New Features in LISREL

Several special features and improvements are available in LISREL. Observed and latent variable names of up to sixteen characters are permitted and path diagram files can be exported as enhanced metafiles which can be imported into other documents. The iterative estimation algorithm for the parameters of single group LISREL models, which uses adaptive quadrature, has been improved. The multilevel generalized linear modeling application includes more link functions and computes estimates for the intra-class correlation coefficients.

LISREL also includes several new statistical methods which are not available in previous versions of LISREL. Two-stage multiple imputation Structural Equation Modeling (SEM) for continuous, ordinal, and a mixture of continuous and ordinal variables, confidence interval estimates for the parameters of LISREL models, and standard error estimates and confidence interval estimates for standardized and completely standardized solutions are implemented in LISREL. In addition, an iterative estimation algorithm for the parameters of single group LISREL models with or without constraints on the variances of the endogenous latent variables is available in LISREL.

The technical details along with illustrative examples for two-stage multiple imputation SEM are provided in section 2. Section 3 contains the statistical theory for standard error and confidence interval estimates for the parameters of LISREL models and includes an illustrative example. In section 4, the estimation theory for estimating the parameters of LISREL models with variance constraints for the endogenous latent variables are provided and demonstrated.

1. Two stage multiple imputation SEM

1.1 Continuous variables

Moment matrices

Suppose that the rows of $\mathbf{X}(n \times p)$ are n observations of p continuous variables $x_1, x_2, \ldots, x_p$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The sample covariance matrix, $\mathbf{S}$, is an unbiased estimator of $\boldsymbol{\Sigma}$ and may be expressed as

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'$$

where $\mathbf{x}_i$ and $\bar{\mathbf{x}}$ denote observation i and the sample mean vector of $\mathbf{x} = \left[x_1\ x_2 \ldots x_p\right]'$, respectively. A typical element of a consistent estimator, $\mathbf{U}$, of the asymptotic covariance matrix, $\boldsymbol{\Upsilon}$, of the sample variances and covariances (Browne 1984) is given by

$$u_{ij,kl} = w_{ijkl} - w_{ij}w_{kl}$$

where

$$w_{ijkl} = n^{-1}\sum_{m=1}^{n}(x_{im} - \bar{x}_i)(x_{jm} - \bar{x}_j)(x_{km} - \bar{x}_k)(x_{lm} - \bar{x}_l)$$

and

$$w_{ij} = n^{-1}\sum_{m=1}^{n}(x_{im} - \bar{x}_i)(x_{jm} - \bar{x}_j)$$

where

$$\bar{x}_i = n^{-1}\sum_{m=1}^{n}x_{im}$$

The robust ML, DWLS, WLS, and ULS methods can be used to fit structural equation models for continuous variables to the sample covariance matrix by using the estimated asymptotic covariance matrix of the sample variances and covariances.

The correlation matrix, $\mathbf{P}$, of $x_1, x_2, \ldots, x_p$ is the covariance matrix of the standardized variables $z_1, z_2, \ldots, z_p$ where

$$\mathbf{P} = \mathbf{D}_{\boldsymbol{\sigma}}^{-1}\boldsymbol{\Sigma}\mathbf{D}_{\boldsymbol{\sigma}}^{-1}$$

and

$$z_i = \frac{x_i - \mu_i}{\sigma_i}$$

where $\mathbf{D}_{\boldsymbol{\sigma}}$ denotes a diagonal matrix with the standard deviations $\sigma_1, \sigma_2, \ldots, \sigma_p$ of $x_1, x_2, \ldots, x_p$ on the diagonal. The sample correlation matrix, $\mathbf{R}$, is an unbiased estimator of $\mathbf{P}$ and may be expressed as

$$\mathbf{R} = \mathbf{D}_{\mathbf{s}}^{-1}\mathbf{R}\mathbf{D}_{\mathbf{s}}^{-1}$$

where $\mathbf{D}_s$ denotes a diagonal matrix with the sample standard deviations $s_1, s_2, \ldots, s_p$ of $x_1, x_2, \ldots, x_p$ on the diagonal. A typical element of a consistent estimator, $\mathbf{U}$, of the asymptotic covariance matrix, $\boldsymbol{\Upsilon}$, of the sample correlations (Steiger and Hakstian 1982) is given by

$$u_{ij,kl} = r_{ijkl} + \frac{1}{4} r_{ij} r_{kl} \left(r_{iikk} + r_{jjkk} + r_{iill} + r_{jjll} \right) - \frac{1}{2} r_{ij} \left(r_{iikl} + r_{jjkl} \right) - \frac{1}{2} r_{kl} \left(r_{ijkk} + r_{ijll} \right)$$

where

$$r_{ijkl} = (n-1)^{-1} \sum_{m=1}^{n} z_{im} z_{jm} z_{km} z_{lm}$$

and

$$r_{ij} = (n-1)^{-1} \sum_{m=1}^{n} z_{im} z_{jm}$$

and

$$z_{im} = \frac{x_{im} - \bar{x}_i}{s_i}$$

The robust DWLS, WLS, and ULS methods can be used to fit structural equation models for continuous variables to the sample correlation matrix by using the estimated asymptotic covariance matrix of the sample correlations.

Multiple imputation

The MCMC method

Suppose now that the n observations of the p continuous variables include missing data values with k missing data value patterns and that the joint distribution of the variables is a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The EM algorithm and the MCMC method for multiple imputation of incomplete data can be used to impute the missing data values of the continuous variables.

Suppose that $\mathbf{X}_o$ denote the observed data values. The EM algorithm (Dempster, Laird, and Rubin 1977) can be used to compute the maximum likelihood estimate of $\boldsymbol{\Sigma}$. The minus two observed-data log likelihood may be expressed as

$$-2 \ln L(\boldsymbol{\Sigma} \mid \mathbf{X}_o) = \sum_{i=1}^{k} n_i \ln |\boldsymbol{\Sigma}_i| + \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(\mathbf{x}_{oij} - \boldsymbol{\mu}_i \right)' \boldsymbol{\Sigma}_i^{-1} \left(\mathbf{x}_{oij} - \boldsymbol{\mu}_i \right)$$

where n_i denotes the number of observations of missing data value pattern $i = 1, 2, \ldots, k$, $\boldsymbol{\Sigma}_i$ denotes the population covariance matrix of missing data value pattern i, $\boldsymbol{\mu}_i$ denotes the mean vector of missing data value pattern i, and $\mathbf{x}_{oij}$ is the j^{th} vector of observed values of missing data value pattern i.

The initial estimate for the M-step is the sample covariance matrix, $\mathbf{S}$, of the complete data or $\mathbf{I}_p$ if the number of complete observations is too small. In the E-step, the conditional covariance matrices of the missing variables given the observed variables of the missing data value patterns are computed and used to compute an updated estimate $\hat{\boldsymbol{\Sigma}}^{(t+1)}$ of $\boldsymbol{\Sigma}$. Iteration of the consecutive M and E steps is terminated when the absolute difference between $\hat{\mathbf{\Sigma}}^{(t+1)}$ and $\hat{\mathbf{\Sigma}}^{(t)}$ is below the tolerance limit $\varepsilon = 10^{-5}$.

The EM estimate, $\hat{\mathbf{\Sigma}}$, of $\mathbf{\Sigma}$ is used as the initial covariance matrix of the multivariate normal distribution in the first step of the Monte Carlo Markov Chain (MCMC) method. In the first step (P-step) of the MCMC method, an estimate of $\mathbf{\Sigma}$ is simulated from an inverse Wishart distribution. In the I-step, observations are simulated from the conditional normal distributions of the missing variables given the observed k missing data value patterns and used to replace the missing data values. The next estimate of $\mathbf{\Sigma}$ is then obtained by computing the sample covariance matrix of the completed data. The P and I steps are repeated for a fixed number of times.

The FCS regression method

Suppose now that the n observations of the p continuous variables include missing data values and that a joint (multivariate) distribution of the variables exists. In this case, the Fully Conditional Specified (FCS) regression method (Brand 1999; Van Buuren 2007) can be used to impute the missing data values. The FCS regression method performs a fixed number of imputations to impute the missing data values. Each imputation consists of a filled-in phase and an imputation phase. In the filled-in phase, the missing data values are filled-in by using a sequence of regression analyses for the p continuous variables. These filled-in data are then used as the initial data for the imputation phase in which the missing data values are imputed by using a sequence of regression analyses for the p continuous variables. These imputed data are then used as the initial data for the next iteration of the imputation phase and a fixed number of iterations are executed for each imputation.

The filled-in stage fits the following p regression models sequentially to the data, namely

$$\begin{aligned}
&x_1 = \beta_{10} + e_1 \\
&x_2 = \beta_{20} + \beta_{21}x_1 + e_2 \\
&x_3 = \beta_{30} + \beta_{31}x_1 + \beta_{32}x_2 + e_3 \\
&\vdots \\
&x_p = \beta_{p0} + \beta_{p1}x_1 + \beta_{p2}x_2 + \cdots + \beta_{p,p-1}x_{p-1} + e_p
\end{aligned}$$

where the elements of $\boldsymbol{\beta} = \left[\beta_{10}\ \beta_{20} \cdots \beta_{p,p-1}\right]'$ denote unknown regression weights and $e_1, e_2, \ldots, e_p$ are p error variables. The first model is fitted to the complete data for x_1. The corresponding estimates are then used to simulate new parameter values from the posterior distributions of the parameters which in turn is used to fill-in the missing data values for x_1. The second model is then fitted to the complete data for x_2 and the filled-in data for x_1. The final model is fitted to the complete data for x_p and the filled-in data for $x_1, x_2, \ldots, x_{p-1}$. The filled-in data for $x_1, x_2, \ldots, x_p$ are used for the first iteration of the imputation phase. The simulation of the new parameter values from the posterior distributions of the parameters and the imputation of the missing data values for each of the p regression models use the same steps as outlined next for each iteration of the imputation stage.

For each iteration of the imputation stage, the following regression models are fitted sequentially either to the filled-in data or the imputed data, namely

$$x_j = \beta_0 + \beta_1 x_1 + \cdots + \beta_{j-1}x_{j-1} + \beta_{j+1}x_{j+1} + \cdots + \beta_p x_p + e_j$$

where $j = 1, 2, \ldots, p$, the elements of $\boldsymbol{\beta}_j = \left[\beta_0\ \beta_1 \ldots \beta_{j-1}\beta_{j+1} \ldots \beta_p\right]'$ denote p unknown regression weights, and e_j denotes an error variable with variance σ_j^2. The estimated covariance matrix of the estimator $\hat{\boldsymbol{\beta}}_j$ of $\boldsymbol{\beta}_j$ may be expressed as

$$\sigma_j^2 \mathbf{V}_j = \sigma_j^2 \left(\mathbf{X}'_{(j)}\mathbf{X}_{(j)}\right)^{-1}$$

where $\mathbf{X}_{(j)}$ denotes rows $1, 2, \ldots, j-1, j, \ldots, p$ of the filled-in or imputed data. New values for the parameters are then simulated from their posterior distributions as

$$\boldsymbol{\beta}_{jt} = \hat{\boldsymbol{\beta}}_j + \sigma_{tj}^2 \mathbf{V}'_{hj}\mathbf{z}$$

$$\sigma_{tj}^2 = \frac{\hat{\sigma}_j^2\left(n_j - p\right)}{c}$$

where $\mathbf{V}_{hj}$ denotes the upper triangular matrix in the Cholesky decomposition of $\mathbf{V}_j = \mathbf{V}'_{hj}\mathbf{V}_{hj}$, $\mathbf{z}$ denotes a $p \times 1$ standard normal vector, and c is a Chi-square variable with $n_j - p$ degrees of freedom. The missing data values are then imputed as

$$x_{ijm} = \boldsymbol{\beta}'_{jt}\mathbf{x}_{i(j)} + \sigma_{tj} z$$

where x_{ijm} denotes a missing data value in row i and column j of $\mathbf{X}$, $\mathbf{x}_{i(j)}$ denotes row i of $\mathbf{X}_{(j)}$, and z is a standard normal variable.

Average unstandardized moment matrices

Suppose that $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_m$ are m imputed data sets for the incomplete data matrix, $\mathbf{X}$, of the p continuous variables $x_1, x_2, \ldots, x_p$ and that $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_m$ and $\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_m$ denote the corresponding sample covariance matrices and the estimated asymptotic covariance matrices of the variances and covariances, respectively. Then, the average sample covariance matrix is

$$\bar{\mathbf{S}} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{S}_i$$

and the average estimated asymptotic covariance matrix is

$$\bar{\mathbf{U}} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{U}_i$$

Chung and Cai (2019) point out that $\bar{\mathbf{U}}$ only captures uncertainty based on complete data. As a result, its inverse cannot be used as a weight matrix for the robust ML, DWLS, WLS, and ULS methods for continuous structural equational modeling. A corrected weight matrix is obtained by correcting for the between-imputation variation in the estimated variances and covariances and is obtained as the inverse of

$$\hat{\boldsymbol{\Upsilon}} = \bar{\mathbf{U}} + \frac{m+1}{m(m-1)}\left[\sum_{i=1}^{m}\left(\mathbf{s}_i - \bar{\mathbf{s}}\right)\left(\mathbf{s}_i - \bar{\mathbf{s}}\right)'\right]$$

where $\mathbf{s}$ denotes the $p\times\left(p+1\right)/2$ vector consisting of the nonduplicated elements of the $p\times p$ symmetric matrix $\mathbf{S}$. $\bar{\mathbf{S}}$ and $\hat{\boldsymbol{\Upsilon}}$ can be used to fit structural equation models to the average sample covariance matrix with the robust ML, DWLS, WLS, and ULS methods. The corrected robust DWLS and ULS Chi-square test statistic proposed by Chung and Cai (2019) is given by

$$T_B=(n-1)(\mathbf{s}-\boldsymbol{\sigma}(\hat{\boldsymbol{\theta}}))'\mathbf{V}(\mathbf{s}-\boldsymbol{\sigma}(\hat{\boldsymbol{\theta}}))$$

where

$$\mathbf{V}=\hat{\boldsymbol{\Upsilon}}^{-1}-\hat{\boldsymbol{\Upsilon}}^{-1}\hat{\boldsymbol{\Delta}}(\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Delta}})^{-1}\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Upsilon}}^{-1}$$

where $\hat{\boldsymbol{\Delta}}$ denotes the Jacobian matrix of $\boldsymbol{\sigma}(\boldsymbol{\theta})$ with respect to the unknown parameters $\boldsymbol{\theta}$ of the structural equation model evaluated at $\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}$. The small sample adjusted T_B test statistic (Yuan and Bentler 1997) is given by

$$T_{YB}=\frac{T_B}{1+nT_B/(n-1)}.$$

Average standardized moment matrices

Suppose that $\mathbf{X}_1,\mathbf{X}_2,\ldots,\mathbf{X}_m$ are m imputed data sets for the incomplete data matrix, $\mathbf{X}$, of the p continuous variables $x_1, x_2, \ldots, x_p$ and that $\mathbf{R}_1,\mathbf{R}_2,\ldots,\mathbf{R}_m$ and $\mathbf{U}_1,\mathbf{U}_2,\ldots,\mathbf{U}_m$ denote the corresponding sample correlation matrices and the estimated asymptotic covariance matrices of the sample correlations, respectively. Then, the average sample correlation matrix is

$$\bar{\mathbf{R}}=\frac{1}{m}\sum_{i=1}^{m}\mathbf{R}_i$$

and the average estimated asymptotic covariance matrix is

$$\bar{\mathbf{U}}=\frac{1}{m}\sum_{i=1}^{m}\mathbf{U}_i$$

Chung and Cai (2019) point out that $\bar{\mathbf{U}}$ only captures uncertainty based on complete data. As a result, its inverse cannot be used as a weight matrix for the robust DWLS, WLS, and ULS methods for continuous structural equational modeling for correlation matrices. A corrected weight matrix is obtained by correcting for the between-imputation variation in the estimated correlations and is obtained as the inverse of

$$\hat{\boldsymbol{\Upsilon}}=\bar{\mathbf{U}}+\frac{m+1}{m(m-1)}\left[\sum_{i=1}^{m}\left(\mathbf{r}_i-\bar{\mathbf{r}}\right)\left(\mathbf{r}_i-\bar{\mathbf{r}}\right)'\right]$$

where $\mathbf{r}$ denotes the $p\times\left(p-1\right)/2$ vector consisting of the nondiagonal and the nonduplicated elements of the $p\times p$ symmetric matrix $\mathbf{R}$. $\bar{\mathbf{R}}$ and $\hat{\boldsymbol{\Upsilon}}$ can be used to fit structural equation models to the average sample correlation matrix with the robust DWLS, WLS, and ULS methods. The corrected robust DWLS and ULS Chi-square test statistic proposed by Chung and Cai (2019) is given by

$$T_B=(n-1)(\mathbf{r}-\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}))'\mathbf{V}(\mathbf{r}-\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}))$$

where

$$\mathbf{V} = \hat{\boldsymbol{\Upsilon}}^{-1} - \hat{\boldsymbol{\Upsilon}}^{-1}\hat{\boldsymbol{\Delta}}(\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Delta}})^{-1}\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Upsilon}}^{-1}$$

where $\hat{\boldsymbol{\Delta}}$ denotes the Jacobian matrix of $\boldsymbol{\rho}(\boldsymbol{\theta})$ with respect to the unknown parameters $\boldsymbol{\theta}$ of the structural equation model evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. The small sample adjusted T_B test statistic (Yuan and Bentler 1997) is given by

$$T_{YB} = \frac{T_B}{1 + nT_B / (n-1)}$$

1.2 Ordinal variables

Polychoric Correlations

Suppose that the rows of $\mathbf{X}(n \times p)$ are n observations of p ordinal variables $x_1, x_2, \ldots, x_p$ with $m_1, m_2, \ldots, m_p$ categories, respectively. Suppose further that these p ordinal variables are the result of the discretization of the underlying p continuous standard normal variables $z_1, z_2, \ldots, z_p$ as such that $\mathbf{z} = \begin{bmatrix} z_1 & z_2 & \ldots & z_p \end{bmatrix}' \sim N(\mathbf{0}, \mathbf{P})$ and

$$\begin{cases} x_i = 1 & \text{if } \tau_{i0} < z_i \le \tau_{i1} \\ x_i = 2 & \text{if } \tau_{i1} < z_i \le \tau_{i2} \\ & \vdots \\ x_i = m_i & \text{if } \tau_{i,m_i-1} < z_i \le \tau_{i,m_i} \end{cases}$$

where $\mathbf{P}$ denotes the population correlation matrix of $\mathbf{z}$ and $-\infty = \tau_{i0} < \tau_{i1} < \tau_{i2} \ldots \le \tau_{i,m_i} = \infty$ are parameters known as thresholds. The model for the univariate marginal of variable x_i is

$$\pi_{ik} = \int_{\tau_{i,k-1}}^{\tau_{i,}} \phi(u)du$$

where $\phi(.)$ denotes the probability density function of the standard normal distribution. The maximum likelihood estimator of τ_{ik} (Jöreskog, 1994) is given by

$$\hat{\tau}_{ik} = \boldsymbol{\Phi}^{-1}(p_{i1} + p_{i2} + \ldots + p_{ik})$$

where $\boldsymbol{\Phi}^{-1}(.)$ denotes the inverse of the cumulative distribution function of the standard normal distribution and $p_{i1}, p_{i2}, \ldots, p_{im_i}$ denote the marginal sample proportions for x_i.

The polychoric correlation matrix, **R**, is a consistent estimator of the population correlation matrix **P**. The model for the bivariate marginal of variables x_i and x_j is

$$\pi_{ijkl} = \int_{\tau_{i,k-1}}^{\tau_{i,}} \int_{\tau_{i,l-1}}^{\tau_{il}} \phi_2(u, v, \rho_{ij})dudv$$

where $\phi_2(u, v, \rho_{ij})$ denotes the probability density function of the bivariate standard normal distribution with correlation ρ_{ij}. The maximization of the bivariate likelihood function is equivalent to minimization of the discrepancy function

$$F\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\boldsymbol{\tau}}_j\right) = \sum_{k=1}^{m_i}\sum_{l=1}^{m_j} p_{ijkl}\left(\ln\left\{p_{ijkl}\right\} - \ln\left\{\pi_{ijkl}\right\}\right)$$

where $\hat{\boldsymbol{\tau}}_i$ and $\hat{\boldsymbol{\tau}}_j$ denote the maximum likelihood estimators of the $m_i - 1$ and $m_j - 1$ thresholds of variables x_i and x_j, respectively and p_{ijkl} is the sample proportion for $x_i = k$ and $x_j = l$. The gradient of $F(\cdot)$ (Olsson 1979) may be expressed as

$$g\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\boldsymbol{\tau}}_j\right) = -\sum_{k=1}^{m_i}\sum_{l=1}^{m_j} \frac{p_{ijkl}}{\pi_{ijkl}}\left[\frac{\partial \pi_{ijkl}}{\partial \rho_{ij}}\right]$$

where (Olsson 1979)

$$\frac{\partial \pi_{ijkl}}{\partial \rho_{ij}} = \phi_2\left(\tau_{ik}, \tau_{il}\right) - \phi_2\left(\tau_{i,k-1}, \tau_{jl}\right) - \phi_2\left(\tau_{i,k}, \tau_{j,l-1}\right) + \phi_2\left(\tau_{i,k-1}, \tau_{j,l-1}\right)$$

where $\phi_2(\cdot)$ denotes the density function of the bivariate standard normal distribution with correlation ρ_{ij}. The information (Jöreskog, 1994) is given by

$$i\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\boldsymbol{\tau}}_j\right) = \left[\frac{\partial \pi_{ijkl}}{\partial \rho_{ij}}\right]\frac{1}{\pi_{ijkl}}\left[\frac{\partial \pi_{ijkl}}{\partial \rho_{ij}}\right]$$

The Fisher scoring algorithm is used to minimize $F(\cdot)$ with respect to ρ_{ij}. Let $\theta = \rho_{ij}$. If $\hat{\theta}^{(t)}$ denotes the t^{th} successive approximation to $\hat{\theta}$, then the $(t+1)^{st}$ approximation is obtained from

$$\hat{\theta}^{(t+1)} = \hat{\theta}^{(t)} - \frac{g\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\boldsymbol{\tau}}_j\right)}{i\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\boldsymbol{\tau}}_j\right)}$$

Iteration is terminated when the absolute gradient value is below the tolerance limit $\varepsilon = 10^{-3}$.

