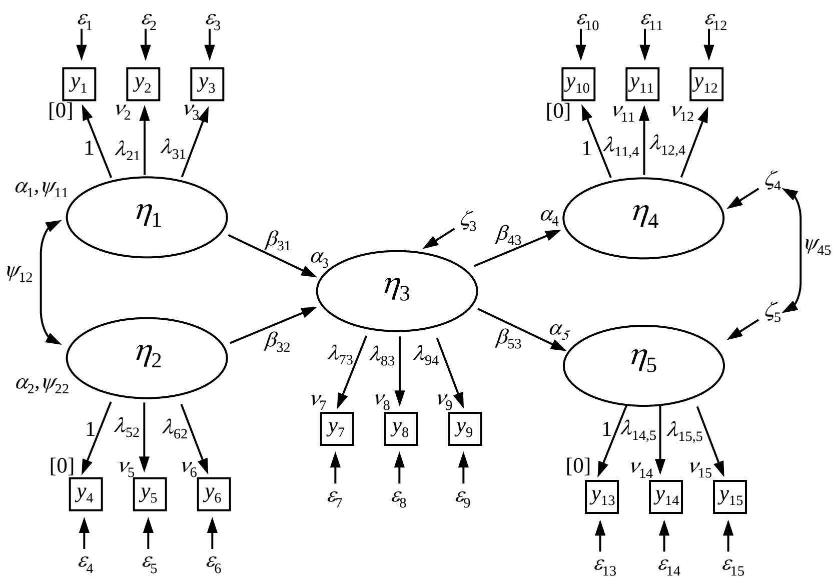
Table A.2
Full LISREL Notation

Exogenous Parameter	English Spelling	Exogenous Matrix	Description	Endogenous Parameter	English Spelling	Endogenous Matrix	Description
λ_x	lambda-x	$\mathbf{\Lambda}_x$	factor loadings for loadings on exogenous latent variables	λ_y	lambda-x	$\mathbf{\Lambda}_y$	factor loadings, λ_x for loadings on exogenous latent variables, *y* for indicators on endogenous latent variables
ϕ	phi	$\mathbf{\Phi}$	variances and covariances of exogenous latent variables, $Var(\xi)$ and $Cov(\xi,\xi)$	ψ	phi	$\mathbf{\Psi}$	disturbance variances and covariances among disturbances, $Var(\zeta)$ and $Cov(\zeta,\zeta)$
γ	gamma	$\mathbf{\Gamma}$	causal paths, endogenous predicted by exogenous	β	beta	$\mathbf{B}$	causal paths, endogenous predicted by endogenous
θ_δ	theta-delta	$\mathbf{\Theta}_\delta$	measurement residual variances for *x* variables	θ_ε	theta-epsilon	$\mathbf{\Theta}_\delta$	measurement residual variances for *y* variables
δ	delta	$\boldsymbol{\delta}$	measurement residual, variances are elements of theta-delta matrix, $Var(\delta) = \theta_\delta$	ε	epsilon	$\boldsymbol{\varepsilon}$	measurement residuals, variances are elements in the theta matrix, $Var(\varepsilon) = \theta$
ξ	ksi	$\boldsymbol{\xi}$	exogenous latent variables	η	eta	$\boldsymbol{\eta}$	endogenous latent variables
				ζ	zeta	$\boldsymbol{\zeta}$	structural disturbances
κ	kappa	$\boldsymbol{\kappa}$	exogenous latent variable mean	α	alpha	$\boldsymbol{\alpha}$	endogenous latent variable mean
ν_x	nu-x	$\boldsymbol{\nu}_x$	measurement intercepts for *x* variables	ν_y	nu-y	$\boldsymbol{\nu}_y$	measurement intercepts for *y* variables

Figure A.1. All-y LISREL Notation

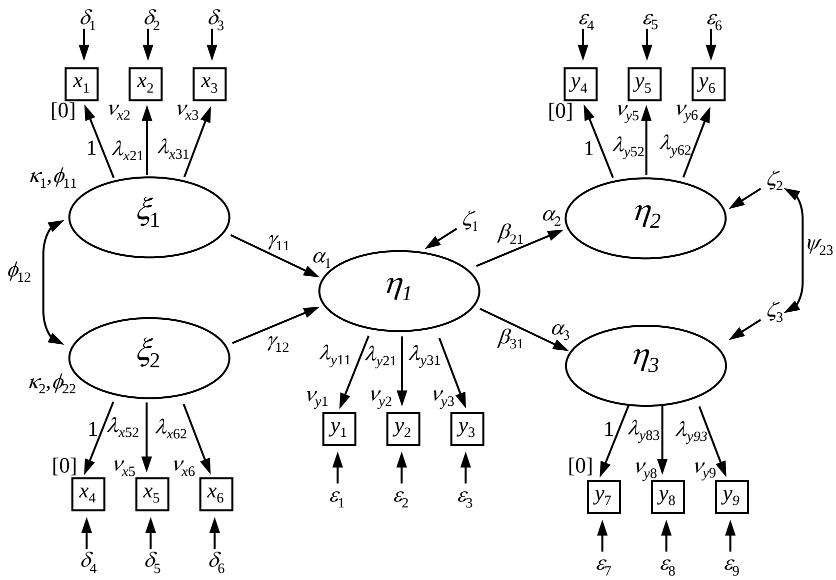
Figure A.2. Full LISREL notation.

Fig A.1



ε1 ε2 ε3
y1 y2 y3
[0] ν2 ν3
1 λ21 λ31
α1, ψ11
η1
ψ12
η2
α2, ψ22
1 λ52 λ62
[0] ν5 ν6
y4 y5 y6
ε4 ε5 ε6
β31
β32
ζ3
α3
η3
λ73 λ83 λ94
ν7 ν8 ν9
y7 y8 y9
ε7 ε8 ε9
β43
β53
ε10 ε11 ε12
y10 y11 y12
[0] ν11 ν12
1 λ11,4 λ12,4
α4
η4
ζ4
ψ45
ζ5
α5
η5
1 λ14,5 λ15,5
[0] ν14 ν15
y13 y14 y15
ε13 ε14 ε15

Fig A.1


δ1 δ2 δ3
x1 x2 x3
[0] νx2 νx3
1 λx21 λx31
κ1,φ11
ξ1
φ12
ξ2
κ2,φ22
1 λx52 λx62
[0] νx5 νx6
x4 x5 x6
δ4 δ5 δ6
γ11
γ12
α1
ζ1
η1
λy11 λy21 λy31
νy1 νy2 νy3
y1 y2 y3
ε1 ε2 ε3
β21
α2
β31
α3
ε4 ε5 ε6
y4 y5 y6
[0] νy5 νy6
1 λy52 λy62
η2
ζ2
ψ23
ζ3
η3
1 λy83 λy93
[0] νy8 νy9
y7 y8 y9
ε7 ε8 ε9