The asymptotic covariance matrix, $\boldsymbol{\Upsilon}$, of the $p^* = p(p-1)/2$ polychoric correlations is a $p^*(p^*+1)/2$ matrix. A typical element of $\hat{\boldsymbol{\Upsilon}}$ (Jöreskog, 1994) may be expressed as

$$[\hat{\boldsymbol{\Upsilon}}]_{ij,rs} = \sum_{c=1}^{n} \kappa_{cijrs}\hat{\gamma}_{ijkl}\hat{\gamma}_{rsno} - \hat{\omega}_{ij}\hat{\omega}_{kl}$$

where $\kappa_{cijrs} = \frac{1}{n}$ if $x_{ci} = k$, $x_{cj} = l$, $x_{cr} = n$, and $x_{cs} = o$ and 0 otherwise, $\omega_{ij} = \sum_{k=1}^{m_i}\sum_{l=1}^{m_j}\gamma_{ijkl}\pi_{ijkl}$, $\omega_{rs} = \sum_{n=1}^{m_i}\sum_{o=1}^{m_j}\gamma_{rsno}\pi_{rsno}$, and γ_{ijkl} denotes a typical element of

$$\boldsymbol{\Gamma}_{ij} = \boldsymbol{\alpha}_{ij} + \mathbf{B}_i\boldsymbol{\beta}_i\mathbf{1}'_j + \mathbf{1}_i\boldsymbol{\beta}'_j\mathbf{B}'_j$$

where $\mathbf{1}_i$ denotes an $m_i \times 1$ column vector and

$$\mathbf{B}_i' = \left(\mathbf{A}_i'\mathbf{D}_{\pi_i}^{-1}\mathbf{A}_i\right)^{-1}\mathbf{A}_i'\mathbf{D}_{\pi_i}^{-1}$$

$$\mathbf{B}_j' = \left(\mathbf{A}_j'\mathbf{D}_{\pi_i}^{-1}\mathbf{A}_j\right)^{-1}\mathbf{A}_j'\mathbf{D}_{\pi_j}^{-1}$$

where $\mathbf{A}_i$ denotes the $m_i \times (m_i - 1)$ matrix given by

$$\mathbf{A}_i = \begin{bmatrix} \phi(\tau_{1k}) & 0 & \dots & 0 \\ -\phi(\tau_{1k}) & \phi(\tau_{2k}) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & -\phi\left(\tau_{m_i-1.k}\right) \end{bmatrix}$$

Typical elements of $\boldsymbol{\alpha}_{ij}$, $\boldsymbol{\beta}_i$, and $\boldsymbol{\beta}_j$ are given by

$$\alpha_{ijkl} = D^{-1}\frac{1}{\pi_{ijkl}}\frac{\partial \pi_{ijkl}}{\partial \rho_{ij}}$$

$$\left[\boldsymbol{\beta}_i\right]_k = \sum_{k=1}^{m_i}\sum_{l=1}^{m_j}\alpha_{ijkl}\left(\frac{\partial \pi_{ijkl}}{\partial \tau_{ik}}\right)$$

$$\left[\boldsymbol{\beta}_j\right]_l = \sum_{k=1}^{m_i}\sum_{l=1}^{m_j}\alpha_{ijkl}\left(\frac{\partial \pi_{ijkl}}{\partial \tau_{il}}\right)$$

where

$$D = \sum_{k=1}^{m_i}\sum_{l=1}^{m_j}\frac{1}{\pi_{ijkl}}\left(\frac{\partial \pi_{ijkl}}{\partial \rho_{ij}}\right)^2.$$

The robust DWLS, WLS, or ULS methods can be used to fit structural equation models for ordinal variables to the polychoric correlation matrix by using the estimated asymptotic covariance matrix of the polychoric correlations (Chung and Cai (2019)).

Multiple Imputation

The MCMC method

Suppose now that the n observations of the p ordinal variables include missing data values with k missing data value patterns. The EM algorithm and the MCMC method for multiple imputation of incomplete data are intended for continuous variables and cannot readily be applied to ordinal variables. However, they can be applied to the underlying continuous variables $z_1, z_2, \dots, z_p$ associated with the ordinal variables $x_1, x_2, \dots, x_p$. Although no observations for these continuous variables are available, these variables are assumed to have a multivariate standard normal distribution with a population covariance matrix $\boldsymbol{\Sigma}$. As a result, we can simulate data from this distribution by using the polychoric correlation matrix of the complete data of the variables if the number of complete cases is large enough and use either the EM algorithm or the MCMC algorithm to impute the missing data values for the underlying continuous variables. After imputation, the estimated thresholds can be used to replace the missing data values for the corresponding ordinal variables by using the relationship between the ordinal variables, the underlying continuous variables, and the thresholds.

Suppose that the rows of $\mathbf{Z}(n \times p)$ are n observations of the p underlying continuous variables $z_1, z_2, \ldots, z_p$ simulated from the $N(\mathbf{0}, \mathbf{\Sigma})$ distribution and that $\mathbf{Z}_o$ denotes the observed data values that corresponds with the observed data values of **X**. The EM algorithm (Dempster, Laird, and Rubin 1977) can be used to compute the maximum likelihood estimate of $\mathbf{\Sigma}$. The minus two observed-data log likelihood may be expressed as

$$-2\ln L(\mathbf{\Sigma} \mid \mathbf{Z}_o) = \sum_{i=1}^{k} n_i \ln\left|\mathbf{\Sigma}_i\right| + \sum_{i=1}^{k}\sum_{j=1}^{n_i} \mathbf{z}'_{oij}\mathbf{\Sigma}_i^{-1}\mathbf{z}_{oij}$$

where n_i denotes the number of observations of missing data value pattern $i = 1,2,\cdots,k$, $\mathbf{\Sigma}_i$ denotes the population covariance matrix for missing data value pattern i, and $\mathbf{z}_{oij}$ is the j^{th} vector of observed values of missing data value pattern i.

The initial estimate for the M-step is the sample covariance matrix, $\mathbf{S}$, of the complete data or $\mathbf{I}_p$ if the number of complete observations is too small. In the E-step, the conditional covariance matrices of the missing variables given the observed variables for the missing data value patterns are computed and used to compute an updated estimate, $\hat{\mathbf{\Sigma}}^{(t+1)}$ of $\mathbf{\Sigma}$. Iteration of the consecutive M and E steps is terminated when the absolute difference between $\hat{\mathbf{\Sigma}}^{(t+1)}$ and $\hat{\mathbf{\Sigma}}^{(t)}$ is below the tolerance limit $\varepsilon = 10^{-5}$.

The correlation matrix of the EM estimate, $\hat{\mathbf{\Sigma}}$, of $\mathbf{\Sigma}$ is used as the initial covariance matrix of the multivariate standard normal distribution in the first step of the Monte Carlo Markov Chain (MCMC) method. In the first step (P-step) of the MCMC method, an estimate of $\mathbf{\Sigma}$ is simulated form an inverse Wishart distribution. In the I-step, observations are simulated from the conditional standard normal distributions of the missing variables given the observed k missing data value patterns and used to replace the missing data values. The next estimate of $\mathbf{\Sigma}$ is then obtained by computing the sample correlation matrix of the completed data. The P and I steps are repeated for a fixed number of times.

Let the rows of $\mathbf{Z}_i : (n \times p)$ contain the observed and the imputed data values for the standard normal variables $z_1, z_2, \ldots, z_p$. The observed data for the ordinal variables are obtained from the corresponding observed data values of **X**. The missing data values of **X** are then replaced by the values obtained from the corresponding imputed data values of **Z** and the estimated thresholds by using the relationship between the ordinal variables, the underlying continuous variables, and the thresholds.

The FCS ordinal logistic regression method

Suppose now that the n observations of the p ordinal variables include missing data values and that a joint (multivariate) distribution of the variables exists. In this case, the Fully Conditional Specified (FCS) ordinal logistic regression method (Brand 1999; Van Buuren 2007) can be used to impute the missing data values. The FCS ordinal logistic regression method performs a fixed number of imputations to impute the missing data values. Each imputation consists of a filled-in phase and an imputation phase. In the filled-in phase, the missing data values are filled-in by using a sequence of ordinal logistic regression analyses for the p ordinal variables. These filled-in data are then used as the initial data for the imputation phase in which the missing data values are imputed by using a sequence of ordinal logistic regression analyses for the p ordinal variables. These imputed data are then used as the initial data for the next iteration of the imputation phase and a fixed number of iterations are executed for each imputation.

The filled-in stage fits the following p ordinal logistic regression models sequentially to the data, namely

$$\begin{aligned}
&\operatorname{logit}(\pi_{1k}) = \alpha_{1k} \\
&\operatorname{logit}(\pi_{2k}) = \alpha_{2k} + \beta_{21}x_1 \\
&\operatorname{logit}(\pi_{3k}) = \alpha_{3k} + \beta_{31}x_1 + \beta_{32}x_2 \\
&\vdots \\
&\operatorname{logit}(\pi_{pk}) = \alpha_{pk} + \beta_{p1}x_1 + \beta_{p2}x_2 + \cdots + \beta_{p,p-1}x_{p-1}
\end{aligned}$$

where $\pi_{ijk} = P(x_j \le k \mid x_1, x_2, \ldots, x_{j-1})$, $\operatorname{logit}(\pi_{jk}) = \ln(\pi_{jk}) - \ln(\pi_{im_j})$, and the elements of $\boldsymbol{\gamma} = \left[\alpha_{11}\ \alpha_{12} \ldots \beta_{21} \ldots \beta_{p,p-1}\right]'$ denote unknown regression weights. The first model is fitted to the complete data for x_1. The corresponding estimates are then used to simulate new parameter values from the posterior distribution of the parameters which in turn is used to fill-in the missing data values for x_1. The second model is then fitted to the complete data for x_2 and the filled-in data for x_1. The final model is fitted to the complete data for x_p and the filled-in data for $x_1, x_2, \ldots, x_{p-1}$. The filled-in data for $x_1, x_2, \ldots, x_p$ are used for the first iteration of the imputation phase. The simulation of the new parameter values from the posterior distribution of the parameters and the imputation of the missing data values for each of the p ordinal logistic regression models use the same steps as outlined next for each iteration of the imputation stage.

For each iteration of the imputation stage, the following ordinal logistic regression models are fitted sequentially either to the filled-in data or the imputed data, namely

$$\operatorname{logit}(\pi_{ijk}) = \alpha_k + \beta_1 x_1 + \cdots + \beta_{j-1}x_{j-1} + \beta_{j+1}x_{j+1} + \cdots + \beta_p x_p$$

where $\pi_{ijk} = P(x_j \le k \mid x_1, \ldots, x_{j-1}, x_{j+1} \ldots, x_p)$, $\operatorname{logit}(\pi_{jk}) = \ln(\pi_{jk}) - \ln(\pi_{im_j})$, the elements of $\boldsymbol{\gamma}_j = \left[\alpha_1\ \alpha_2 \ldots \alpha_{m_i-1}\ \beta_1 \ldots \beta_{j-1}\beta_{j+1} \ldots \beta_p\right]'$ denote $p + m_j - 1$ unknown regression weights, $j = 1, 2, \ldots, p$, and $k = 1, 2, \ldots, m_j - 1$. Let $\mathbf{V}_j$ denote the estimated covariance matrix of the estimator $\hat{\gamma}$ of $\boldsymbol{\gamma}_j$.

New values for the parameters are then simulated from their posterior distribution as

$$\boldsymbol{\gamma}_{jt} = \hat{\gamma}_j + \mathbf{V}'_{hj}\mathbf{z}$$

where $\mathbf{V}'_{hj}$ denotes the upper triangular matrix in the Cholesky decomposition of $\mathbf{V}_j = \mathbf{V}'_{hj}\mathbf{V}_{hj}$, and $\mathbf{z}$ is a $(p + m_j - 1) \times 1$ standard normal vector. These new parameter values are then used to compute the predicted cumulative probability $\hat{\pi}_{jk}$ for $k = 1, 2, \ldots, m_j - 1$. A random uniform variable, u, between 0 and 1 is simulated and the missing data values for x_j are imputed as 1 if $u < \hat{\pi}_{j1}$, as k if $\hat{\pi}_{j,k-1} \le u < \hat{\pi}_{jk}$, and as m_j if $u \ge \hat{\pi}_{jm_j}$.

Average moment matrices

Suppose that $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_m$ are m imputed data sets for the incomplete data matrix, $\mathbf{X}$, of the p ordinal variables $x_1, x_2, \ldots, x_p$ and that $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_m$ and $\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_m$ denote the corresponding polychoric correlation matrices and the estimated asymptotic covariance matrices of the polychoric correlations, respectively. Then, the average polychoric correlation matrix is

$$\bar{\mathbf{R}} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{R}_i$$

and the average estimated asymptotic covariance matrix is

$$\bar{\mathbf{U}} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{U}_i$$

Chung and Cai (2019) point out that $\bar{\mathbf{U}}$ only captures uncertainty based on complete data. As a result, its inverse cannot be used as a weight matrix for the robust DWLS, WLS, and ULS methods for ordinal structural equational modeling. A corrected weight matrix is obtained by correcting for the between-imputation variation in the estimated polychoric correlations and is obtained as the inverse of

$$\hat{\boldsymbol{\Upsilon}} = \bar{\mathbf{U}} + \frac{m+1}{m(m-1)}\left[\sum_{i=1}^{m}(\mathbf{r}_i - \bar{\mathbf{r}})(\mathbf{r}_i - \bar{\mathbf{r}})'\right]$$

where $\mathbf{r}$ denotes the $p \times (p-1)/2$ vector consisting of the nondiagonal and nonduplicated elements of the $p \times p$ symmetric matrix $\mathbf{R}$. $\bar{\mathbf{R}}$ and $\hat{\boldsymbol{\Upsilon}}$ can be used to fit structural equation models to the average polychoric correlation matrix with the robust DWLS, WLS, and ULS methods. The corrected robust DWLS and ULS Chi-square test statistic proposed by Chung and Cai (2019) is given by

$$T_B = (n-1)(\bar{\mathbf{r}} - \boldsymbol{\rho}(\hat{\boldsymbol{\theta}}))'\mathbf{V}(\bar{\mathbf{r}} - \boldsymbol{\rho}(\hat{\boldsymbol{\theta}}))$$

where

$$\mathbf{V} = \hat{\boldsymbol{\Upsilon}}^{-1} - \hat{\boldsymbol{\Upsilon}}^{-1}\hat{\boldsymbol{\Delta}}(\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Delta}})^{-1}\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Upsilon}}^{-1}$$

where $\hat{\boldsymbol{\Delta}}$ denotes the Jacobian matrix of $\boldsymbol{\rho}(\boldsymbol{\theta})$ with respect to the unknown parameters, $\boldsymbol{\theta}$, of the structural equation model evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. The small sample adjusted T_B test statistic (Yuan and Bentler 1997) is given by

$$T_{YB} = \frac{T_B}{1 + nT_B/(n-1)}$$

1.3 Mixed variables

Correlations

Polychoric correlations

Suppose that the rows of $\mathbf{X}(n \times p)$ are n observations of p ordinal variables $x_1, x_2, \ldots, x_p$ with $m_1, m_2, \ldots, m_p$ categories, respectively. Suppose further that these p ordinal variables are the result of the discretization of the underlying p continuous standard normal variables $z_1, z_2, \ldots, z_p$ as such that $\mathbf{z} = \begin{bmatrix} z_1 & z_2 & \ldots & z_p \end{bmatrix}' \sim N(\mathbf{0}, \mathbf{P})$ and

$$\begin{cases} x_i = 1 & \text{if } \tau_{i0} < z_i \le \tau_{i1} \\ x_i = 2 & \text{if } \tau_{i1} < z_i \le \tau_{i2} \\ & \vdots \\ x_i = m_i & \text{if } \tau_{i,m_i-1} < z_i \le \tau_{i,m_i} \end{cases}$$

where **P** denotes the population correlation matrix of **z** and $-\infty = \tau_{i0} < \tau_{i1} < \tau_{i2} \ldots \le \tau_{i,m_i} = \infty$ are parameters known as thresholds. The model for the univariate marginal of variable x_i is

$$\pi_{ik} = \int_{\tau_{i,k-1}}^{\tau_{i,}} \phi(u)du$$

where $\phi(.)$ denotes the probability density function of the standard normal distribution. The maximum likelihood estimator of τ_{ik} (Jöreskog, 1994) is given by

$$\hat{\tau}_{ik} = \mathbf{\Phi}^{-1}(p_{i1} + p_{i2} + \ldots + p_{ik})$$

where $\mathbf{\Phi}^{-1}(.)$ denotes the inverse of the cumulative distribution function of the standard normal distribution and $p_{i1}, p_{i2}, \ldots, p_{im_i}$ denote the marginal sample proportions for x_i.

The polychoric correlation matrix, **R**, is a consistent estimator of the population correlation matrix **P**. The model for the bivariate marginal of variables x_i and x_j is

$$\pi_{ijkl} = \int_{\tau_{i,k-1}}^{\tau_{i,}} \int_{\tau_{i,l-1}}^{\tau_{il}} \phi_2(u, v, \rho_{ij}) dudv$$

where $\phi_2(u, v, \rho_{ij})$ denotes the probability density function of the bivariate standard normal distribution with correlation ρ_{ij}. The maximization of the bivariate likelihood function is equivalent to minimization of the discrepancy function

$$F\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\boldsymbol{\tau}}_j\right) = \sum_{k=1}^{m_i} \sum_{l=1}^{m_j} p_{ijkl} \left(\ln\left\{p_{ijkl}\right\} - \ln\left\{\pi_{ijkl}\right\}\right)$$

where $\hat{\boldsymbol{\tau}}_i$ and $\hat{\boldsymbol{\tau}}_j$ denote the maximum likelihood estimators of the $m_i - 1$ and $m_j - 1$ thresholds of variables x_i and x_j, respectively and p_{ijkl} is the sample proportion for $x_i = k$ and $x_j = l$. The gradient of $F(\cdot)$ (Olsson 1979) may be expressed as

$$g\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\boldsymbol{\tau}}_j\right) = -\sum_{k=1}^{m_i} \sum_{l=1}^{m_j} \frac{p_{ijkl}}{\pi_{ijkl}} \left[\frac{\partial \pi_{ijkl}}{\partial \rho_{ij}}\right]$$

where (Olsson 1979)

$$\frac{\partial \pi_{ijkl}}{\partial \rho_{ij}} = \phi_2\left(\tau_{ik}, \tau_{il}\right) - \phi_2\left(\tau_{i,k-1}, \tau_{jl}\right) - \phi_2\left(\tau_{i,k}, \tau_{j,l-1}\right) + \phi_2\left(\tau_{i,k-1}, \tau_{j,l-1}\right)$$

where $\phi_2(\cdot)$ denotes the density function of the bivariate standard normal distribution with correlation ρ_{ij}. The information (Jöreskog, 1994) is given by

$$i\left(\rho_{ij},\hat{\boldsymbol{\tau}}_i,\hat{\boldsymbol{\tau}}_j\right)=\sum_{k=1}^{m_i}\sum_{l=1}^{m_j}\frac{1}{\pi_{ijkl}}\left[\frac{\partial\pi_{ijkl}}{\partial\rho_{ij}}\right]^2$$

The Fisher scoring algorithm is used to minimize $F(\cdot)$ with respect to ρ_{ij}. Let $\theta=\rho_{ij}$. If $\hat{\theta}^{(t)}$ denotes the t^{th} successive approximation to $\hat{\theta}$, then the $(t+1)^{st}$ approximation is obtained from

$$\hat{\theta}^{(t+1)}=\hat{\theta}^{(t)}-\frac{g\left(\rho_{ij},\hat{\boldsymbol{\tau}}_i,\hat{\boldsymbol{\tau}}_j\right)}{i\left(\rho_{ij},\hat{\boldsymbol{\tau}}_i,\hat{\boldsymbol{\tau}}_j\right)}.$$

Iteration is terminated when the absolute gradient value is below the tolerance limit $\varepsilon=10^{-3}$.

Pearson product-moment correlations

Suppose that the rows of $\mathbf{X}(n\times p)$ are n observations of p continuous variables $x_1,x_2,\ldots,x_p$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The sample covariance matrix, $\mathbf{S}$, is an unbiased estimator of $\boldsymbol{\Sigma}$ and may be expressed as

$$\mathbf{S}=\frac{1}{n-1}\sum_{i=1}^{n}\left(\mathbf{x}_i-\bar{\mathbf{x}}\right)\left(\mathbf{x}_i-\bar{\mathbf{x}}\right)'$$

where $\mathbf{x}_i$ and $\bar{\mathbf{x}}$ denote observation i and the sample mean vector of $\mathbf{x}=\left[x_1\ x_2\ \ldots\ x_p\right]'$, respectively.

The correlation matrix, $\mathbf{P}$, of $x_1,x_2,\ldots,x_p$ is the covariance matrix of the standardized variables $z_1,z_2,\ldots,z_p$ where

$$\mathbf{P}=\mathbf{D}_{\boldsymbol{\sigma}}^{-1}\boldsymbol{\Sigma}\mathbf{D}_{\boldsymbol{\sigma}}^{-1}$$

and

$$z_i=\frac{x_i-\mu_i}{\sigma_i}$$

where $\mathbf{D}_{\boldsymbol{\sigma}}$ denotes a diagonal matrix with the standard deviations $\sigma_1,\sigma_2,\ldots,\sigma_p$ of $x_1,x_2,\ldots,x_p$ on the diagonal. The sample correlation matrix, $\mathbf{R}$, which contains the Pearson product-moment correlations (Pearson 1896), is an unbiased estimator of $\mathbf{P}$ and may be expressed as

$$\mathbf{R}=\mathbf{D}_{\mathbf{s}}^{-1}\mathbf{R}\mathbf{D}_{\mathbf{s}}^{-1}$$

where $\mathbf{D}_{\mathbf{s}}$ denotes a diagonal matrix with the sample standard deviations $s_1,s_2,\ldots,s_p$ of $x_1,x_2,\ldots,x_p$ on the diagonal.

Polyserial correlations

Suppose that the rows of $\mathbf{X}(n\times p)=\left[\mathbf{X}_o\quad\mathbf{X}_c\right]$ are n observations of p_o ordinal variables $x_1,x_2,\ldots,x_{p_o}$ with $m_1,m_2,\ldots,m_{p_o}$ categories, respectively and p_c continuous variables $x_1,x_2,\ldots,x_{p_c}$ as such that $p_o+p_c=p$. Suppose further that the p_o ordinal variables are the result of the discretization of the underlying p_o continuous standard normal variables $z_1,z_2,\ldots,z_{p_o}$ as such that $\mathbf{z}=\left[z_1\quad z_2\quad\ldots\quad z_{p_o}\right]'\sim N(\mathbf{0},\mathbf{P}_o)$ and

$$\begin{cases} x_i = 1 & \text{if } \tau_{i0} < z_i \leq \tau_{i1} \\ x_i = 2 & \text{if } \tau_{i1} < z_i \leq \tau_{i2} \\ & \vdots \\ x_i = m_i & \text{if } \tau_{i,m_i-1} < z_i \leq \tau_{i,m_i} \end{cases}$$

where $\mathbf{P}_o$ denotes the population correlation matrix of $\mathbf{z}$ and $-\infty = \tau_{i0} < \tau_{i1} < \tau_{i2} \ldots \leq \tau_{i,m_i} = \infty$ are parameters known as thresholds. The model for the univariate marginal of variable x_i is

$$\pi_{ik} = \int_{\tau_{i,k-1}}^{\tau_{i,}} \phi(u)du$$

where $\phi(.)$ denotes the probability density function of the standard normal distribution. The maximum likelihood estimator of τ_{ik} (Jöreskog, 1994) is given by

$$\hat{\tau}_{ik} = \mathbf{\Phi}^{-1}(p_{i1} + p_{i2} + \ldots + p_{ik})$$

where $\mathbf{\Phi}^{-1}(.)$ denotes the inverse of the cumulative distribution function of the standard normal distribution and $p_{i1}, p_{i2}, \ldots, p_{im_i}$ denote the marginal sample proportions for x_i .

If x_i denotes the i-th ordinal variable and x_j denotes the j-th continuous variable with mean μ_j and standard deviation σ_j and ρ_{ij} is the polyserial correlation of x_i and x_j, the corresponding bivariate log-likelihood function (Olsson, Drasgow, and Dorans 1982) is given by

$$l\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\mu}_j, \hat{\sigma}_j\right) = \sum_{m=1}^{n} \ln\left(\pi_{ikjm}\right) - \frac{n}{2}\left[\ln\left(2\pi\right) + \ln\left[\hat{\sigma}_j\right]\right] - \frac{1}{2}\sum_{m=1}^{n} z_{jm}^2$$

where

$$z_{jm} = \frac{x_{jm} - \hat{\mu}_j}{\hat{\sigma}_i}$$

and

$$\pi_{ijkm} = \Phi\left(\tau_{ikjm}^*\right) - \Phi\left(\tau_{i,k-1,jm}^*\right)$$

where k denotes the observed category of x_i , Φ denotes the cumulative distribution function of the standard normal distribution, and

$$\tau_{ikjm}^* = \frac{\hat{\tau}_{ik} - \rho_{ij} z_{jm}}{\sqrt{1 - \rho_{ij}^2}}$$

The maximization of the log-likelihood function is equivalent to minimizing the following discrepancy function

$$F\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\mu}_j, \hat{\sigma}_j\right) = -\sum_{m=1}^{n} \ln\left(\pi_{ikjm}\right)$$

The gradient of $F\left(.\right)$ follows as

$$g\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\mu}_j, \hat{\sigma}_j\right) = -\sum_{m=1}^{n} \frac{1}{\pi_{ikjm}} \frac{\partial \pi_{ikjm}}{\partial \rho_{ij}}$$

where (Olsson, Drasgow, and Dorans 1982)

$$\frac{\partial \pi_{ikjm}}{\partial \rho_{ij}} = (1-\rho_{ij}^2)^{-\frac{3}{2}} \left[\phi\left(\tau_{ikjm}^*\right)\left(\hat{\tau}_{ik}\rho_{ij} - z_{jm}\right) - \phi\left(\tau_{i,k-1,jm}^*\right)\left(\hat{\tau}_{i,k-1}\rho_{ij} - z_{jm}\right)\right]$$

where ϕ denotes the probability density function of the standard normal distribution. The information follows as

$$i\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\mu}_j, \hat{\sigma}_j\right) = \pi_{ikjm}^{-2} \left[\frac{\partial \pi_{ikjm}}{\partial \rho_{ij}}\right]^2$$

The Fisher scoring algorithm is used to minimize $F\left(\cdot\right)$ with respect to ρ_{ij}. Let $\theta = \rho_{ij}$. If $\hat{\theta}^{(t)}$ denotes the t^{th} successive approximation to $\hat{\theta}$, then the $(t+1)^{st}$ approximation is obtained from

$$\hat{\theta}^{(t+1)} = \hat{\theta}^{(t)} - \frac{g\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\mu}_j, \hat{\sigma}_j\right)}{i\left(\rho_{ij}, \hat{\boldsymbol{\tau}}_i, \hat{\mu}_j, \hat{\sigma}_j\right)}.$$

Iteration is terminated when the absolute gradient value is below the tolerance limit $\varepsilon = 10^{-3}$.

Mixed correlation and asymptotic covariance matrices

Suppose that the rows of $\mathbf{X}(n \times p) = \left[\mathbf{X}_o \quad \mathbf{X}_c\right]$ are n observations of p_o ordinal variables $x_1, x_2, \ldots, x_{p_o}$ with $m_1, m_2, \ldots, m_{p_o}$ categories, respectively and p_c continuous variables $x_1, x_2, \ldots, x_{p_c}$ as such that $p_o + p_c = p$. Let $\mathbf{R}_o(p_o \times p_o)$ denote the polychoric correlation matrix of the p_o ordinal variables, $\mathbf{R}_c(p_c \times p_c)$ denote the Pearson product-moment correlation matrix of the p_c continuous variables $x_1, x_2, \ldots, x_{p_c}$, and $\mathbf{R}_{oc}(p_o \times p_c)$ denote the polyserial correlation matrix of the ordinal and continuous variables. The correlation matrix, $\mathbf{R}$, of the ordinal and continuous variables may then be expressed as

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_o & \mathbf{R}_{oc} \\ \mathbf{R}'_{oc} & \mathbf{R}_c \end{bmatrix}$$

If F_{ij} denotes the discrepancy function which is minimized with respect to ρ_{ij} to obtain the maximum likelihood estimate of ρ_{ij}, then the asymptotic covariance matrix, $\boldsymbol{\Upsilon}$, of the polychoric, polyserial, and Pearson product-moment correlations (Muthen 1984) may be approximated by the matrix, $\mathbf{U}$, with typical element given by

$$u_{ij,kl} = n^{-1} \sum_{m=1}^{n} g_{ijm} g_{klm}$$

where g_{ijm} denotes the gradient of F_{ij} for observation m evaluated at $\rho_{ij} = r_{ij}$. If r_{ij} is a polychoric correlation, this gradient is given by

$$g_{ijm} = \frac{1}{\pi_{ijkl}}\left[\phi_2\left(\tau_{ik},\tau_{il}\right)-\phi_2\left(\tau_{i,k-1},\tau_{jl}\right)-\phi_2\left(\tau_{i,k},\tau_{j,l-1}\right)+\phi_2\left(\tau_{i,k-1},\tau_{j,l-1}\right)\right]$$

where $\phi_2(\cdot)$ denotes the density function of the bivariate standard normal distribution with correlation ρ_{ij} and k and l denote the observed category of x_i and x_j for observation m, respectively. In the case of a Pearson product-moment correlation, the gradient for observation m may be expressed as

$$g_{ijm} = \frac{z_{im}z_{jm} + \left(1 - z_{im}^2 - z_{jm}^2\right)r_{ij} + z_{im}z_{jm}r_{ij}^2 - r_{ij}^3}{\left(1 - r_{ij}^2\right)^2}$$

If r_{ij} denotes the polyserial correlation of ordinal variable x_i and continuous variable x_j, the gradient for observation m is given by

$$g_{ijm} = \frac{\phi\left(\tau_{ik}\right)\left(\tau_{ik}r_{ij} - z_{jm}\right) - \phi\left(\tau_{i,k-1}\right)\left(\tau_{i,k-1}r_{ij} - z_{jm}\right)}{\left(1 - r_{ij}^2\right)^{3/2}}$$

where ϕ denotes the probability density function of the standard normal distribution and k denotes the observed category of x_i.

Multiple Imputation

The MCMC method

Suppose now that the n observations of the p_o ordinal variables include missing data values with k_o missing data value patterns. The EM algorithm and the MCMC method for multiple imputation of incomplete data are intended for continuous variables and cannot readily be applied to ordinal variables. However, they can be applied to the underlying continuous variables $z_1, z_2, \ldots, z_{p_o}$ associated with the ordinal variables $x_1, x_2, \ldots, x_{p_o}$. Although no observations for these continuous variables are available, these variables are assumed to have a multivariate standard normal distribution with a population covariance matrix $\mathbf{\Sigma}_o$. As a result, we can simulate data from this distribution by using the polychoric correlation matrix of the complete data of the variables if the number of complete cases is large enough and use either the EM algorithm or the MCMC algorithm to impute the missing data values for the underlying continuous variables. After imputation, the estimated thresholds can be used to replace the missing data values for the corresponding ordinal variables by using the relationship between the ordinal variables, the underlying continuous variables, and the thresholds.

Suppose that the rows of $\mathbf{Z}(n\times p_o)$ are n observations of the p_o underlying continuous variables $z_1, z_2, \ldots, z_{p_o}$ simulated from the $N(\mathbf{0}, \mathbf{\Sigma}_o)$ distribution and that $\mathbf{Z}_o$ denotes the observed data values that corresponds with the observed data values of $\mathbf{X}_o$. The EM algorithm (Dempster, Laird, and Rubin 1977) can be used to compute the maximum likelihood estimate of $\mathbf{\Sigma}_o$. The minus two observed-data log likelihood may be expressed as

$$-2\ln L(\mathbf{\Sigma}_o \mid \mathbf{Z}_o) = \sum_{i=1}^{k_o} n_i \ln\left|\mathbf{\Sigma}_{oi}\right| + \sum_{i=1}^{k_o}\sum_{j=1}^{n_i} \mathbf{z}'_{oij}\mathbf{\Sigma}_{oi}^{-1}\mathbf{z}_{oij}$$

where n_i denotes the number of observations of missing data value pattern $i = 1,2,\cdots,k_o$, $\boldsymbol{\Sigma}_{oi}$ denotes the population covariance matrix for missing data value pattern i, and $\mathbf{z}_{oij}$ is the j^{th} vector of observed values of missing data value pattern i.

The initial estimate for the M-step is the sample covariance matrix, $\mathbf{S}_{p_o}$, of the complete ordinal data or $\mathbf{I}_{p_o}$ if the number of complete observations is too small. In the E-step, the conditional covariance matrices of the missing variables given the observed variables for the missing data value patterns are computed and used to compute an updated estimate, $\hat{\boldsymbol{\Sigma}}_o^{(t+1)}$ of $\boldsymbol{\Sigma}_o$. Iteration of the consecutive M and E steps is terminated when the absolute difference between $\hat{\boldsymbol{\Sigma}}_o^{(t+1)}$ and $\hat{\boldsymbol{\Sigma}}_o^{(t)}$ is below the tolerance limit $\varepsilon = 10^{-5}$.

The correlation matrix of the EM estimate, $\hat{\boldsymbol{\Sigma}}_o$, of $\boldsymbol{\Sigma}_o$ is used as the initial covariance matrix of the multivariate standard normal distribution in the first step of the Monte Carlo Markov Chain (MCMC) method. In the first step (P-step) of the MCMC method, an estimate of $\boldsymbol{\Sigma}_o$ is simulated form an inverse Wishart distribution. In the I-step, observations are simulated from the conditional standard normal distributions of the missing variables given the observed k missing data value patterns and used to replace the missing data values. The next estimate of $\boldsymbol{\Sigma}_o$ is then obtained by computing the sample correlation matrix of the completed data. The P and I steps are repeated for a fixed number of times.

Let the rows of $\mathbf{Z}_i(n \times p)$ contain the observed and the imputed data values for the standard normal variables $z_1, z_2, \ldots, z_{p_o}$. The observed data for the ordinal variables are obtained from the corresponding observed data values of $\mathbf{X}_o$. The missing data values of $\mathbf{X}_o$ are then replaced by the values obtained from the corresponding imputed data values of $\mathbf{Z}$ and the estimated thresholds by using the relationship between the ordinal variables, the underlying continuous variables, and the thresholds.

Suppose further that the n observations of the p_c continuous variables include missing data values with k_c missing data value patterns and that the joint distribution of the variables is a multivariate normal distribution with mean vector $\boldsymbol{\mu}_c$ and covariance matrix $\boldsymbol{\Sigma}_c$. The EM algorithm and the MCMC method for multiple imputation of incomplete data can be used to impute the missing data values of the continuous variables.

Suppose that $\mathbf{X}_{co}$ denote the observed data values. The EM algorithm (Dempster, Laird, and Rubin 1977) can be used to compute the maximum likelihood estimate of $\boldsymbol{\Sigma}_c$. The minus two observed-data log likelihood may be expressed as

$$-2\ln L(\boldsymbol{\Sigma}_c \mid \mathbf{X}_{co}) = \sum_{i=1}^{k} n_i \ln\left|\boldsymbol{\Sigma}_{ci}\right| + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\mathbf{x}_{coij} - \boldsymbol{\mu}_{ci}\right)' \boldsymbol{\Sigma}_{ci}^{-1}\left(\mathbf{x}_{coij} - \boldsymbol{\mu}_{ci}\right)$$

where n_i denotes the number of observations of missing data value pattern $i = 1, 2, \ldots, k_c$, $\boldsymbol{\Sigma}_{ci}$ denotes the population covariance matrix of missing data value pattern i, $\boldsymbol{\mu}_{ci}$ denotes the mean vector of missing data value pattern i, and $\mathbf{x}_{coij}$ is the j^{th} vector of observed values of missing data value pattern i.

The initial estimate for the M-step is the sample covariance matrix, $\mathbf{S}_{p_c}$, of the complete data or $\mathbf{I}_{p_c}$ if the number of complete observations is too small. In the E-step, the conditional covariance matrices of the missing variables given the observed variables of the missing data value patterns are computed and used to compute an updated estimate $\hat{\boldsymbol{\Sigma}}_c^{(t+1)}$ of $\boldsymbol{\Sigma}_c$. Iteration of the consecutive M and E steps is terminated when the absolute difference between $\hat{\boldsymbol{\Sigma}}_c^{(t+1)}$ and $\hat{\boldsymbol{\Sigma}}_o^{(t)}$ is below the tolerance limit $\varepsilon = 10^{-5}$.

The EM estimate, $\hat{\boldsymbol{\Sigma}}_c$, of $\boldsymbol{\Sigma}_c$ is used as the initial covariance matrix of the multivariate normal distribution in the first step of the Monte Carlo Markov Chain (MCMC) method. In the first step (P-step) of the MCMC method, an estimate of $\boldsymbol{\Sigma}_c$ is simulated from an inverse Wishart distribution. In the I-step, observations are simulated from the conditional normal distributions of the missing variables given the observed k_c missing data value patterns and used to replace the missing data values. The next estimate of $\boldsymbol{\Sigma}_c$ is then obtained by computing the sample covariance matrix of the completed data. The P and I steps are repeated for a fixed number of times.

The FCS ordinal logistic regression method

Suppose that the n observations of the p_o ordinal variables include missing data values and that a joint (multivariate) distribution of the variables exists. In this case, the Fully Conditional Specified (FCS) ordinal logistic regression method (Brand 1999; Van Buuren 2007) can be used to impute the missing data values. The FCS ordinal logistic regression method performs a fixed number of imputations to impute the missing data values. Each imputation consists of a filled-in phase and an imputation phase. In the filled-in phase, the missing data values are filled-in by using a sequence of ordinal logistic regression analyses for the p_o ordinal variables. These filled-in data are then used as the initial data for the imputation phase in which the missing data values are imputed by using a sequence of ordinal logistic regression analyses for the p_o ordinal variables. These imputed data are then used as the initial data for the next iteration of the imputation phase and a fixed number of iterations are executed for each imputation.

The filled-in stage fits the following p_o ordinal logistic regression models sequentially to the data, namely

$$\begin{aligned}
&\text{logit}(\pi_{1k}) = \alpha_{1k} \\
&\text{logit}(\pi_{2k}) = \alpha_{2k} + \beta_{21}x_1 \\
&\text{logit}(\pi_{3k}) = \alpha_{3k} + \beta_{31}x_1 + \beta_{32}x_2 \\
&\vdots \\
&\text{logit}(\pi_{p_o k}) = \alpha_{p_o k} + \beta_{p_o 1}x_1 + \beta_{p_o 2}x_2 + \cdots + \beta_{p_o, p_o-1}x_{p_o-1}
\end{aligned}$$

where $\pi_{ijk} = P(x_j \le k \,|\, x_1, x_2, \ldots, x_{j-1})$, $\text{logit}(\pi_{jk}) = \ln(\pi_{jk}) - \ln(\pi_{im_j})$, and the elements of $\boldsymbol{\gamma} = \left[\alpha_{11}\ \alpha_{12} \ldots \beta_{21} \ldots \beta_{p_o, p_o-1}\right]'$ denote unknown regression weights. The first model is fitted to the complete data for x_1. The corresponding estimates are then used to simulate new parameter values from the posterior distribution of the parameters which in turn is used to fill-in the missing data values for x_1. The second model is then fitted to the complete data for x_2 and the filled-in data for x_1. The final model is fitted to the complete data for x_{p_o} and the filled-in data for $x_1, x_2, \ldots, x_{p_o-1}$. The filled-in data for $x_1, x_2, \ldots, x_{p_o}$ are used for the first iteration of the imputation phase. The simulation of the new parameter values from the posterior distribution of the parameters and the imputation of the missing data values for each of the p ordinal logistic regression models use the same steps as outlined next for each iteration of the imputation stage.

For each iteration of the imputation stage, the following ordinal logistic regression models are fitted sequentially either to the filled-in data or the imputed data, namely

$$\operatorname{logit}(\pi_{ijk}) = \alpha_k + \beta_1 x_1 + \cdots + \beta_{j-1} x_{j-1} + \beta_{j+1} x_{j+1} + \cdots + \beta_{p_o} x_{p_o}$$

where $\pi_{ijk} = P(x_j \le k \mid x_1, \ldots, x_{j-1}, x_{j+1} \ldots, x_{p_o})$, $\operatorname{logit}(\pi_{jk}) = \ln(\pi_{jk}) - \ln(\pi_{im_j})$, the elements of $\boldsymbol{\gamma}_j = \left[\alpha_1\ \alpha_2 \ldots \alpha_{m_i-1}\ \beta_1 \ldots \beta_{j-1} \beta_{j+1} \ldots \beta_{p_o}\right]'$ denote $p_o + m_j - 1$ unknown regression weights, $j = 1, 2, ..., p_o$, and $k = 1, 2, ..., m_j - 1$. Let $\mathbf{V}_j$ denote the estimated covariance matrix of the estimator $\hat{\boldsymbol{\gamma}}_j$ of $\boldsymbol{\gamma}_j$.

New values for the parameters are then simulated from their posterior distribution as

$$\boldsymbol{\gamma}_{jt} = \hat{\boldsymbol{\gamma}}_j + \mathbf{V}'_{hj}\mathbf{z}$$

where $\mathbf{V}'_{hj}$ denotes the upper triangular matrix in the Cholesky decomposition of $\mathbf{V}_j = \mathbf{V}'_{hj}\mathbf{V}_{hj}$, and $\mathbf{z}$ is a $(p_o + m_j - 1) \times 1$ standard normal vector. These new parameter values are then used to compute the predicted cumulative probability $\hat{\pi}_{jk}$ for $k = 1, 2, ..., m_j - 1$. A random uniform variable, u, between 0 and 1 is simulated and the missing data values for x_j are imputed as 1 if $u < \hat{\pi}_{j1}$, as k if $\hat{\pi}_{j,k-1} \le u < \hat{\pi}_{jk}$, and as m_j if $u \ge \hat{\pi}_{jm_j}$.

The FCS regression method

Suppose now that the n observations of the p_c continuous variables include missing data values and that a joint (multivariate) distribution of the variables exists. In this case, the Fully Conditional Specified (FCS) regression method (Brand 1999; Van Buuren 2007) can be used to impute the missing data values. The FCS regression method performs a fixed number of imputations to impute the missing data values. Each imputation consists of a filled-in phase and an imputation phase. In the filled-in phase, the missing data values are filled-in by using a sequence of regression analyses for the p_c continuous variables. These filled-in data are then used as the initial data for the imputation phase in which the missing data values are imputed by using a sequence of regression analyses for the p_c continuous variables. These imputed data are then used as the initial data for the next iteration of the imputation phase and a fixed number of iterations are executed for each imputation.

The filled-in stage fits the following p_c regression models sequentially to the data, namely

$$\begin{aligned}
x_1 &= \beta_{01} + e_1 \\
x_2 &= \beta_{02} + \beta_{21} x_1 + e_2 \\
x_3 &= \beta_{03} + \beta_{31} x_1 + \beta_{32} x_2 + e_3 \\
&\vdots \\
x_{p_c} &= \beta_{0 p_c} + \beta_{p_c 1} x_1 + \beta_{p_c 2} x_2 + \cdots + \beta_{p_c, p_c - 1} x_{p_c - 1} + e_{p_c}
\end{aligned}$$

where the elements of $\boldsymbol{\beta} = \left[\beta_{01}\ \beta_{02} \cdots \beta_{p_c, p_c-1}\right]'$ denote unknown regression weights and $e_1, e_2, \ldots, e_{p_c}$ are p_c error variables. The first model is fitted to the complete data for x_1. The corresponding estimates are then used to simulate new parameter values from the posterior distributions of the parameters which in turn is used to fill-in the missing data values for x_1. The second model is then fitted to the complete data for x_2 and the filled-in data for x_1. The final model is fitted to the complete data for x_{p_c} and the filled-in data for $x_1, x_2, \ldots, x_{p_c-1}$. The filled-in data for $x_1, x_2, \ldots, x_{p_c}$ are used for the first iteration of the imputation phase. The simulation of the new parameter values from the posterior distributions of the parameters and the imputation of the missing data values for each of the p_c regression models use the same steps as outlined next for each iteration of the imputation stage.

For each iteration of the imputation stage, the following regression models are fitted sequentially either to the filled-in data or the imputed data, namely

$$x_j = \beta_0 + \beta_1 x_1 + \cdots + \beta_{j-1} x_{j-1} + \beta_{j+1} x_{j+1} + \cdots + \beta_{p_c} x_{p_c} + e_j$$

where $j = 1, 2, \ldots, p_c$, the elements of $\boldsymbol{\beta}_j = \left[\beta_0\ \beta_1 \ldots \beta_{j-1}\beta_{j+1} \ldots \beta_{p_c}\right]'$ denote p_c unknown regression weights, and e_j denotes an error variable with variance σ_j^2. The estimated covariance matrix of the estimator $\hat{\boldsymbol{\beta}}_j$ of $\boldsymbol{\beta}_j$ may be expressed as

$$\sigma_j^2 \mathbf{V}_j = \sigma_j^2 \left(\mathbf{X}'_{c(j)} \mathbf{X}_{c(j)}\right)^{-1}$$

where $\mathbf{X}_{c(j)}$ denotes rows $1, 2, \ldots, j-1, j, \ldots, p_c$ of the filled-in or imputed data. New values for the parameters are then simulated from their posterior distributions as

$$\boldsymbol{\beta}_{jt} = \hat{\boldsymbol{\beta}}_j + \sigma_{tj}^2 \mathbf{V}'_{hj} \mathbf{z}$$
$$\sigma_{tj}^2 = \frac{\hat{\sigma}_j^2 \left(n_j - p_c\right)}{c}$$

where $\mathbf{V}_{hj}$ denotes the upper triangular matrix in the Cholesky decomposition of $\mathbf{V}_j = \mathbf{V}'_{hj} \mathbf{V}_{hj}$, $\mathbf{z}$ denotes a $p_c \times 1$ standard normal vector, and c is a Chi-square variable with $n_j - p_c$ degrees of freedom. The missing data values are then imputed as

$$x_{cijm} = \boldsymbol{\beta}'_{jt} \mathbf{x}_{ci(j)} + \sigma_{ij} z$$

where x_{cijm} denotes a missing data value in row i and column j of $\mathbf{X}_c$, $\mathbf{x}_{ci(j)}$ denotes row i of $\mathbf{X}_{c(j)}$, and z is a standard normal variable.

Average moment matrices

Suppose that $\mathbf{X}_{1i}, \mathbf{X}_{2i}, \ldots, \mathbf{X}_{mi}$ are m imputed data sets for the incomplete data matrix, $\mathbf{X}$, of the of p_o ordinal variables $x_1, x_2, \ldots, x_{p_o}$ and the p_c continuous variables $x_1, x_2, \ldots, x_{p_c}$ and that $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_m$ and $\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_m$ denote the corresponding mixed correlation matrices and the estimated asymptotic covariance matrices of the mixed correlations, respectively. Then, the average mixes correlation matrix is

$$\bar{\mathbf{R}} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{R}_i$$

and the average estimated asymptotic covariance matrix is

$$\bar{\mathbf{U}} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{U}_i$$

Chung and Cai (2019) point out that $\bar{\mathbf{U}}$ only captures uncertainty based on complete data. As a result, its inverse cannot be used as a weight matrix for the robust DWLS, WLS, and ULS methods for structural equational modeling. A corrected weight matrix is obtained by correcting for the between-imputation variation in the estimated mixed correlations and is obtained as the inverse of

$$\hat{\boldsymbol{\Upsilon}} = \bar{\mathbf{U}} + \frac{m+1}{m(m-1)}\left[\sum_{i=1}^{m}\left(\mathbf{r}_i - \bar{\mathbf{r}}\right)\left(\mathbf{r}_i - \bar{\mathbf{r}}\right)'\right]$$

where $\mathbf{r}$ denotes the $p \times (p-1)/2$ vector consisting of the nondiagonal and nonduplicated elements of the $p \times p$ symmetric matrix $\mathbf{R}$. $\bar{\mathbf{R}}$ and $\hat{\boldsymbol{\Upsilon}}$ can be used to fit structural equation models to the average mixed correlation matrix with the robust DWLS, WLS, and ULS methods. The corrected robust DWLS and ULS Chi-square test statistic proposed by Chung and Cai (2019) is given by

$$T_B = (n-1)(\bar{\mathbf{r}} - \boldsymbol{\rho}(\hat{\boldsymbol{\theta}}))'\mathbf{V}(\bar{\mathbf{r}} - \boldsymbol{\rho}(\hat{\boldsymbol{\theta}}))$$

where

$$\mathbf{V} = \hat{\boldsymbol{\Upsilon}}^{-1} - \hat{\boldsymbol{\Upsilon}}^{-1}\hat{\boldsymbol{\Delta}}(\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Delta}})^{-1}\hat{\boldsymbol{\Delta}}'\hat{\boldsymbol{\Upsilon}}^{-1}$$

where $\hat{\boldsymbol{\Delta}}$ denotes the Jacobian matrix of $\boldsymbol{\rho}(\boldsymbol{\theta})$ with respect to the unknown parameters, $\boldsymbol{\theta}$, of the structural equation model evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. The small sample adjusted T_B test statistic (Yuan and Bentler 1997) is given by

$$T_{YB} = \frac{T_B}{1 + nT_B/(n-1)}.$$

1.4 Measurement model for visual and verbal ability

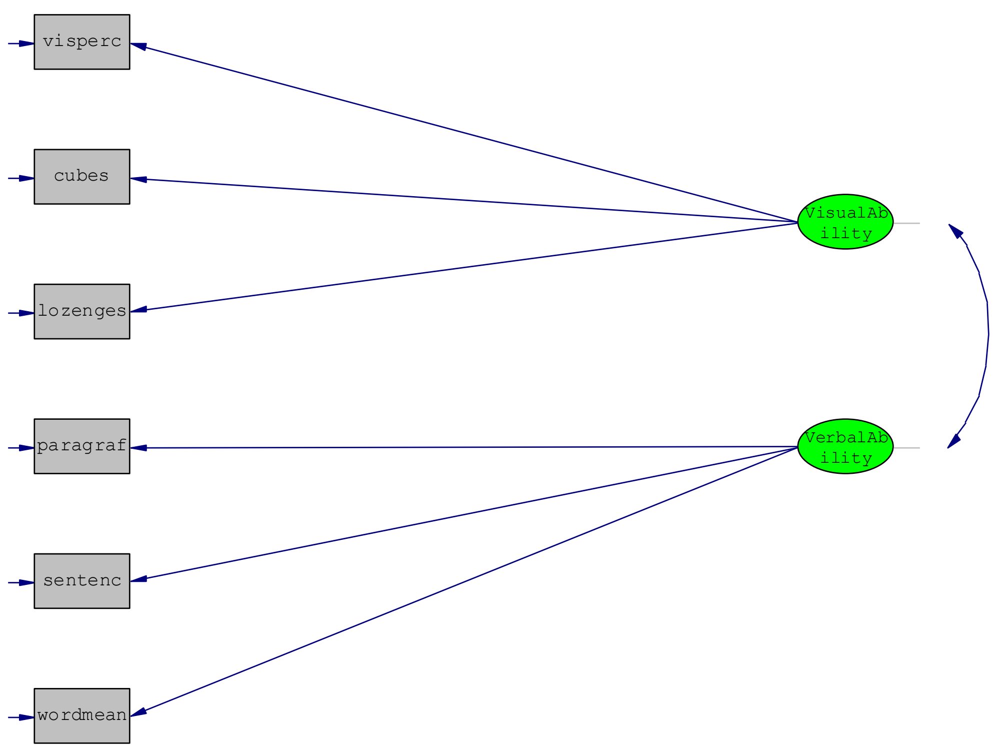
The data are the simulated scores of 1250 girls on six psychological tests (visual perception, cubes, lozenges, paragraph completion, sentence completion, and word meaning). The corresponding data file is **GIRLS.LSF**, and the first few observations are shown in the image below.

GIRLS.lsf

	visperc	cubes	lozenges	paragraf	sentenc	wordmean
1	-0.71	-1.96	-6.87	-999999.00	-0.24	7.60
2	-6.42	-8.64	-4.37	0.59	1.53	9.63
3	8.79	-2.15	10.21	2.29	2.81	8.92
4	9.72	-2.38	14.83	0.48	4.17	6.80
5	-999999.00	0.65	-8.02	-999999.00	-7.51	-12.61
6	-10.57	-999999.00	-2.18	0.48	6.70	-999999.00
7	-999999.00	7.80	-999999.00	-999999.00	-1.22	5.38
8	-999999.00	-2.14	5.24	-1.01	2.72	3.73
9	2.60	5.72	-0.64	-0.04	-1.32	-999999.00
10	0.73	-7.94	-6.89	-999999.00	-999999.00	-27.54
11	-3.66	-4.20	3.51	6.26	9.25	12.35
12	-6.26	0.88	-14.83	4.86	0.88	12.27
13	6.53	-999999.00	-8.80	-999999.00	-999999.00	-1.79
14	8.77	-3.62	-999999.00	-3.47	-0.57	-999999.00
15	-2.61	-6.94	1.41	-0.65	-5.81	-999999.00

Note that the data values of -999999.00 are missing data values. If a different global missing data value code is used, it should be assigned using the **Define Variables** dialog box.

The theoretical model is a measurement model that specifies that the six psychological tests are indicators of visual ability and verbal ability of Junior High students. A path diagram for this model is depicted in the image below.



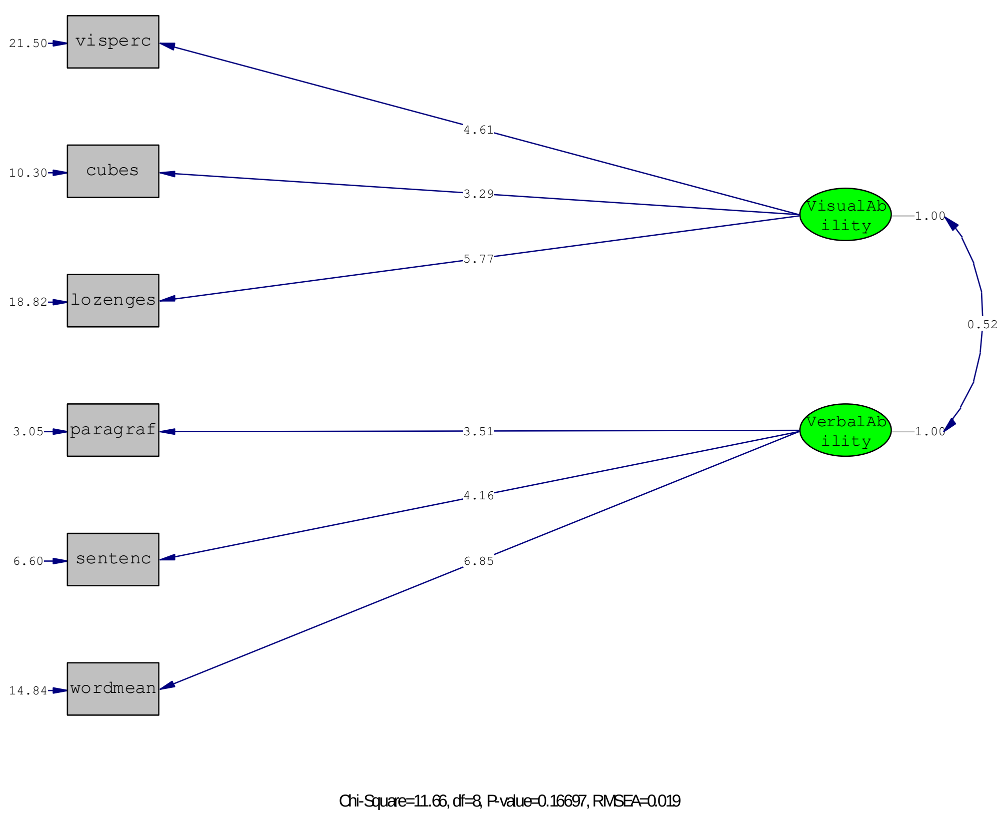
The SIMPLIS syntax file to fit the theoretical model to the average sample covariance matrix of 30 MCMC imputations is shown in the image below.

GIRLS4A.SPL

```
Raw Data from File GIRLS.LSF
Latent Variables
VisualAbility VerbalAbility
Relationships
visperc cubes lozenges = VisualAbility
paragraf sentenc wordmean = VerbalAbility
LISREL Output: SC ME=ML MI2S NM=30 IX=103829 IM=MC
Path Diagram
End of Problem
```

- Line 1 specifies the data file.
- Lines 2 and 3 specify the labels for the two latent variables.
- Lines 4 to 6 specify the measurement model for the six psychological tests.
- Line 7 requests that the results in the output file should be given in terms of the LISREL model for the measurement model (LISREL Output). It also requests that the completely standardized solution should be written to the output file (SC), and robust maximum likelihood estimation (ME = ML). The MI2S option invokes the two-stage multiple imputation SEM method to fit the model to the average sample covariance matrix of the NM = 30 MCMC imputations (IM = MC) based on an initial random seed of IX = 103829.
- Line 8 requests a path diagram of the model.
- Line 9 indicates that no more SIMPLIS commands are to be processed.

When the SPL file above is opened in LISREL and the **Run LISREL** icon is clicked, the following path diagram is obtained.



The corresponding output file, **GIRLS4A.OUT**, is opened in a separate window. The Chi-square test statistic values listed in this file are shown in the image below.

GIRLS4A.OUT

```
                        Goodness-of-Fit Statistics

Degrees of Freedom for (C1)-(C3),(C5)             8
Maximum Likelihood Ratio Chi-Square (C1)          11.804 (P = 0.16019)
Browne's (1984) ADF Chi-Square (C2_NT)            11.666 (P = 0.16673)
Browne's (1984) ADF Chi-Square (C2_NNT)           12.005 (P = 0.15098)
Satorra-Bentler (1988) Scaled Chi-Square (C3)     11.661 (P = 0.16697)
Satorra-Bentler (1988) Adjusted Chi-Square (C4)   11.490 (P = 0.16815)
Degrees of Freedom for C4                          7.882
Chi-Square Scaled and Shifted (C5)                11.634 (P = 0.16828)
P-Value of C1 under Non-Normality                        = 0.1677
```

These Chi-square test statistic values indicate that the theoretical measurement model for numerical and verbal ability is supported by the data.

1.5 Two-wave model for political efficacy and political responsiveness

This example is based on panel data of the six political efficacy measurements described in Aish and Jöreskog (1990) observed in two different calendar years. The data file, **PANELUSA.LSF**, consists of 933 cases obtained in a USA sample. The first few observations of this data file are shown below.

panelusa.LSF

	NOSAY1	VOTING1	COMPLEX1	NOCARE1	TOUCH1	INTERES1	NOSAY2	VOTING2
1	2.00	2.00	1.00	1.00	1.00	1.00	-999999.00	2.00
2	2.00	3.00	3.00	3.00	2.00	3.00	2.00	3.00
3	3.00	2.00	2.00	3.00	3.00	3.00	3.00	2.00
4	2.00	2.00	1.00	1.00	2.00	1.00	2.00	2.00
5	3.00	2.00	2.00	3.00	3.00	3.00	3.00	2.00
6	2.00	2.00	2.00	2.00	1.00	2.00	3.00	2.00
7	3.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00
8	2.00	1.00	2.00	2.00	1.00	1.00	3.00	3.00
9	3.00	3.00	2.00	2.00	3.00	3.00	3.00	3.00
10	2.00	2.00	3.00	1.00	1.00	1.00	2.00	2.00
11	3.00	2.00	1.00	1.00	2.00	2.00	3.00	2.00
12	1.00	1.00	1.00	1.00	1.00	1.00	3.00	3.00
13	2.00	2.00	2.00	1.00	2.00	2.00	1.00	1.00
14	3.00	3.00	2.00	3.00	2.00	2.00	3.00	2.00
15	3.00	3.00	3.00	3.00	3.00	3.00	2.00	3.00
16	3.00	3.00	2.00	2.00	3.00	2.00	2.00	2.00
17	3.00	3.00	4.00	2.00	1.00	1.00	2.00	2.00
18	4.00	2.00	3.00	4.00	-999999.00	-999999.00	3.00	2.00

The data values of -999999.00 are missing data values. If a different global missing data value code is used, it should be assigned using the **Define Variables** dialog box.

The data are the responses to the following statements:

- People like me have no say in what the government does (NOSAY)
- Voting is the only way that people like me can have any say about how the government runs things (VOTING)
- Sometimes politics and government seem so complicated that a person like me cannot really understand what is
- going on (COMPLEX)
- I don't think that public officials care much about what people like me think (NOCARE)
- Generally speaking, those we elect to Parliament lose touch with the people pretty quickly (TOUCH)

- Parties are only interested in people's votes but not in their opinions (INTEREST)

The ordered categories are:

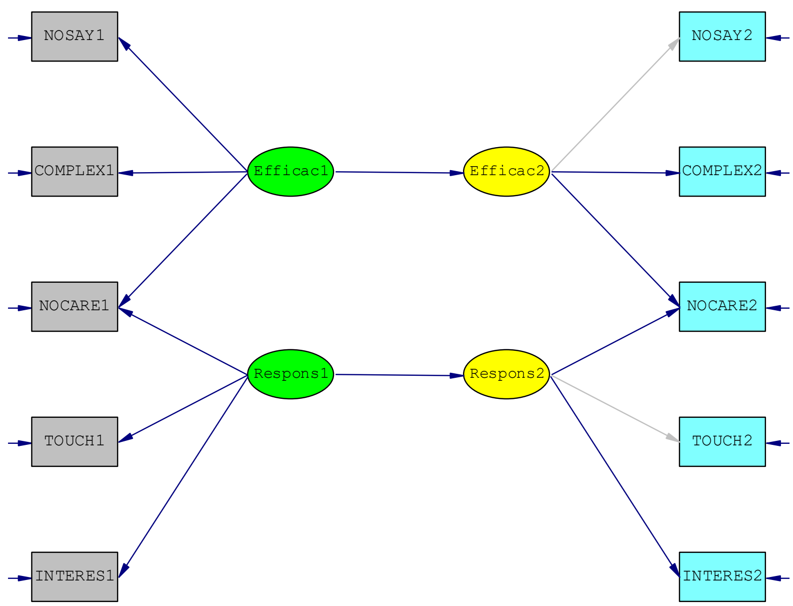
1: agree strongly

2: agree

3: disagree

4: disagree strongly

The theoretical model is a two-wave model for political efficacy and political responsiveness. A path diagram of the theoretical model is shown in the image below.

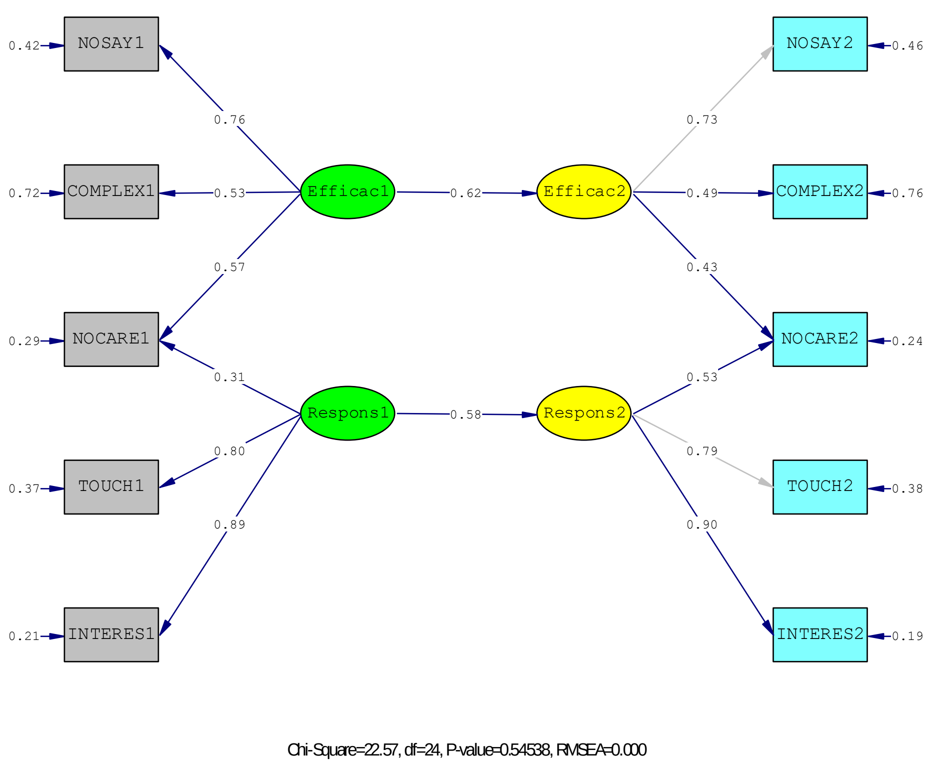


The SIMPLIS syntax file to fit the model reflected in the path diagram above to the average polychoric correlation matrix of 10 FCS imputations is depicted in the image below. The two-stage multiple imputation SEM syntax is reflected on the LISREL Output command as MI2S which requests the method, NM = 10 which requests 10 FCS imputations (IM = FC), and IX = 18957 which requests a starting random seed of 18957.

```
Raw Data from File panelusa.lsf
Latent Variables
Efficac1 Respons1 Efficac2 Respons2
Relationships
NOSAY1 COMPLEX1 NOCARE1 = Efficac1
NOCARE1 TOUCH1 INTERES1 = Respons1
NOSAY2 COMPLEX2 NOCARE2 = Efficac2
NOCARE2 TOUCH2 INTERES2 = Respons2
Let the errors of NOSAY1 and NOSAY2 correlate
Let the errors of COMPLEX1 and COMPLEX2 correlate
Let the errors of NOCARE1 and NOCARE2 correlate
Let the errors of TOUCH1 and TOUCH2 correlate
Let the errors of INTERES1 and INTERES2 correlate
Efficac2 = Efficac1
Respons2 = Respons1
Let the errors of Efficac2 and Respons2 correlate
LISREL Output: SC MI2S ME=WLS IX=18957 NM=10 IM=FC
Path Diagram
End of Problem
```

- Line 1 specifies the raw data file.
- Lines 2 and 3 specify labels for the latent variables of the model.
- Lines 4 to 16 specify the two-wave model for political efficacy and political responsiveness.
- Line 17 requests that the results in the output file should be given in terms of the LISREL model for the structural equation model (LISREL Output). It also requests that the completely standardized solution should be written to the output file (SC) and weighted least squares estimation (ME = WLS).
- Line 18 requests a path diagram of the model.
- Line 19 indicates that no more SIMPLIS commands are to be processed.

When the SPL file above is opened in LISREL and the **Run LISREL** icon is clicked, the following path diagram is opened.



The corresponding output file, **PANELUSA4A.OUT**, is opened in a separate window. A small portion of this file is shown in the following image.

```
PANELUSA4A.OUT
                          Goodness-of-Fit Statistics

 Degrees of Freedom for C(1),C(6)                         24
 Weighted Least Squares Chi-Square (C1)                   22.568 (P = 0.54538)
 Yuan-Bentler (1997) Chi-Square for C1 (C6)               22.035 (P = 0.57718)

 Estimated Non-centrality Parameter (NCP)                 0.0
 90 Percent Confidence Interval for NCP                   (0.0 ; 13.674)

 Minimum Fit Function Value                               0.0242
 Population Discrepancy Function Value (F0)               0.0
 90 Percent Confidence Interval for F0                    (0.0 ; 0.0147)
 Root Mean Square Error of Approximation (RMSEA)          0.0
 90 Percent Confidence Interval for RMSEA                 (0.0 ; 0.0247)
 P-Value for Test of Close Fit (RMSEA < 0.05)             1.00
```

These goodness-of-fit statistic values indicate that the theoretical two-wave model for political efficacy and political responsiveness is supported by the data.

.

2. Standard error and confidence interval estimates

2.1 Standard error estimates for standardized solutions

The LISREL model for observed and latent variables

The LISREL model (Jöreskog 1973, 1977) for population covariance matrices may be expressed as

$$\mathbf{y} = \mathbf{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

$$\mathbf{x} = \mathbf{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}$$

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \mathbf{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}$$

where $\mathbf{y}$ and $\mathbf{x}$ denote p_η and p_ξ indicators of the m_η endogenous latent variables, $\boldsymbol{\eta}$, and the m_ξ exogenous latent variables, $\boldsymbol{\xi}$, respectively, $\mathbf{\Lambda}_y$ and $\mathbf{\Lambda}_x$ are $p_y \times m_\eta$ and $p_x \times m_\xi$ matrices of factor loadings, respectively, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$ denote p_η and p_ξ measurement errors, respectively, $\mathbf{B}$ and $\mathbf{\Gamma}$ are $m_\eta \times m_\eta$ and $m_\eta \times m_\xi$ matrices of regression weights, respectively, and the elements of $\boldsymbol{\zeta}$ denote m_η error variables.

The $t \times 1$ vector, $\mathbf{z}$, consisting of all the variables of the LISREL model follows as

$$\mathbf{z} = \begin{pmatrix} \mathbf{y} \\ \mathbf{x} \\ \boldsymbol{\eta} \\ \boldsymbol{\xi} \\ \boldsymbol{\varepsilon} \\ \boldsymbol{\delta} \\ \boldsymbol{\zeta} \end{pmatrix}$$

The model for the relationships between all the variables of the LISREL model may then be expressed as

$$\mathbf{z} = \mathbf{B}_t \mathbf{z} + \mathbf{z}_e$$

where

$$\mathbf{B}_t = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \boldsymbol{\Lambda}_y & \mathbf{0} & \mathbf{I}_{p_y} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Lambda}_x & \mathbf{0} & \mathbf{I}_{p_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B} & \boldsymbol{\Gamma} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{m_\eta} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

where $\mathbf{I}_n$ denotes the $n \times n$ identity matrix and

$$\mathbf{z}_e = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \boldsymbol{\xi} \\ \boldsymbol{\varepsilon} \\ \boldsymbol{\delta} \\ \boldsymbol{\zeta} \end{pmatrix}$$

The covariance matrix, $\boldsymbol{\Phi}_t$, of $\mathbf{z}_e$ follows as

$$\boldsymbol{\Phi}_t = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Phi} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Theta}_\varepsilon & \boldsymbol{\Theta}_{\varepsilon\delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Theta}_{\delta\varepsilon} & \boldsymbol{\Theta}_\delta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Psi} \end{pmatrix}$$

where $\mathbf{\Phi}$, $\mathbf{\Theta}_\varepsilon$, $\mathbf{\Theta}_\delta$, and $\mathbf{\Psi}$ denote the covariance matrices of $\boldsymbol{\xi}$, $\boldsymbol{\varepsilon}$, $\boldsymbol{\delta}$, and $\boldsymbol{\zeta}$, respectively and $\mathbf{\Theta}_{\varepsilon\delta} = \mathbf{\Theta}'_{\delta\varepsilon}$ denotes the covariance matrix between $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$. The $t \times t$ covariance matrix of $\mathbf{z}$, Υ_t, may then be expressed as

$$\Upsilon_t = \left(\mathbf{I}_t - \mathbf{B}_t\right)^{-1} \mathbf{\Phi}_t \left(\mathbf{I}_t - \mathbf{B}_t\right)^{-1\prime} = \mathbf{\Lambda}_t \mathbf{\Phi}_t \mathbf{\Lambda}'_t$$

where $\mathbf{\Lambda}_t = \left(\mathbf{I}_t - \mathbf{B}_t\right)^{-1}$. The $(p+m)\times(p+m)$ covariance matrix, Υ, of the $p = p_y + p_x$ observed variables and $m = m_\eta + m_\xi$ latent variables, follows as

$$\Upsilon = \left[\mathbf{I}_{p+m} \quad \mathbf{0}\right]\left(\mathbf{I}_t - \mathbf{B}_t\right)^{-1} \mathbf{\Phi}_t \left(\mathbf{I}_t - \mathbf{B}_t\right)^{-1\prime} \left[\mathbf{I}_{p+m} \quad \mathbf{0}\right]' = \mathbf{\Lambda} \mathbf{\Phi}_t \mathbf{\Lambda}'$$

where $\mathbf{\Lambda}$ denotes the $(p+m)\times t$ matrix consisting of the first $p+m$ rows of $\mathbf{\Lambda}_t$. In terms of the parameter matrices of the LISREL model, we obtain that

$$\Upsilon = \begin{pmatrix} \mathbf{\Sigma}_{yy} & \mathbf{\Sigma}_{yx} & \mathbf{\Sigma}_{y\eta} & \mathbf{\Sigma}_{y\xi} \\ \mathbf{\Sigma}_{xy} & \mathbf{\Sigma}_{xx} & \mathbf{\Sigma}_{x\eta} & \mathbf{\Sigma}_{x\xi} \\ \mathbf{\Sigma}_{\eta y} & \mathbf{\Sigma}_{\eta x} & \mathbf{\Sigma}_{\eta\eta} & \mathbf{\Sigma}_{\eta\xi} \\ \mathbf{\Sigma}_{\xi y} & \mathbf{\Sigma}_{\xi x} & \mathbf{\Sigma}_{\xi\eta} & \mathbf{\Sigma}_{\xi\xi} \end{pmatrix}$$

where

$$\mathbf{\Sigma}_{yy} = \mathbf{\Lambda}_y \left((\mathbf{I}-\mathbf{B})^{-1} \mathbf{\Gamma\Phi\Gamma}'(\mathbf{I}-\mathbf{B})^{-1\prime} + \mathbf{\Psi}^* \right) \mathbf{\Lambda}'_y + \mathbf{\Theta}_\varepsilon$$

$$\mathbf{\Sigma}_{xy} = \mathbf{\Sigma}'_{yx} = \mathbf{\Lambda}_x \mathbf{\Phi\Gamma}' \left(\mathbf{I}-\mathbf{B}\right)^{-1\prime} \mathbf{\Lambda}'_y + \mathbf{\Theta}_{\delta\varepsilon}$$

$$\mathbf{\Sigma}_{xx} = \mathbf{\Lambda}_x \mathbf{\Phi} \mathbf{\Lambda}'_x + \mathbf{\Theta}_\delta$$

$$\mathbf{\Sigma}_{\eta y} = \left((\mathbf{I}-\mathbf{B})^{-1} \mathbf{\Gamma\Phi\Gamma}'(\mathbf{I}-\mathbf{B})^{-1\prime} + \mathbf{\Psi}^* \right) \mathbf{\Lambda}'_y$$

$$\mathbf{\Sigma}_{\eta x} = \mathbf{\Sigma}'_{\eta x} = \left(\mathbf{I}-\mathbf{B}\right)^{-1} \mathbf{\Gamma\Phi\Lambda}'_x$$

$$\mathbf{\Sigma}_{\eta\eta} = (\mathbf{I}-\mathbf{B})^{-1} \mathbf{\Gamma\Phi\Gamma}'(\mathbf{I}-\mathbf{B})^{-1\prime} + \mathbf{\Psi}^*$$

$$\mathbf{\Sigma}_{\xi y} = \mathbf{\Phi\Gamma}'(\mathbf{I}-\mathbf{B})^{-1\prime} \mathbf{\Lambda}'_y$$

$$\mathbf{\Sigma}_{\xi x} = \mathbf{\Sigma}'_{x\xi} = \mathbf{\Phi\Lambda}'_x$$

$$\mathbf{\Sigma}_{\xi\eta} = \mathbf{\Sigma}'_{\eta\xi} = \mathbf{\Phi\Gamma}'(\mathbf{I}-\mathbf{B})^{-1\prime}$$

$$\mathbf{\Sigma}_{\xi\xi} = \mathbf{\Phi}_\xi$$

where

$$\mathbf{\Psi}^* = (\mathbf{I}-\mathbf{B})^{-1} \mathbf{\Psi} (\mathbf{I}-\mathbf{B})^{-1\prime}$$

Completely standardized solution

Suppose now that $\hat{\mathbf{B}}_t$ and $\hat{\mathbf{\Phi}}_t$ denote the unstandardized estimators of $\mathbf{B}_t$ and $\mathbf{\Phi}_t$, respectively. The reproduced covariance matrix of the observed and latent variables may then be expressed as

$$\hat{\Upsilon} = \hat{\mathbf{\Lambda}}\hat{\mathbf{\Phi}}_t\hat{\mathbf{\Lambda}}'$$

The completely standardized covariance matrix of the observed and latent variables follows as

$$\hat{\Upsilon}^* = \mathbf{D}_{\hat{\Upsilon}}^{-1/2}\hat{\mathbf{\Lambda}}\hat{\mathbf{\Phi}}_t\hat{\mathbf{\Lambda}}'\mathbf{D}_{\hat{\Upsilon}}^{-1/2}$$

where $\mathbf{D}_{\hat{\Upsilon}}^{-1/2}$ denotes a $(p+m)\times(p+m)$ diagonal matrix with the reciprocals of the estimated standard deviations of the observed and latent variables on the diagonal.

The relationships between the unstandardized and the completely standardized estimators are given by

$$\hat{\mathbf{B}}_t^* = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_y^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_x^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\eta^{-1} & \hat{\mathbf{D}}_\eta^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \hat{\mathbf{B}}_t \begin{pmatrix} \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\eta & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\xi & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\eta & \hat{\mathbf{D}}_\xi & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

and

$$\hat{\mathbf{\Phi}}_t^* = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\xi^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_y^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_x^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\eta^{-1} \end{pmatrix} \hat{\mathbf{\Phi}}_t \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\xi^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_y^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_x^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\eta^{-1} \end{pmatrix}$$

where $\hat{\mathbf{D}}_y$, $\hat{\mathbf{D}}_x$, $\hat{\mathbf{D}}_\eta$, and $\hat{\mathbf{D}}_\xi$ denote diagonal matrices with the estimated standard deviations of the elements of $\mathbf{y}$, $\mathbf{x}$, $\boldsymbol{\eta}$, and $\boldsymbol{\xi}$ on the diagonal, respectively.

Suppose that the vector $\boldsymbol{\theta}$ consists of the q unknown elements of $\mathbf{B}$ and $\mathbf{\Phi}$. Let $\hat{\boldsymbol{\theta}}$ denotes the unstandardized estimator of $\boldsymbol{\theta}$ as such that asymptotically

$$\hat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta}, \mathbf{H}(\boldsymbol{\theta}))$$

By using the Delta method (Bishop, Fienberg, and Holland 1988), the asymptotic distribution of the completely standardized estimator, $\hat{\boldsymbol{\theta}}^*$, of $\boldsymbol{\theta}$ follows as

$$\hat{\boldsymbol{\theta}}^* \sim N\left(\boldsymbol{\theta}^*, \boldsymbol{\Delta}\mathbf{H}(\boldsymbol{\theta})\boldsymbol{\Delta}'\right)$$

where $\boldsymbol{\Delta} = \dfrac{\delta\boldsymbol{\theta}^*}{\delta\boldsymbol{\theta}'}$ and the elements of $\boldsymbol{\theta}^*$ are the unknown elements of $\mathbf{B}^*$ and $\boldsymbol{\Phi}^*$. Typical elements of $\mathbf{B}^*$ and $\boldsymbol{\Phi}^*$ are given by

$$\beta_{ij}^* = \upsilon_{ii}^{-1/2}\upsilon_{jj}^{1/2}\beta_{ij}$$

and

$$\phi_{ij}^* = \upsilon_{rr}^{-1/2}\upsilon_{ss}^{-1/2}\phi_{ij}$$

respectively where r and s are defined as

$$r = \begin{cases} i & \text{if } i \le p+m \\ i-p-m & \text{if } i > p+m \end{cases} \text{ and } s = \begin{cases} j & \text{if } j \le p+m \\ j-p-m & \text{if } j > p+m \end{cases}$$

respectively. Suppose that the sets $I_{\mathbf{B}}$ and $I_{\boldsymbol{\Phi}}$ are sets containing the row and column positions of the unknown elements of $\mathbf{B}$ and $\boldsymbol{\Phi}$, respectively, i.e.

$$I_{\mathbf{B}} = \left\{(i,j) : \beta_{ij} \in \boldsymbol{\theta}\right\} \text{ and } I_{\boldsymbol{\Phi}} = \left\{(i,j) : \phi_{ij} \in \boldsymbol{\theta}\right\}$$

respectively. The partial derivatives of the diagonal elements of Υ with respect to the elements of $\mathbf{B}$ and $\boldsymbol{\Phi}$ may then be expressed as

$$\frac{\delta\upsilon_{ii}}{\delta\beta_{kl}} = \begin{cases} 2\lambda_{ik}\upsilon_{il} & \text{if } (k,l) \in I_{\mathbf{B}} \\ 0 & \text{if } (k,l) \notin I_{\mathbf{B}} \end{cases}$$

and

$$\frac{\delta\upsilon_{ii}}{\delta\phi_{kl}} = \begin{cases} 2\left(1+\delta_{kl}\right)^{-1}\lambda_{ik}\lambda_{il} & \text{if } (k,l) \in I_{\boldsymbol{\Phi}} \\ 0 & \text{if } (k,l) \notin I_{\boldsymbol{\Phi}} \end{cases}$$

respectively where δ_{kl} denotes the Kronecker delta, i.e.

$$\delta_{kl} = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{if } k \ne l \end{cases}$$

Typical elements of $\boldsymbol{\Delta}$ may be expressed as

$$[\boldsymbol{\Delta}]_{ij,kl} = \begin{cases} \dfrac{\delta\beta_{ij}^*}{\delta\beta_{kl}} & \text{if } (i,j) \in I_{\mathbf{B}} \text{ and } (k,l) \in I_{\mathbf{B}} \\ \dfrac{\delta\beta_{ij}^*}{\delta\phi_{kl}} & \text{if } (i,j) \in I_{\mathbf{B}} \text{ and } (k,l) \in I_{\boldsymbol{\Phi}} \\ \dfrac{\delta\phi_{ij}^*}{\delta\beta_{kl}} & \text{if } (i,j) \in I_{\boldsymbol{\Phi}} \text{ and } (k,l) \in I_{\mathbf{B}} \\ \dfrac{\delta\phi_{ij}^*}{\delta\phi_{kl}} & \text{if } (i,j) \in I_{\boldsymbol{\Phi}} \text{ and } (k,l) \in I_{\boldsymbol{\Phi}} \end{cases}$$

where

$$\frac{\delta\beta_{ij}^*}{\delta\beta_{kl}} = \lambda_{jk}\upsilon_{jl}\upsilon_{ii}^{-1/2}\upsilon_{jj}^{-1/2}\beta_{ij} - \lambda_{ik}\upsilon_{il}\upsilon_{ii}^{-3/2}\upsilon_{jj}^{1/2}\beta_{ij} + \upsilon_{ii}^{-1/2}\upsilon_{jj}^{1/2}\frac{\delta\beta_{ij}}{\delta\beta_{kl}}$$

$$\frac{\delta\beta_{ij}^*}{\delta\phi_{kl}} = \left[\lambda_{jk}\lambda_{jl}\upsilon_{ii}^{-1/2}\upsilon_{jj}^{-1/2} - \lambda_{ik}\lambda_{il}\upsilon_{ii}^{-3/2}\upsilon_{jj}^{1/2}\right](1+\delta_{kl})^{-1}\beta_{ij}$$

$$\frac{\delta\phi_{ij}^*}{\delta\beta_{kl}} = -\left[\lambda_{ik}\upsilon_{il}\upsilon_{ii}^{-3/2}\upsilon_{jj}^{-1/2} + \lambda_{jk}\upsilon_{jl}\upsilon_{ii}^{-1/2}\upsilon_{jj}^{-3/2}\right]\phi_{ij}$$

$$\frac{\delta\phi_{ij}^*}{\delta\phi_{kl}} = -\left[\lambda_{ik}\lambda_{il}\upsilon_{ii}^{-3/2}\upsilon_{jj}^{-1/2} + \lambda_{jk}\lambda_{jl}\upsilon_{ii}^{-1/2}\upsilon_{jj}^{-3/2}\right]\left(1+\delta_{kl}\right)^{-1}\phi_{ij} + \upsilon_{ii}^{-1/2}\upsilon_{jj}^{-1/2}\frac{\delta\phi_{ij}}{\delta\phi_{kl}}$$

The standard error estimates of the completely standardized estimators of the elements of $\boldsymbol{\theta}$ are obtained as the positive square roots of the diagonal elements of the estimated asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}^*$ which is given by

$$\mathbf{H}^*\left(\hat{\boldsymbol{\theta}}^*\right) = \boldsymbol{\Delta}\left(\hat{\boldsymbol{\theta}}^*,\hat{\boldsymbol{\theta}}\right)\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right)\boldsymbol{\Delta}\left(\hat{\boldsymbol{\theta}}^*,\hat{\boldsymbol{\theta}}\right)'$$

These standard error estimates are numerically equivalent to those obtained by transforming correlation structures to covariance structures and fitting the transformed covariance structures correctly to the sample correlation matrices by using the theory and methods for covariance structures proposed by Shapiro and Browne (1990) which are implemented in Steiger (1995) and Browne and Mels (1996). Whenever a LISREL model without parameter equality constraints is fitted to a sample correlation matrix, the standard error estimates of the completely standardized solution are the correct standard error estimates which addresses the issue of incorrect standard error estimates for correlation matrices pointed out by Cudeck (1989).

Standardized solution

Let $\hat{\mathbf{B}}_t$ and $\hat{\boldsymbol{\Phi}}_t$ denote the unstandardized estimators of $\mathbf{B}_t$ and $\boldsymbol{\Phi}_t$, respectively and let $\hat{\mathbf{B}}_t^*$ and $\hat{\boldsymbol{\Phi}}_t^*$ denote the corresponding standardized estimators. The relationships between the unstandardized and the standardized estimators of $\mathbf{B}_t$ and $\boldsymbol{\Phi}_t$ are given by

$$\hat{\mathbf{B}}_t^* = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I}_{p_y} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{p_x} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\eta^{-1} & \hat{\mathbf{D}}_\eta^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \hat{\mathbf{B}}_t \begin{pmatrix} \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\eta & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\xi & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\eta & \hat{\mathbf{D}}_\xi & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

and

$$\hat{\mathbf{\Phi}}_t^* = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\xi^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{p_y} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{p_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\eta^{-1} \end{pmatrix} \hat{\mathbf{\Phi}}_t \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\xi^{-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{p_y} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{p_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \hat{\mathbf{D}}_\eta^{-1} \end{pmatrix}$$

respectively where $\hat{\mathbf{D}}_\eta$ and $\hat{\mathbf{D}}_\xi$ denote diagonal matrices with the estimated standard deviations of the elements of $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ on the diagonal, respectively.

Suppose that the vector $\boldsymbol{\theta}$ consists of the q unknown elements of $\mathbf{B}$ and $\mathbf{\Phi}$. Let $\hat{\boldsymbol{\theta}}$ denotes the unstandardized estimator of $\boldsymbol{\theta}$ as such that asymptotically

$$\hat{\boldsymbol{\theta}} \sim N\left(\boldsymbol{\theta}, \mathbf{H}\left(\boldsymbol{\theta}\right)\right)$$

By using the Delta method (Bishop, Fienberg, and Holland 1988), the asymptotic distribution of the standardized estimator, $\hat{\boldsymbol{\theta}}^*$, of $\boldsymbol{\theta}$ follows as

$$\hat{\boldsymbol{\theta}}^* \sim N\left(\boldsymbol{\theta}^*, \boldsymbol{\Delta}\mathbf{H}\left(\boldsymbol{\theta}\right)\boldsymbol{\Delta}'\right)$$

where $\boldsymbol{\Delta} = \dfrac{\delta\boldsymbol{\theta}^*}{\delta\boldsymbol{\theta}'}$ and the elements of $\boldsymbol{\theta}^*$ are the unknown elements of $\mathbf{B}^*$ and $\mathbf{\Phi}^*$. Typical elements of $\mathbf{B}^*$ and $\mathbf{\Phi}^*$ are given by

$$\beta_{ij}^* = \begin{cases} \upsilon_{jj}^{1/2}\beta_{ij} & \text{if } i \le p \\ \upsilon_{ii}^{-1/2}\upsilon_{jj}^{1/2}\beta_{ij} & \text{if } i > p \end{cases}$$

and

$$\phi_{ij}^* = \begin{cases} \upsilon_{rr}^{-1/2}\upsilon_{ss}^{-1/2}\phi_{ij} & \text{if } \left(i, j\right) \in I_1 \\ \phi_{ij} & \text{if } \left(i, j\right) \in I_2 \end{cases}$$

respectively where r and s are defined as

$$r = \begin{cases} i & \text{if } i \le p+m \\ i-p-m & \text{if } i > p+m \end{cases} \text{ and } s = \begin{cases} j & \text{if } j \le p+m \\ j-p-m & \text{if } j > p+m \end{cases}$$

respectively and where the sets I_1 and I_2 are defined as

$$I_1 = \{(i,j): i,j \le p+m \text{ or } i,j > 2p+m\} \text{ and } I_2 = \{(i,j): p+m < i,j \le 2p+m\}$$

respectively.

Suppose that the sets $I_{\mathbf{B}}$ and $I_{\mathbf{\Phi}}$ are sets containing the row and column positions of the unknown elements of $\mathbf{B}$ and $\mathbf{\Phi}$, respectively, i.e.

$$I_{\mathbf{B}} = \{(i,j): \beta_{ij} \in \boldsymbol{\theta}\} \text{ and } I_{\mathbf{\Phi}} = \{(i,j): \phi_{ij} \in \boldsymbol{\theta}\}$$

respectively. The partial derivatives of the diagonal elements of Υ with respect to the elements of $\mathbf{B}$ and $\mathbf{\Phi}$ may then be expressed as

$$\frac{\delta \upsilon_{ii}}{\delta \beta_{kl}} = \begin{cases} 2\lambda_{ik}\upsilon_{il} & \text{if } (k,l) \in I_{\mathbf{B}} \\ 0 & \text{if } (k,l) \notin I_{\mathbf{B}} \end{cases}$$

and

$$\frac{\delta \upsilon_{ii}}{\delta \phi_{kl}} = \begin{cases} 2(1+\delta_{kl})^{-1} \lambda_{ik}\lambda_{il} & \text{if } (k,l) \in I_{\mathbf{\Phi}} \\ 0 & \text{if } (k,l) \notin I_{\mathbf{\Phi}} \end{cases}$$

respectively where δ_{kl} denotes the Kronecker delta, i.e.

$$\delta_{kl} = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{if } k \ne l \end{cases}$$

Typical elements of $\mathbf{\Delta}$ may be expressed as

$$[\mathbf{\Delta}]_{ij,kl} = \begin{cases} \dfrac{\delta \beta_{ij}^*}{\delta \beta_{kl}} & \text{if } (i,j) \in I_{\mathbf{B}} \text{ and } (k,l) \in I_{\mathbf{B}} \\ \dfrac{\delta \beta_{ij}^*}{\delta \phi_{kl}} & \text{if } (i,j) \in I_{\mathbf{B}} \text{ and } (k,l) \in I_{\mathbf{\Phi}} \\ \dfrac{\delta \phi_{ij}^*}{\delta \beta_{kl}} & \text{if } (i,j) \in I_{\mathbf{\Phi}} \text{ and } (k,l) \in I_{\mathbf{B}} \\ \dfrac{\delta \phi_{ij}^*}{\delta \phi_{kl}} & \text{if } (i,j) \in I_{\mathbf{\Phi}} \text{ and } (k,l) \in I_{\mathbf{\Phi}} \end{cases}$$

where

$$\frac{\delta\beta_{ij}^*}{\delta\beta_{kl}} = \begin{cases} \lambda_{jk}\upsilon_{jl}\upsilon_{jj}^{-1/2}\beta_{ij} + \upsilon_{jj}^{1/2}\dfrac{\delta\beta_{ij}}{\delta\beta_{kl}} & \text{if } i \le p \\ \lambda_{jk}\upsilon_{jl}\upsilon_{ii}^{-1/2}\upsilon_{jj}^{-1/2}\beta_{ij} - \lambda_{ik}\upsilon_{il}\upsilon_{ii}^{-3/2}\upsilon_{jj}^{1/2}\beta_{ij} + \upsilon_{ii}^{-1/2}\upsilon_{jj}^{1/2}\dfrac{\delta\beta_{ij}}{\delta\beta_{kl}} & \text{if } i > p \end{cases}$$

$$\frac{\delta\beta_{ij}^*}{\delta\phi_{kl}} = \begin{cases} (1+\delta_{kl})^{-1}\lambda_{jk}\lambda_{jl}\upsilon_{jj}^{-1/2}\beta_{ij} & \text{if } i \le p \\ \left[\lambda_{jk}\lambda_{jl}\upsilon_{ii}^{-1/2}\upsilon_{jj}^{-1/2} - \lambda_{ik}\lambda_{il}\upsilon_{ii}^{-3/2}\upsilon_{jj}^{1/2}\right](1+\delta_{kl})^{-1}\beta_{ij} & \text{if } i > p \end{cases}$$

$$\frac{\delta\phi_{ij}^*}{\delta\beta_{kl}} = \begin{cases} -\left[\lambda_{ik}\upsilon_{il}\upsilon_{ii}^{-3/2}\upsilon_{jj}^{-1/2} + \lambda_{jk}\upsilon_{jl}\upsilon_{ii}^{-1/2}\upsilon_{jj}^{-3/2}\right]\phi_{ij} & \text{if } (i,j) \in I_1 \\ 0 & \text{if } (i,j) \in I_2 \end{cases}$$

$$\frac{\delta\phi_{ij}^*}{\delta\phi_{kl}} = \begin{cases} -\left[\lambda_{ik}\lambda_{il}\upsilon_{ii}^{-3/2}\upsilon_{jj}^{-1/2} + \lambda_{jk}\lambda_{jl}\upsilon_{ii}^{-1/2}\upsilon_{jj}^{-3/2}\right](1+\delta_{kl})^{-1}\phi_{ij} + \upsilon_{ii}^{-1/2}\upsilon_{jj}^{-1/2}\dfrac{\delta\phi_{ij}}{\delta\phi_{kl}} & \text{if } (i,j) \in I_1 \\ \dfrac{\delta\phi_{ij}}{\delta\phi_{kl}} & \text{if } (i,j) \in I_2 \end{cases}$$

The standard error estimates of the completely standardized estimators of the elements of $\boldsymbol{\theta}$ are obtained as the positive square roots of the diagonal elements of the estimated asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}^*$ which is given by

$$\mathbf{H}^*\left(\hat{\boldsymbol{\theta}}^*\right) = \boldsymbol{\Delta}\left(\hat{\boldsymbol{\theta}}^*, \hat{\boldsymbol{\theta}}\right)\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right)\boldsymbol{\Delta}\left(\hat{\boldsymbol{\theta}}^*, \hat{\boldsymbol{\theta}}\right)'$$

2.2 Confidence interval estimates

The extended LISREL model

In the extended LISREL model (Jöreskog and Sörbom 1999), the three sets of relationships between the observed and latent variables are given by

$$\mathbf{y} = \boldsymbol{\tau}_y + \boldsymbol{\Lambda}_y\boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

$$\mathbf{x} = \boldsymbol{\tau}_x + \boldsymbol{\Lambda}_x\boldsymbol{\xi} + \boldsymbol{\delta}$$

$$\boldsymbol{\eta} = \boldsymbol{\alpha} + \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}$$

where the elements of $\mathbf{y}$ denote p_y observed indicators of the m_η endogenous latent variables $\boldsymbol{\eta}$, the elements of $\mathbf{x}$ denote p_x observed indicators of the m_ξ exogenous latent variables $\boldsymbol{\xi}$, the elements of $\boldsymbol{\varepsilon}$ denote p_y measurement errors, the elements of $\boldsymbol{\delta}$ denote p_x measurement errors, the elements of $\boldsymbol{\zeta}$ denote m_η error variables, the elements of $\boldsymbol{\tau}_y$ are p_y intercepts, the elements of $\boldsymbol{\tau}_x$ are p_x intercepts, the elements of $\boldsymbol{\alpha}$ are m_η intercepts, the elements of $\boldsymbol{\Lambda}_y$ are $p_y \times m_\eta$ measurement weights, the elements of $\boldsymbol{\Lambda}_x$ are $p_x \times m_\xi$ measurement weights, the elements of $\mathbf{B}$ are $m_\eta \times m_\eta$ regression weights, and the elements of $\boldsymbol{\Gamma}$ are $m_\eta \times m_\xi$ regression weights. We assume that $\boldsymbol{\zeta}$ is uncorrelated with $\boldsymbol{\xi}$, $\boldsymbol{\varepsilon}$ is uncorrelated with $\boldsymbol{\eta}$, and that $\boldsymbol{\delta}$ is uncorrelated with $\boldsymbol{\xi}$. We also assume that the means of $\boldsymbol{\varepsilon}$, and $\boldsymbol{\delta}$, and $\boldsymbol{\zeta}$ are zero, but it is not assumed that the means of $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are zero. If the mean of $\boldsymbol{\xi}$ is denoted by $\boldsymbol{\kappa}$, the mean of $\boldsymbol{\eta}$ follows as

$$\boldsymbol{\mu}_\eta = (\mathbf{I}-\mathbf{B})^{-1}(\boldsymbol{\alpha}+\boldsymbol{\Gamma}\boldsymbol{\kappa})$$

The mean vectors of the observed indicators are given by

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_y \\ \boldsymbol{\mu}_x \end{bmatrix} = \begin{bmatrix} \boldsymbol{\tau}_y + \boldsymbol{\Lambda}_y(\mathbf{I}-\mathbf{B})^{-1}(\boldsymbol{\alpha}+\boldsymbol{\Gamma}\boldsymbol{\kappa}) \\ \boldsymbol{\tau}_x + \boldsymbol{\Lambda}_x\boldsymbol{\kappa} \end{bmatrix}$$

The covariance matrix of the observed indicators follows as

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Lambda}_y\left((\mathbf{I}-\mathbf{B})^{-1}\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}'(\mathbf{I}-\mathbf{B})^{-1\prime}+\boldsymbol{\Psi}^*\right)\boldsymbol{\Lambda}_y'+\boldsymbol{\Theta}_\varepsilon & \boldsymbol{\Lambda}_y(\mathbf{I}-\mathbf{B})^{-1}\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Lambda}_x'+\boldsymbol{\Theta}_{\varepsilon\delta} \\ \boldsymbol{\Lambda}_x\boldsymbol{\Phi}\boldsymbol{\Gamma}'\left(\mathbf{I}-\mathbf{B}\right)^{-1\prime}\boldsymbol{\Lambda}_y'+\boldsymbol{\Theta}_{\delta\varepsilon} & \boldsymbol{\Lambda}_x\boldsymbol{\Phi}\boldsymbol{\Lambda}_x'+\boldsymbol{\Theta}_\delta \end{bmatrix}$$

where

$$\boldsymbol{\Psi}^* = (\mathbf{I}-\mathbf{B})^{-1}\boldsymbol{\Psi}(\mathbf{I}-\mathbf{B})^{-1\prime}$$

and $\boldsymbol{\Phi}$ denotes the covariance matrix of $\boldsymbol{\xi}$, $\boldsymbol{\Theta}_\varepsilon$ denotes the covariance matrix of $\boldsymbol{\varepsilon}$, $\boldsymbol{\Theta}_\delta$ denotes the covariance matrix of $\boldsymbol{\delta}$, $\boldsymbol{\Theta}_{\delta\varepsilon} = \boldsymbol{\Theta}_{\varepsilon\delta}'$ denotes the covariance matrix between $\boldsymbol{\delta}$ and $\boldsymbol{\varepsilon}$, and $\boldsymbol{\Psi}$ denotes the covariance matrix of $\boldsymbol{\zeta}$.

The q unknown parameters $\boldsymbol{\theta}$ of the extended LISREL model consist of the unknown elements of $\boldsymbol{\Lambda}_y$, $\boldsymbol{\Lambda}_x$, $\mathbf{B}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}$, $\boldsymbol{\Theta}_\varepsilon$, $\boldsymbol{\Theta}_\delta$, $\boldsymbol{\Theta}_{\delta\varepsilon} = \boldsymbol{\Theta}_{\varepsilon\delta}'$, $\boldsymbol{\tau}_y$, $\boldsymbol{\tau}_x$, $\boldsymbol{\alpha}$, and $\boldsymbol{\kappa}$.

Unstandardized solution

Suppose that $\hat{\boldsymbol{\theta}}$ denote the unstandardized estimators of the parameters $\boldsymbol{\theta}$ of the extended LISREL model as such that asymptotically

$$\hat{\boldsymbol{\theta}} \sim N\left(\boldsymbol{\theta}, \mathbf{H}\left(\boldsymbol{\theta}\right)\right)$$

The elements of $\boldsymbol{\theta}$ consists of intercepts, measurement weights, regression weights, variances, and covariances. The intercepts, measurement weights, regression weights, and covariances are unbounded parameters. As a result, the $100(1-\alpha)\%$ approximate confidence interval estimates of these parameters (Browne 1982) are given by

$$\left(\hat{\theta}_i - z_{\alpha/2}s\left(\hat{\theta}_i\right); \hat{\theta}_i + z_{\alpha/2}s\left(\hat{\theta}_i\right)\right)$$

where $\hat{\theta}_i$ denotes the estimate of θ_i, $z_{\alpha/2}$ denotes the $100(1-\alpha/2)\%$ critical value of the standard normal distribution and $s\left(\hat{\theta}_i\right) = \sqrt{\left[\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right)\right]_{ii}}$ denotes the estimate of the standard error of the estimator of θ_i. If θ_i denotes a variance, then θ_i is a bounded parameter as such that $0 < \theta_i < \infty$. In this case, a logarithmic transformation is used as such that the $100(1-\alpha)\%$ approximate confidence interval estimate of θ_i (Browne 1982) follows as

$$\left(\hat{\theta}_i \exp\left(-z_{\alpha/2} s\left(\hat{\theta}_i\right)\Big/\hat{\theta}_i\right); \hat{\theta}_i \exp\left(z_{\alpha/2} s\left(\hat{\theta}_i\right)\Big/\hat{\theta}_i\right)\right)$$

Standardized solution

Let $\hat{\boldsymbol{\theta}}^*$ denote the standardized estimators of the parameters $\boldsymbol{\theta}$ of the extended LISREL model. By using the Delta method (Bishop, Feinberg, and Holland 1988), it follows that asymptotically

$$\hat{\boldsymbol{\theta}}^* \sim N\left(\boldsymbol{\theta}^*, \boldsymbol{\Delta}\mathbf{H}\left(\boldsymbol{\theta}\right)\boldsymbol{\Delta}'\right)$$

where $\boldsymbol{\Delta}$ denotes is Jacobian matrix of $\boldsymbol{\theta}^*$ with respect to $\boldsymbol{\theta}$. The elements of $\boldsymbol{\theta}^*$ consists of intercepts of the observed indicators, measurement weights, standardized regression weights, variances, standardized variances, covariances, and correlations. The intercepts of the observed variables, measurement weights, and covariances are unbounded parameters. As a result, the $100(1-\alpha)\%$ approximate confidence interval estimates of these parameters (Browne 1982) are given by

$$\left(\hat{\theta}_i^* - z_{\alpha/2} s\left(\hat{\theta}_i^*\right); \hat{\theta}_i^* + z_{\alpha/2} s\left(\hat{\theta}_i^*\right)\right)$$

where $\hat{\theta}_i^*$ denotes the estimate of θ_i^*, $z_{\alpha/2}$ denotes the $100(1-\alpha/2)\%$ critical value of the standard normal distribution and $s\left(\hat{\theta}_i^*\right) = \sqrt{\left[\boldsymbol{\Delta}\left(\hat{\boldsymbol{\theta}}^*, \hat{\boldsymbol{\theta}}\right)\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right)\boldsymbol{\Delta}\left(\hat{\boldsymbol{\theta}}^*, \hat{\boldsymbol{\theta}}\right)'\right]_{ii}}$ denotes the estimate of the standard error of the estimator of θ_i^*. If θ_i^* denotes a variance, then θ_i^* is a bounded parameter as such that $0 < \theta_i^* < \infty$. In this case, a logarithmic transformation is used as such that the $100(1-\alpha)\%$ approximate confidence interval estimate of θ_i^* (Browne 1982) follows as

$$\left(\hat{\theta}_i \exp\left(-z_{\alpha/2} s\left(\hat{\theta}_i\right)\Big/\hat{\theta}_i\right); \hat{\theta}_i \exp\left(z_{\alpha/2} s\left(\hat{\theta}_i\right)\Big/\hat{\theta}_i\right)\right)$$

If θ_i^* denotes a standardized variance, then θ_i^* is a bounded parameter as such that $0 < \theta_i^* < 1$. The $100(1-\alpha)\%$ approximate confidence interval estimate of θ_i^* (Browne 1982) may be expressed as

$$\left(1\Bigg/\left(1 + \frac{1}{\left(\hat{\theta}_i^* - 1\right)}\exp\left(\frac{z_{\alpha/2} s\left(\hat{\theta}_i^*\right)}{\hat{\theta}_i^*\left(1-\hat{\theta}_i^*\right)}\right)\right); 1\Bigg/\left(1 + \frac{1}{\left(\hat{\theta}_i^* - 1\right)}\exp\left(\frac{-z_{\alpha/2} s\left(\hat{\theta}_i^*\right)}{\hat{\theta}_i^*\left(1-\hat{\theta}_i^*\right)}\right)\right)\right)$$

When θ_i^* is a standardized regression weight or a correlation, θ_i^* is bounded as such that $-1 < \theta_i^* < 1$. In this case, the Fisher z-transformation is used as such that the $100(1-\alpha)\%$ approximate confidence interval estimate of θ_i^* (Browne 1982) is given by

$$\left(\frac{\left(\frac{1+\hat{\theta}_i^*}{1-\hat{\theta}_i^*}\right)\exp\left(\frac{-2z_{\alpha/2} s\left(\hat{\theta}_i^*\right)}{\left(1-\hat{\theta}_i^{*2}\right)}\right) - 1}{\left(\frac{1+\hat{\theta}_i^*}{1-\hat{\theta}_i^*}\right)\exp\left(\frac{-2z_{\alpha/2} s\left(\hat{\theta}_i^*\right)}{\left(1-\hat{\theta}_i^{*2}\right)}\right) + 1}; \frac{\left(\frac{1+\hat{\theta}_i^*}{1-\hat{\theta}_i^*}\right)\exp\left(\frac{2z_{\alpha/2} s\left(\hat{\theta}_i^*\right)}{\left(1-\hat{\theta}_i^{*2}\right)}\right) - 1}{\left(\frac{1+\hat{\theta}_i^*}{1-\hat{\theta}_i^*}\right)\exp\left(\frac{2z_{\alpha/2} s\left(\hat{\theta}_i^*\right)}{\left(1-\hat{\theta}_i^{*2}\right)}\right) + 1}\right)$$

Completely standardized solution

If $\hat{\boldsymbol{\theta}}^*$ denote the completely standardized estimators of the parameters $\boldsymbol{\theta}$ of the extended LISREL model, the asymptotic distribution of $\hat{\boldsymbol{\theta}}^*$ is obtained by means of the Delta method (Bishop, Feinberg, and Holland 1988) as

$$\hat{\boldsymbol{\theta}}^* \sim N\left(\boldsymbol{\theta}^*, \boldsymbol{\Delta}\mathbf{H}\left(\boldsymbol{\theta}\right)\boldsymbol{\Delta}'\right)$$

where $\boldsymbol{\Delta}$ denotes is Jacobian matrix of $\boldsymbol{\theta}^*$ with respect to $\boldsymbol{\theta}$. The elements of $\boldsymbol{\theta}^*$ consists of standardized measurement weights, standardized regression weights, standardized variances, and correlations.

If θ_i^* denotes a standardized variance, then θ_i^* is a bounded parameter as such that $0 < \theta_i^* < 1$. The $100(1-\alpha)\%$ approximate confidence interval estimate of θ_i^* (Browne 1982) may be expressed as

$$\left(1\middle/\left(1+\frac{1}{\left(\hat{\theta}_i^*-1\right)}\exp\left(\frac{z_{\alpha/2}s\left(\hat{\theta}_i^*\right)}{\hat{\theta}_i^*\left(1-\hat{\theta}_i^*\right)}\right)\right); 1\middle/\left(1+\frac{1}{\left(\hat{\theta}_i^*-1\right)}\exp\left(\frac{-z_{\alpha/2}s\left(\hat{\theta}_i^*\right)}{\hat{\theta}_i^*\left(1-\hat{\theta}_i^*\right)}\right)\right)\right)$$

where $\hat{\theta}_i^*$ denotes the estimate of θ_i^*, $z_{\alpha/2}$ denotes the $100(1-\alpha/2)\%$ critical value of the standard normal distribution and $s\left(\hat{\theta}_i^*\right) = \sqrt{\left[\boldsymbol{\Delta}\left(\hat{\boldsymbol{\theta}}^*, \hat{\boldsymbol{\theta}}\right)\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right)\boldsymbol{\Delta}\left(\hat{\boldsymbol{\theta}}^*, \hat{\boldsymbol{\theta}}\right)'\right]_{ii}}$ denotes the estimate of the standard error of the estimator of θ_i^*.

When θ_i^* is a standardized measurement weight, a standardized regression weight, or a correlation, θ_i^* is bounded as such that $-1 < \theta_i^* < 1$. In this case, the Fisher *z*-transformation is used a such that the $100(1-\alpha)\%$ approximate confidence interval estimate of θ_i^* (Browne 1982) is given by

$$\left(\frac{\left(\frac{1+\hat{\theta}_i^*}{1-\hat{\theta}_i^*}\right)\exp\left(\frac{-2z_{\alpha/2}s\left(\hat{\theta}_i^*\right)}{\left(1-\hat{\theta}_i^{*2}\right)}\right)-1}{\left(\frac{1+\hat{\theta}_i^*}{1-\hat{\theta}_i^*}\right)\exp\left(\frac{-2z_{\alpha/2}s\left(\hat{\theta}_i^*\right)}{\left(1-\hat{\theta}_i^{*2}\right)}\right)+1}; \frac{\left(\frac{1+\hat{\theta}_i^*}{1-\hat{\theta}_i^*}\right)\exp\left(\frac{2z_{\alpha/2}s\left(\hat{\theta}_i^*\right)}{\left(1-\hat{\theta}_i^{*2}\right)}\right)-1}{\left(\frac{1+\hat{\theta}_i^*}{1-\hat{\theta}_i^*}\right)\exp\left(\frac{2z_{\alpha/2}s\left(\hat{\theta}_i^*\right)}{\left(1-\hat{\theta}_i^{*2}\right)}\right)+1}\right)$$

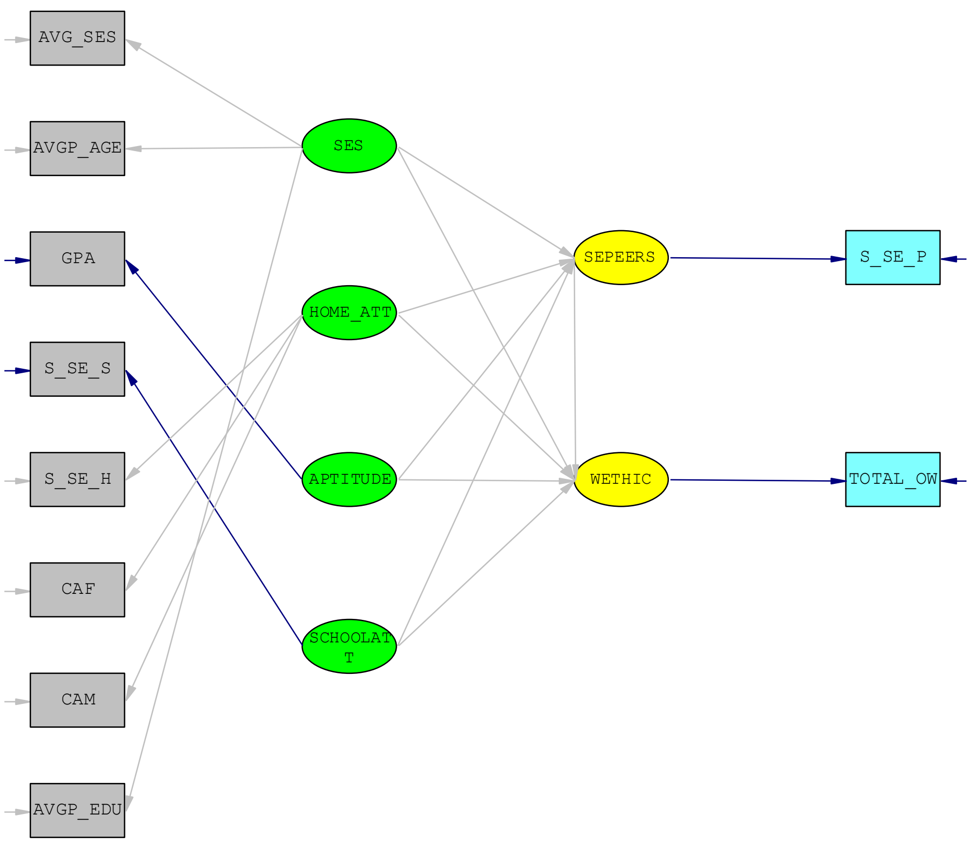
2.3 Structural equation model for work ethic

The data are the scores of 194 freshman students at a high school in Bainbridge, Georgia on ten observed scores (average socio-economic index, average age of parents, grade point average, self-esteem at school, self-esteem at home, self-esteem around peers, attitude towards father, attitude towards mother, work ethic, and average education level of parents). The first couple of observations of the corresponding data file, **STUDENTS.LSF**, are shown below.

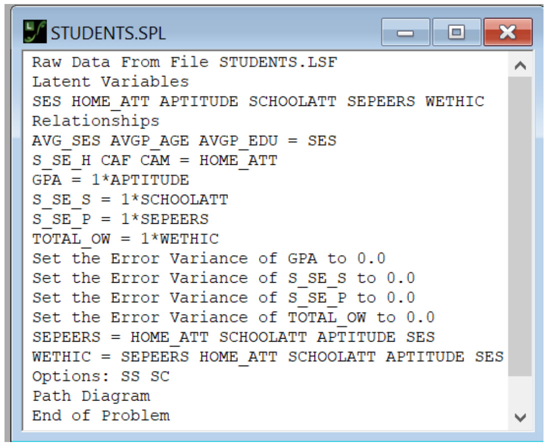
STUDENTS.lsf

	AVG_SES	AVGP_AGE	GPA	S_SE_S	S_SE_H	S_SE_P	CAF	CAM	TOTAL_OW	AVGP_EDU
1	35.55	40.50	85.00	20.00	25.00	22.00	15.33	14.00	12.69	3.00
2	34.62	41.50	75.00	26.00	21.00	29.00	12.00	2.00	12.93	2.00
3	30.13	50.50	85.00	24.00	25.00	23.00	19.33	21.33	12.43	1.50
4	35.37	37.00	95.00	23.00	26.00	26.00	5.33	0.00	19.82	7.50
5	22.40	43.50	75.00	20.00	24.00	23.00	24.67	34.00	14.63	1.50
6	15.96	29.50	85.00	24.00	26.00	25.00	7.33	0.00	12.89	1.50
7	22.75	49.00	95.00	24.00	25.00	31.00	56.67	4.00	18.29	3.00
8	37.31	47.00	75.00	24.00	24.50	17.00	7.33	50.00	10.56	2.00
9	22.70	37.00	85.00	27.00	30.00	24.00	50.00	34.00	11.13	2.00
10	22.54	42.50	85.00	26.00	27.00	27.00	6.67	4.67	13.02	2.00
11	23.50	30.50	75.00	24.00	29.00	26.00	50.00	50.00	11.74	1.00
12	23.28	29.00	85.00	27.00	27.50	22.00	0.00	4.17	16.82	3.00
13	23.43	35.50	85.00	25.00	21.00	26.00	6.00	6.67	13.55	2.00
14	30.91	32.50	85.00	25.00	25.00	24.00	36.67	42.00	11.99	4.00
15	49.20	45.00	85.00	34.00	25.00	32.00	14.67	7.33	14.95	2.00

The theoretical model is a structural equation model that suggests that socio-economic status, home environment, grade point average, and self-esteem at school are antecedents of self-esteem around peers and work ethic. A path diagram for this model is depicted in the image below.



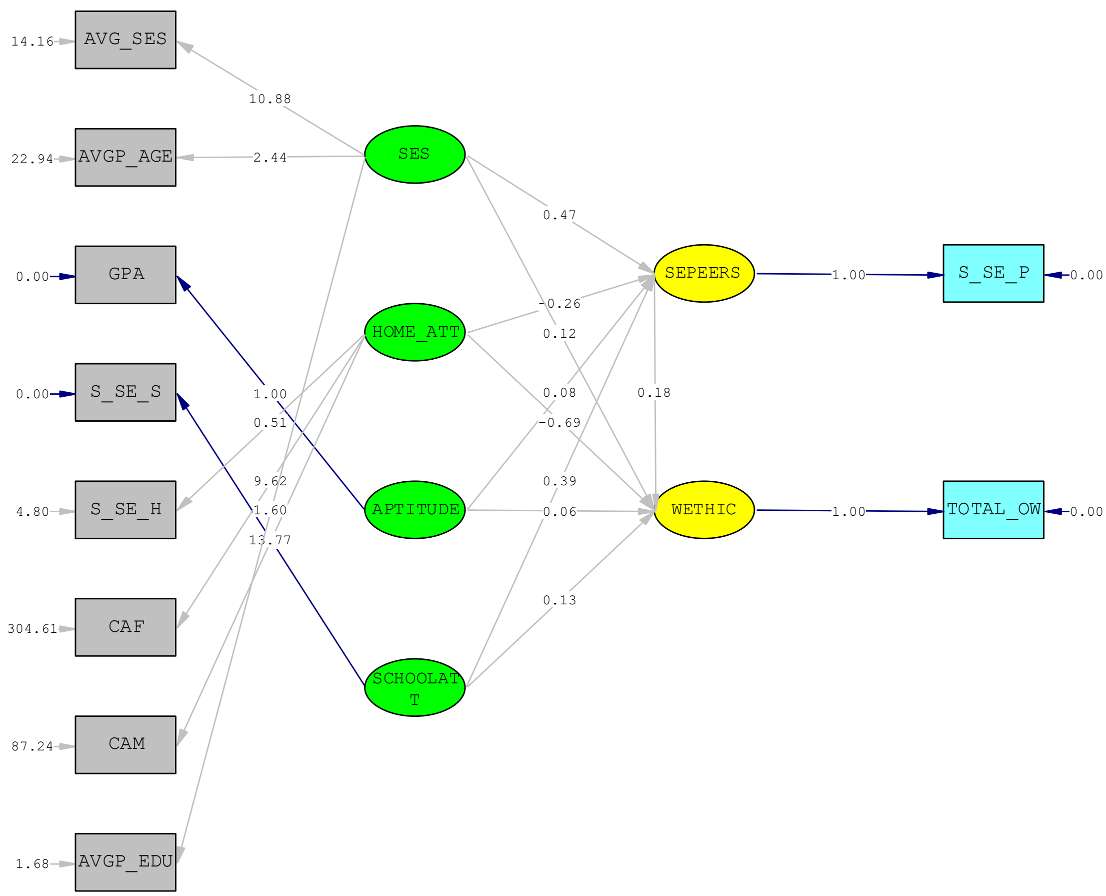
The SIMPLIS syntax file for the theoretical model above is shown in the image below.



STUDENTS.SPL

```
Raw Data From File STUDENTS.LSF
Latent Variables
SES HOME_ATT APTITUDE SCHOOLATT SEPEERS WETHIC
Relationships
AVG_SES AVGP_AGE AVGP_EDU = SES
S_SE_H CAF CAM = HOME_ATT
GPA = 1*APTITUDE
S_SE_S = 1*SCHOOLATT
S_SE_P = 1*SEPEERS
TOTAL_OW = 1*WETHIC
Set the Error Variance of GPA to 0.0
Set the Error Variance of S_SE_S to 0.0
Set the Error Variance of S_SE_P to 0.0
Set the Error Variance of TOTAL_OW to 0.0
SEPEERS = HOME_ATT SCHOOLATT APTITUDE SES
WETHIC = SEPEERS HOME_ATT SCHOOLATT APTITUDE SES
Options: SS SC
Path Diagram
End of Problem
```

- Line 1 specifies data file to be used.
- Lines 2 to 3 specify the labels for the latent variables of the model.
- Lines 4 to 16 specify the model to be fitted to the data.
- Line 17 requests the standardized and completely standardized solutions.
- Line 18 requests a path diagram of the model.
- Line 19 indicates that no more SIMPLIS commands are to be processed.



Chi-Square=28.71, df=24, P-value=0.23138, RMSEA=0.032

If the SPL file is opened in LISREL and the **Run LISREL** icon is clicked, the path diagram shown above is obtained.

The corresponding output file, **STUDENTS.OUT**, is opened in a separate window. The confidence interval estimates of the structural equation model parameters for the unstandardized solution, the standardized solution, and the completely standardized solution, which are listed in this file, are shown in the images below.

STUDENTS.OUT

```
        Structural Equations


 SEPEERS=0.475*SES - 0.257*HOME_ATT + 0.0801*APTITUDE + 0.391*SCHOOLATT, Errorvar.= 7.557 , R²=0.214
90% CILL(0.106;      (-0.668;        (0.0272;          (0.255;                      (6.385;
90% CIUL 0.844)       0.154)          0.133)            0.527)                       8.945)
Standerr(0.224)      (0.250)         (0.0322)          (0.0827)                     (0.774)
Z-value  2.116       -1.030           2.488             4.728                        9.758
P-value  0.034        0.303           0.013             0.000                        0.000

  WETHIC=0.177*SEPEERS + 0.120*SES - 0.692*HOME_ATT + 0.0610*APTITUDE + 0.125*SCHOOLATT, Errorvar.= 3.240 , R²=0.355
90% CILL(0.0970;         (-0.130;     (-0.997;         (0.0240;          (0.0289;                     (2.698;
90% CIUL 0.258)           0.371)       -0.387)          0.0980)           0.222)                       3.892)
Standerr(0.0489)         (0.152)      (0.185)          (0.0225)          (0.0587)                     (0.361)
Z-value  3.631            0.792       -3.734            2.709             2.138                        8.980
P-value  0.000            0.428        0.000            0.007             0.033                        0.000
```

STUDENTS.OUT

```
        Structural Equations


 SEPEERS=0.153*SES - 0.0830*HOME_ATT + 0.180*APTITUDE + 0.307*SCHOOLATT, Errorvar.= 0.786
90% CILL(0.0341;     (-0.215;          (0.0603;          (0.202;                      (0.686;
90% CIUL 0.268)       0.0499)           0.295)            0.406)                       0.861)
Standerr(0.0714)     (0.0804)          (0.0717)          (0.0623)                     (0.0532)
Z-value  2.145       -1.033             2.514             4.932                        14.789
P-value  0.032        0.302             0.012             0.000                        0.000

  WETHIC=0.245*SEPEERS + 0.0537*SES - 0.309*HOME_ATT + 0.190*APTITUDE + 0.136*SCHOOLATT, Errorvar.= 0.645
90% CILL(0.133;          (-0.058;      (-0.440;          (0.0734;          (0.0308;                     (0.537;
90% CIUL 0.351)           0.164)        -0.178)           0.301)            0.239)                       0.740)
Standerr(0.0665)         (0.0677)      (0.0797)          (0.0695)          (0.0635)                     (0.0625)
Z-value  3.689            0.793        -3.871             2.731             2.147                        10.311
P-value  0.000            0.428         0.000             0.006             0.032                        0.000
```

STUDENTS.OUT

```
        Structural Equations


 SEPEERS=0.153*SES - 0.0830*HOME_ATT + 0.180*APTITUDE + 0.307*SCHOOLATT, Errorvar.= 0.786
90% CILL(0.0341;     (-0.215;          (0.0603;          (0.202;                      (0.686;
90% CIUL 0.268)       0.0499)           0.295)            0.406)                       0.861)
Standerr(0.0714)     (0.0804)          (0.0717)          (0.0623)                     (0.0532)
Z-value  2.145       -1.033             2.514             4.932                        14.789
P-value  0.032        0.302             0.012             0.000                        0.000

  WETHIC=0.245*SEPEERS + 0.0537*SES - 0.309*HOME_ATT + 0.190*APTITUDE + 0.136*SCHOOLATT, Errorvar.= 0.645
90% CILL(0.133;          (-0.058;      (-0.440;          (0.0734;          (0.0308;                     (0.537;
90% CIUL 0.351)           0.164)        -0.178)           0.301)            0.239)                       0.740)
Standerr(0.0665)         (0.0677)      (0.0797)          (0.0695)          (0.0635)                     (0.0625)
Z-value  3.689            0.793        -3.871             2.731             2.147                        10.311
P-value  0.000            0.428         0.000             0.006             0.032                        0.000
```

Note that the confidence interval estimates for the standardized and completely standardized solutions are identical since the variances of all the latent variables of the model are scaled to be equal to unity for both solutions. However, the confidence interval estimates of the parameters of the measurement model of the standardized and completely standardized solutions differ since the variances of the observed variables are not scaled for the standardized solution but are scaled to be equal to unity for the completely standardized solution.

3. Variance constraints for endogenous latent variables

3.1 Estimation

The LISREL model for observed and latent variables

The LISREL model (Jöreskog 1973, 1977) for population covariance matrices may be expressed as

$$\mathbf{y} = \mathbf{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

$$\mathbf{x} = \mathbf{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}$$

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \mathbf{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}$$

where $\mathbf{y}$ and $\mathbf{x}$ denote p_η and p_ξ indicators of the m_η endogenous latent variables, $\boldsymbol{\eta}$, and the m_ξ exogenous latent variables, $\boldsymbol{\xi}$, respectively, $\mathbf{\Lambda}_y$ and $\mathbf{\Lambda}_x$ are $p_y \times m_\eta$ and $p_x \times m_\xi$ matrices of factor loadings, respectively, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$ denote p_η and p_ξ measurement errors, respectively, $\mathbf{B}$ and $\mathbf{\Gamma}$ are $m_\eta \times m_\eta$ and $m_\eta \times m_\xi$ matrices of regression weights, respectively, and the elements of $\boldsymbol{\zeta}$ denote m_η error variables.

The $t \times 1$ vector, $\mathbf{z}$, consisting of all the variables of the LISREL model follows as

$$\mathbf{z} = \begin{pmatrix} \mathbf{y} \\ \mathbf{x} \\ \boldsymbol{\eta} \\ \boldsymbol{\xi} \\ \boldsymbol{\varepsilon} \\ \boldsymbol{\delta} \\ \boldsymbol{\zeta} \end{pmatrix}$$

The model for the relationships between all the variables of the LISREL model may then be expressed as

$$\mathbf{z} = \mathbf{B}_t \mathbf{z} + \mathbf{z}_e$$

where

$$\mathbf{B}_t = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{\Lambda}_y & \mathbf{0} & \mathbf{I}_{p_y} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{\Lambda}_x & \mathbf{0} & \mathbf{I}_{p_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B} & \mathbf{\Gamma} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{m_\eta} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

where $\mathbf{I}_n$ denotes the $n \times n$ identity matrix and

$$\mathbf{z}_e = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \boldsymbol{\xi} \\ \boldsymbol{\varepsilon} \\ \boldsymbol{\delta} \\ \boldsymbol{\zeta} \end{pmatrix}$$

The covariance matrix, $\boldsymbol{\Phi}_t$, of $\mathbf{z}_e$ follows as

$$\boldsymbol{\Phi}_t = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Phi} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Theta}_{\varepsilon} & \boldsymbol{\Theta}_{\varepsilon\delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Theta}_{\delta\varepsilon} & \boldsymbol{\Theta}_{\delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Psi} \end{pmatrix}$$

where $\boldsymbol{\Phi}$, $\boldsymbol{\Theta}_{\varepsilon}$, $\boldsymbol{\Theta}_{\delta}$, and $\boldsymbol{\Psi}$ denote the covariance matrices of $\boldsymbol{\xi}$, $\boldsymbol{\varepsilon}$, $\boldsymbol{\delta}$, and $\boldsymbol{\zeta}$, respectively and $\boldsymbol{\Theta}_{\varepsilon\delta} = \boldsymbol{\Theta}'_{\delta\varepsilon}$ denotes the covariance matrix between $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$. The $t \times t$ covariance matrix of $\mathbf{z}$, Υ_t, may then be expressed as

$$\Upsilon_t = \left(\mathbf{I}_t - \mathbf{B}_t\right)^{-1} \boldsymbol{\Phi}_t \left(\mathbf{I}_t - \mathbf{B}_t\right)^{-1\prime} = \boldsymbol{\Lambda}_t \boldsymbol{\Phi}_t \boldsymbol{\Lambda}'_t$$

where $\boldsymbol{\Lambda}_t = \left(\mathbf{I}_t - \mathbf{B}_t\right)^{-1}$. The $\left(p+m\right)\times\left(p+m\right)$ covariance matrix, Υ, of the $p = p_y + p_x$ observed variables and $m = m_\eta + m_\xi$ latent variables, follows as

$$\Upsilon = \left[\mathbf{I}_{p+m} \quad \mathbf{0}\right]\left(\mathbf{I}_t - \mathbf{B}_t\right)^{-1} \boldsymbol{\Phi}_t \left(\mathbf{I}_t - \mathbf{B}_t\right)^{-1\prime} \left[\mathbf{I}_{p+m} \quad \mathbf{0}\right]' = \boldsymbol{\Lambda} \boldsymbol{\Phi}_t \boldsymbol{\Lambda}'$$

where $\boldsymbol{\Lambda}$ denotes the $\left(p+m\right)\times t$ matrix consisting of the first $p+m$ rows of $\boldsymbol{\Lambda}_t$. The $p \times p$ covariance matrix, $\boldsymbol{\Sigma}$, of the p observed variables may be expressed as

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}_p \boldsymbol{\Phi}_t \boldsymbol{\Lambda}'_p$$

where $\boldsymbol{\Lambda}_p$ denotes the $p \times t$ matrix consisting of the first p rows of $\boldsymbol{\Lambda}_t$.

In terms of the parameter matrices of the general LISREL model, we obtain that

$$\Upsilon_t = \begin{pmatrix} \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_{y\eta} & \boldsymbol{\Sigma}_{y\xi} \\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx} & \boldsymbol{\Sigma}_{x\eta} & \boldsymbol{\Sigma}_{x\xi} \\ \boldsymbol{\Sigma}_{\eta y} & \boldsymbol{\Sigma}_{\eta x} & \boldsymbol{\Sigma}_{\eta\eta} & \boldsymbol{\Sigma}_{\eta\xi} \\ \boldsymbol{\Sigma}_{\xi y} & \boldsymbol{\Sigma}_{\xi x} & \boldsymbol{\Sigma}_{\xi\eta} & \boldsymbol{\Sigma}_{\xi\xi} \end{pmatrix}$$

where

$$\mathbf{\Sigma}_{yy} = \mathbf{\Lambda}_y \left((\mathbf{I}-\mathbf{B})^{-1}\mathbf{\Gamma\Phi\Gamma}'(\mathbf{I}-\mathbf{B})^{-1'} + \mathbf{\Psi}^* \right) \mathbf{\Lambda}_y' + \mathbf{\Theta}_\varepsilon$$

$$\mathbf{\Sigma}_{xy} = \mathbf{\Sigma}_{yx}' = \mathbf{\Lambda}_x \mathbf{\Phi\Gamma}' \left(\mathbf{I}-\mathbf{B}\right)^{-1'} \mathbf{\Lambda}_y' + \mathbf{\Theta}_{\delta\varepsilon}$$

$$\mathbf{\Sigma}_{xx} = \mathbf{\Lambda}_x \mathbf{\Phi\Lambda}_x' + \mathbf{\Theta}_\delta$$

$$\mathbf{\Sigma}_{\eta y} = \left((\mathbf{I}-\mathbf{B})^{-1}\mathbf{\Gamma\Phi\Gamma}'(\mathbf{I}-\mathbf{B})^{-1'} + \mathbf{\Psi}^* \right) \mathbf{\Lambda}_y'$$

$$\mathbf{\Sigma}_{\eta x} = \mathbf{\Sigma}_{\eta x}' = \left(\mathbf{I}-\mathbf{B}\right)^{-1} \mathbf{\Gamma\Phi\Lambda}_x'$$

$$\mathbf{\Sigma}_{\eta\eta} = (\mathbf{I}-\mathbf{B})^{-1}\mathbf{\Gamma\Phi\Gamma}'(\mathbf{I}-\mathbf{B})^{-1'} + \mathbf{\Psi}^*$$

$$\mathbf{\Sigma}_{\xi y} = \mathbf{\Phi\Gamma}'(\mathbf{I}-\mathbf{B})^{-1'} \mathbf{\Lambda}_y'$$

$$\mathbf{\Sigma}_{\xi x} = \mathbf{\Sigma}_{x\xi}' = \mathbf{\Phi\Lambda}_x'$$

$$\mathbf{\Sigma}_{\xi\eta} = \mathbf{\Sigma}_{\eta\xi}' = \mathbf{\Phi\Gamma}'(\mathbf{I}-\mathbf{B})^{-1'}$$

$$\mathbf{\Sigma}_{\xi\xi} = \mathbf{\Phi}$$

where

$$\mathbf{\Psi}^* = (\mathbf{I}-\mathbf{B})^{-1}\mathbf{\Psi}(\mathbf{I}-\mathbf{B})^{-1'}$$

Variances of the latent variables

The covariance matrices of the p_η endogenous latent variables, $\boldsymbol{\eta}$, and the p_ξ exogenous latent variables, $\boldsymbol{\xi}$, are given by

$$\mathbf{\Sigma}_{\eta\eta} = (\mathbf{I}-\mathbf{B})^{-1}\left[\mathbf{\Gamma\Phi\Gamma}' + \mathbf{\Psi}\right](\mathbf{I}-\mathbf{B})^{-1'}$$

and

$$\mathbf{\Sigma}_{\xi\xi} = \mathbf{\Phi}$$

respectively.

The variances of the exogenous latent variables are the diagonal elements of $\mathbf{\Phi}$ and are parameters of the LISREL model. As a result, these variances can be fixed to unity to ensure that the corresponding factor loadings (elements of $\mathbf{\Lambda}_x$) are standardized.

The variances of the endogenous latent variables are the diagonal elements of $\mathbf{\Sigma}_{\eta\eta}$ and are not parameters of the general LISREL model. Instead, they are complex functions of the regression weights, variances and covariances of the exogenous variables, and the regression error variances and covariances of the general LISREL model. As a result, equality constraints are required to constrain the variances of the endogenous latent variables to unity to ensure that the corresponding factor loadings (elements of $\mathbf{\Lambda}_y$) are standardized. In the case of the general LISREL model, these constraints may be expressed as

$$[\Upsilon]_{p+i,p+i} - 1 = 0 \quad \forall\ i = 1, 2, \cdots, m_\eta$$

Typical elements of the Jacobian matrix of these equality constraints (Mels 1988) follow as

$$[\mathbf{L}]_{ik} = 2c_i(\theta_k)\, d_i(\theta_k)$$

where

$$\mathbf{c}(\theta_k) = \begin{cases} [\mathbf{\Lambda}]_{.,p+i} & \text{if } \theta_k \text{ denotes a regression weight} \\ (1+\partial_{ij})^{-1}[\mathbf{\Lambda}]_{.,p+i} & \text{if } \theta_k \text{ denotes a variance or covariance} \end{cases}$$

and

$$\mathbf{d}(\theta_k) = \begin{cases} [\Upsilon]_{.,p+j} & \text{if } \theta_k \text{ denotes a regression weight} \\ [\mathbf{\Lambda}]_{.,p+j} & \text{if } \theta_k \text{ denotes a variance or covariance} \end{cases}$$

where i and j denote the row and column of the parameter of the general LISREL model, respectively, $[\mathbf{A}]_{.j}$ denotes the j^{th} column of the matrix $\mathbf{A}$, and ∂_{ij} denotes the Kronecker delta.

Gauss-Newton algorithm

Suppose that the elements of $\boldsymbol{\theta}$ consist of the q unknown parameters of the general LISREL model and that $f(\cdot)$ denotes the discrepancy function to be minimized with respect to the elements of $\boldsymbol{\theta}$ subject to the constraints that the variances of the endogenous latent variables are equal to unity.

Let $\boldsymbol{\lambda}$ denote the vector of Lagrange multipliers associated with the m_η equality constraints for the variances of the endogenous latent variables. If $\hat{\boldsymbol{\theta}}^{(t)}$ and $\hat{\boldsymbol{\lambda}}^{(t)}$ denote the t^{th} successive approximation to the estimators $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\lambda}}$ respectively, then the $(t+1)^{\text{st}}$ approximation (Browne and Du Toit, 1992) is obtained from

$$\begin{bmatrix} \hat{\boldsymbol{\theta}}^{(t+1)} \\ \hat{\boldsymbol{\lambda}}^{(t+1)} \end{bmatrix} = \begin{bmatrix} \hat{\boldsymbol{\theta}}^{(t)} \\ \hat{\boldsymbol{\lambda}}^{(t)} \end{bmatrix} - \alpha_t \begin{bmatrix} \mathbf{H}\left(\hat{\boldsymbol{\theta}}^{(t)}\right) & \mathbf{L}'\left(\hat{\boldsymbol{\theta}}^{(t)}\right) \\ \mathbf{L}\left(\hat{\boldsymbol{\theta}}^{(t)}\right) & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{g}\left(\hat{\boldsymbol{\theta}}^{(t)}\right) \\ \mathbf{c}\left(\hat{\boldsymbol{\lambda}}^{(t)}\right) \end{bmatrix}$$

where $\mathbf{g}(\cdot)$ denotes the gradient vector of $f(\cdot)$, $\mathbf{H}(\cdot)$ denotes the approximate Hessian matrix of $f(\cdot)$, $\mathbf{L}(\cdot)$ denotes the Jacobian matrix of the equality constraints for the variances of the endogenous latent variables with respect to the elements of $\boldsymbol{\theta}$, and α_t denotes a selected step-size parameter (Browne 1982) to ensure that

$$f\left(\boldsymbol{\theta}^{(t+1)}\right) + 2\sum_{i=1}^{m_\eta}\left|\lambda_i\left(\boldsymbol{\theta}^{(t)}\right)c_i\left(\boldsymbol{\theta}^{(t)}\right)\right| < f\left(\boldsymbol{\theta}^{(t)}\right) + 2\sum_{i=1}^{m_\eta}\left|\lambda_i\left(\boldsymbol{\theta}^{(t)}\right)c_i\left(\boldsymbol{\theta}^{(t)}\right)\right|$$

with $\alpha_t = 1$ in most cases. The maximum absolute residual cosine is the maximum absolute value of

$$\begin{bmatrix} \mathbf{g}\left(\hat{\boldsymbol{\theta}}^{(t)}\right) \\ \mathbf{c}\left(\hat{\boldsymbol{\lambda}}^{(t)}\right) \end{bmatrix}_i \Bigg/ \sqrt{\begin{bmatrix} \mathbf{H}\left(\hat{\boldsymbol{\theta}}^{(t)}\right) & \mathbf{L}'\left(\hat{\boldsymbol{\theta}}^{(t)}\right) \\ \mathbf{L}\left(\hat{\boldsymbol{\theta}}^{(t)}\right) & \mathbf{0} \end{bmatrix}_{ii} f\left(\hat{\boldsymbol{\theta}}^{(t)}\right)} \quad \forall\ i = 1, 2, \cdots, q + m_\eta$$

Iteration is terminated when the maximum absolute residual cosine falls below the tolerance limit of $\varepsilon_{rc} = 10^{-4}$ and the maximum absolute constraint falls below the tolerance limit of $\varepsilon_c = 10^{-6}$ (Browne 1982).

After convergence, estimated standard errors of the estimators and the Lagrange multipliers are computed as the positive square roots of the diagonal elements of the matrix

$$\frac{1}{n-1}\begin{bmatrix} \mathbf{H}\left(\hat{\boldsymbol{\theta}}^{(t)}\right) & \mathbf{L}'\left(\hat{\boldsymbol{\theta}}^{(t)}\right) \\ \mathbf{L}\left(\hat{\boldsymbol{\theta}}^{(t)}\right) & \mathbf{0} \end{bmatrix}^{-1}$$

where n denotes the sample size.

If the data distribution of the continuous observed variables is assumed to be a multivariate Normal distribution, typical elements of the gradient vector and the approximate Hessian matrix (Mels 1988) may be expressed as

$$\left[\mathbf{g}\right]_k = -2\mathbf{a}'(\theta_k)\boldsymbol{\Omega}\mathbf{b}(\theta_k)$$

and

$$\left[\mathbf{H}\right]_{kl} = 2\left(\mathbf{a}'(\theta_k)\mathbf{b}(\theta_k)\mathbf{a}'(\theta_l)\mathbf{b}(\theta_l) + \mathbf{a}'(\theta_k)\mathbf{b}(\theta_l)\mathbf{b}'(\theta_k)\mathbf{a}(\theta_l)\right)$$

where

$$\boldsymbol{\Omega} = \mathbf{V}^{-1/2}\left(\mathbf{S} - \boldsymbol{\Lambda\Phi\Lambda}'\right)\mathbf{V}^{-1/2\prime}$$

where $\mathbf{V} = \left\{\boldsymbol{\Sigma}\left(\hat{\boldsymbol{\theta}}\right)\right\}^{-1}$ for maximum likelihood estimation, $\mathbf{V} = \mathbf{S}^{-1}$ for generalized least squares estimation, and $\mathbf{V} = \mathbf{I}_p$ for unweighted least squares estimation and

$$\mathbf{a}\left(\theta_k\right) = \begin{cases} \left[\boldsymbol{\Lambda}_p\right]_{.i} & \text{if } \theta_k \text{ denotes a regression weight} \\ \left(1 + \partial_{ij}\right)^{-1}\left[\boldsymbol{\Lambda}_p\right]_{.i} & \text{if } \theta_k \text{ denotes a variance or covariance} \end{cases}$$

and

$$\mathbf{b}\left(\theta_k\right) = \begin{cases} \left[\Upsilon_p\right]_{.j} & \text{if } \theta_k \text{ denotes a regression weight} \\ \left[\boldsymbol{\Lambda}_p\right]_{.j} & \text{if } \theta_k \text{ denotes a variance or covariance} \end{cases}$$

where $\boldsymbol{\Lambda}_p$ and Υ_p contain the first p rows of $\boldsymbol{\Lambda}$ and Υ, respectively, i and j denote the row and column of the parameter of the general LISREL model, respectively, $\left[\mathbf{A}\right]_{.j}$ denotes the j^{th} column of the matrix $\mathbf{A}$, and ∂_{ij} denotes the Kronecker delta.

If the data distribution of the continuous observed variables is a multivariate distribution with finite eight-order moments, the gradient vector, and the approximate Hessian matrix (Browne 1982, 1984) of the weighted least squares (distribution free) discrepancy function

$$f(\boldsymbol{\theta}) = (\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta}))' \mathbf{W}^{-1} (\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta}))$$

are given by

$$\mathbf{g}(\boldsymbol{\theta}) = -\boldsymbol{\Delta}'\mathbf{W}^{-1}(\mathbf{s} - \boldsymbol{\sigma}(\boldsymbol{\theta}))$$

and

$$\mathbf{H}(\boldsymbol{\theta}) = \boldsymbol{\Delta}'\mathbf{W}^{-1}\boldsymbol{\Delta}'$$

where $\mathbf{s}$ denotes the $p \times (p+1)/2 \times 1$ vector consisting of the nonduplicated elements of the sample covariance matrix $\mathbf{S}$, $\boldsymbol{\sigma}(\boldsymbol{\theta})$ denotes the $p \times (p+1)/2 \times 1$ vector consisting of the nonduplicated elements of the general LISREL model for the population covariance matrix, $\mathbf{W}$ is a weight matrix with typical elements (Browne 1984) given by

$$w_{ij,kl} = w_{ijkl} - w_{ij} w_{kl}$$

$\boldsymbol{\Delta}$ is the Jacobian matrix for $\boldsymbol{\sigma}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$, and

$$w_{ijkl} = \frac{1}{n}\sum_{m=1}^{n}(x_{im} - \bar{x}_i)(x_{jm} - \bar{x}_j)(x_{km} - \bar{x}_k)(x_{lm} - \bar{x}_l)$$

$$w_{ij} = \frac{1}{n}\sum_{m=1}^{n}(x_{im} - \bar{x}_i)(x_{jm} - \bar{x}_j)$$

where

$$\bar{x}_i = \frac{1}{n}\sum_{m=1}^{n} x_{im}$$

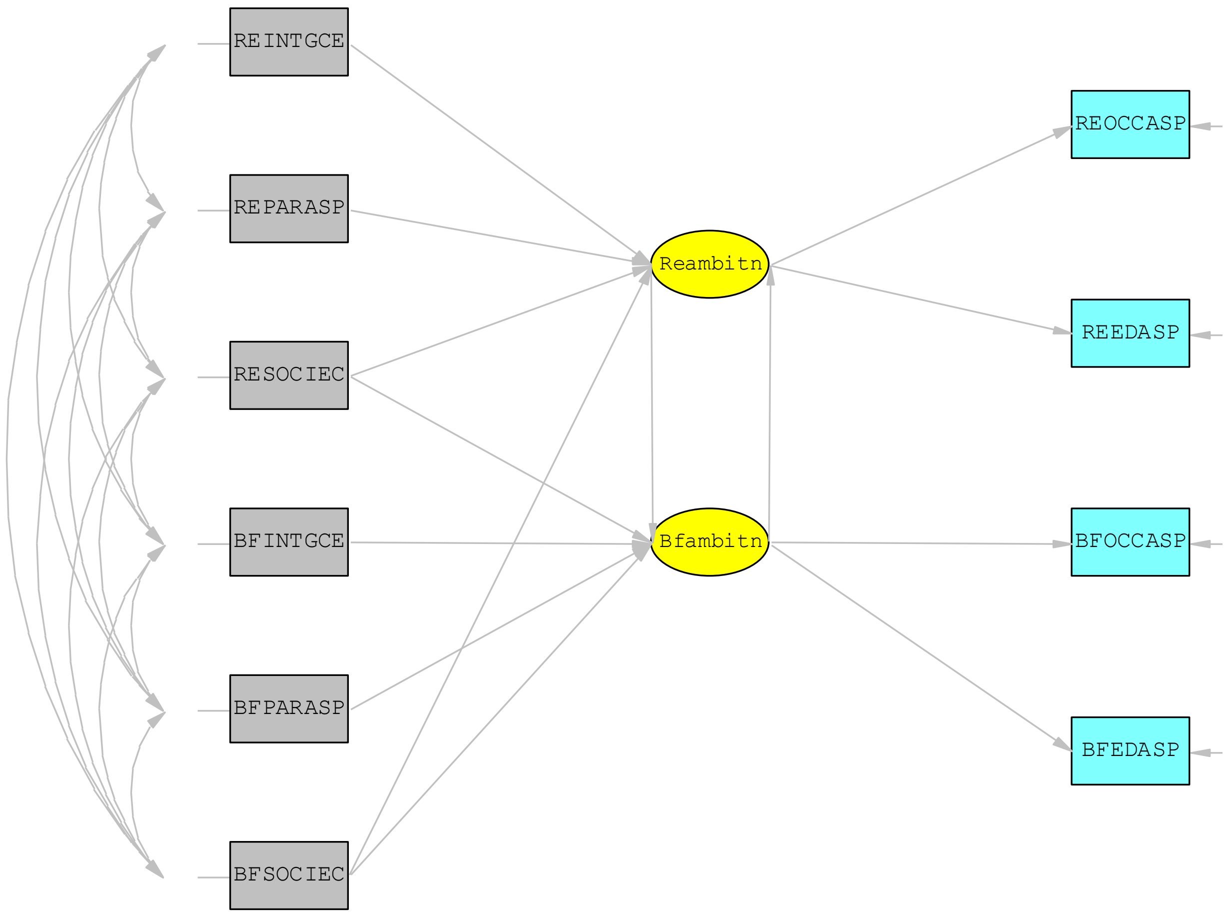
3.2 MIMIC model for peer influences on ambition

The data are scores for the occupational aspiration, the educational aspiration, the intelligence, the socio-economic status, and the parental aspiration of 329 students and their best friends at a Michigan high school used in a study by Duncan, Haller, and Portes (1971). The corresponding data file is **PEERS.LSF**, and the first few observations are depicted below.

PEERS.LSF

	REINTGCE	REPARASP	RESOCIEC	REOCCASP	REEDASP	BFINTGCE	BFPARASP	BFSOCIEC	BFOCCASP
1	-0.159	-0.191	1.338	0.106	0.764	0.696	1.162	0.199	-0.536
2	0.624	-0.503	0.409	-1.018	0.004	-0.492	0.392	0.648	1.159
3	-0.427	-0.349	-1.425	-0.794	-0.724	-0.151	1.008	-2.476	0.124
4	-0.150	1.249	0.356	-0.436	0.011	-0.923	-0.395	-1.509	-1.252
5	0.537	0.673	-1.394	0.697	1.094	-0.101	-0.353	0.523	1.522
6	-0.509	0.857	0.780	0.104	-0.553	-1.191	-0.601	0.033	-0.458
7	-1.680	-0.234	-0.508	-1.240	-1.268	0.302	0.047	-1.796	0.944
8	-0.866	-1.236	1.250	-0.955	-0.418	-0.041	-0.490	0.378	-0.229
9	0.924	-0.948	-0.559	0.243	0.257	0.999	-0.020	-0.040	0.351
10	0.509	-0.607	1.740	1.038	-0.128	1.070	-0.619	1.192	1.266

The theoretical model is a Multiple Indicators, Multiple Causes (MIMIC) model that suggests that the respondent's parental aspiration, intelligence, and socio-economic status along with the best friend's socio-economic status are causes of the respondent's ambition and that the best friend's parental aspiration, intelligence, and socio-economic status along with the respondent's socio-economic status are causes of the best friend's ambition. A path diagram for this model is shown in the image below.



The SIMPLIS syntax file to fit the theoretical model above to the student data is shown in the image below.

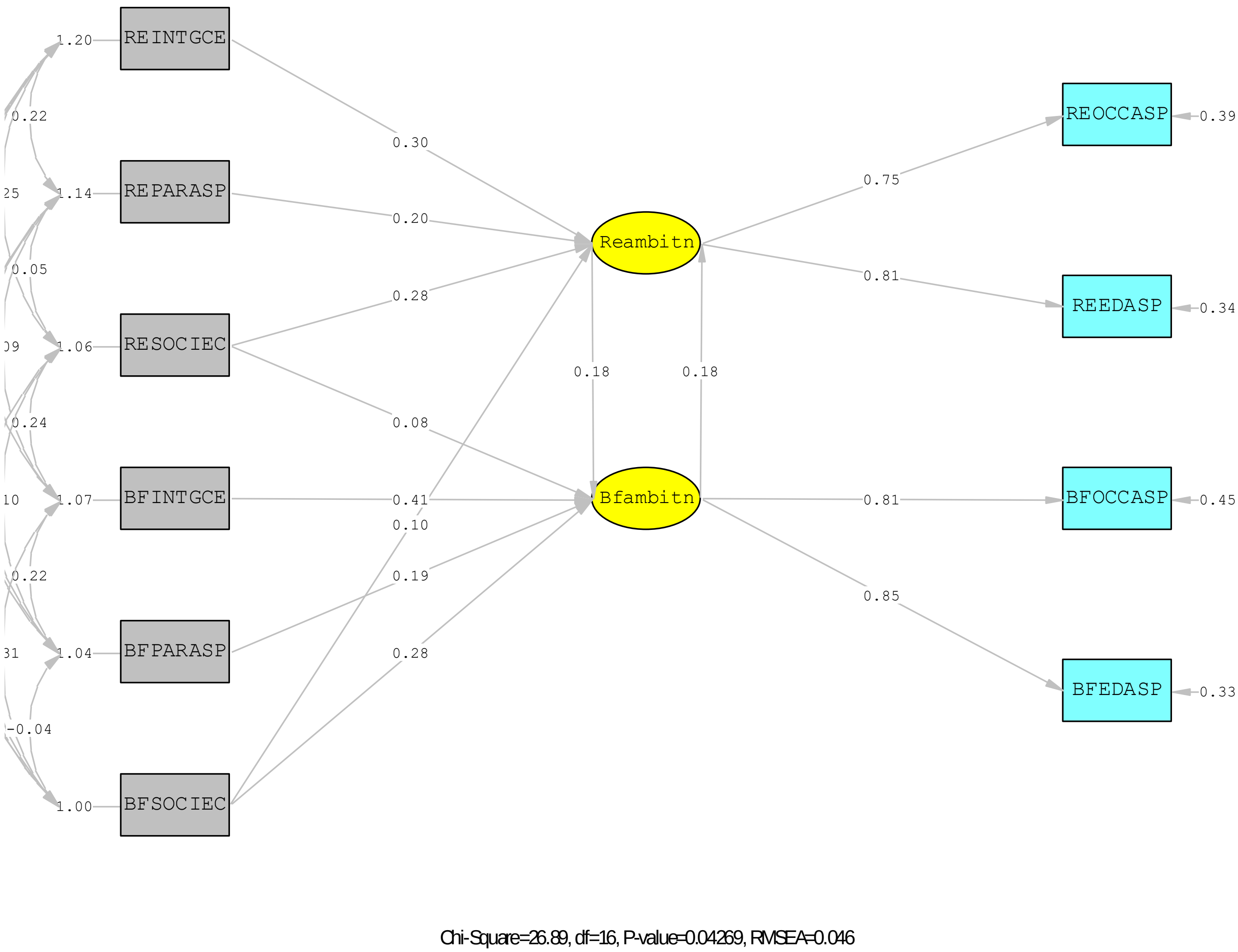
PEERS.SPL

```
Raw Data from File PEERS.LSF
Latent Variables
Reambitn Bfambitn
Relationships
REOCCASP REEDASP = Reambitn
BFOCCASP BFEDASP = Bfambitn
Reambitn = Bfambitn REPARASP REINTGCE RESOCIEC BFSOCIEC
Bfambitn = Reambitn RESOCIEC BFSOCIEC BFINTGCE BFPARASP
Options: SO SC
Path Diagram
End of Problem
```

- Line 1 specifies the data file.
- Lines 2 to 3 specify the labels for the latent variables of the model.
- Lines 4 to 8 specify the theoretical model.
- Line 9 requests that all the factor loadings for the two endogenous latent variables are estimated by constraining their variances to be equal to unity (SO option) and the completely standardized solution (SC option).

- Line 10 requests a path diagram of the model.
- Line 11 indicates that no more SIMPLIS commands are to be processed.

If this SPL file is opened in LISREL and the **Run LISREL** icon is clicked, the following path diagram is obtained.



The corresponding output file, **PEERS.OUT**, is opened in a separate window. The completely standardized estimates along with the standard error estimates, the test statistic values, and the exceedance probabilities for the free parameters of the measurement and structural models, which are listed in this file, are shown in the images below.

PEERS.OUT

```
Multiple Indicators, Multiple Causes (MIMIC) model

Completely Standardized Solution

        Measurement Equations


REOCCASP=0.766*Reambitn, Errorvar.= 0.413
90% CILL(0.703;                    (0.329;
90% CIUL 0.817)                     0.501)
Standerr(0.0345)                   (0.0529)
Z-value   22.214                    7.804
P-value   0.000                     0.000

 REEDASP=0.814*Reambitn, Errorvar.= 0.338
90% CILL(0.752;                    (0.255;
90% CIUL 0.862)                     0.431)
Standerr(0.0332)                   (0.0540)
Z-value   24.523                    6.249
P-value   0.000                     0.000

BFOCCASP=0.772*Bfambitn, Errorvar.= 0.404
90% CILL(0.715;                    (0.328;
90% CIUL 0.818)                     0.485)
Standerr(0.0312)                   (0.0482)
Z-value   24.747                    8.389
P-value   0.000                     0.000

 BFEDASP=0.828*Bfambitn, Errorvar.= 0.314
90% CILL(0.774;                    (0.240;
90% CIUL 0.871)                     0.398)
Standerr(0.0291)                   (0.0482)
Z-value   28.487                    6.511
P-value   0.000                     0.000
```

PEERS.OUT

```
         Structural Equations


Reambitn=0.175*Bfambitn + 0.332*REINTGCE + 0.214*REPARASP + 0.290*RESOCIEC + 0.103*BFSOCIEC, Errorvar.= 0.479
90% CILL(0.0310;         (0.245;          (0.132;          (0.199;          (0.00175;                   (0.390;
90% CIUL 0.313)           0.414)           0.293)           0.376)           0.202)                      0.570)
Standerr(0.0862)         (0.0514)         (0.0490)         (0.0538)         (0.0612)                    (0.0555)
Z-value   2.035           6.465            4.363            5.386            1.685                       8.640
P-value   0.042           0.000            0.000            0.000            0.092                       0.000

Bfambitn=0.184*Reambitn + 0.0867*RESOCIEC + 0.428*BFINTGCE + 0.197*BFPARASP + 0.282*BFSOCIEC, Errorvar.= 0.384
90% CILL(0.0528;         (-0.005;          (0.346;          (0.120;          (0.198;                    (0.305;
90% CIUL 0.309)           0.177)            0.503)           0.272)           0.363)                     0.470)
Standerr(0.0783)         (0.0558)          (0.0475)         (0.0461)         (0.0502)                   (0.0506)
Z-value   2.350           1.553             8.995            4.272            5.620                      7.591
P-value   0.019           0.120             0.000            0.000            0.000                      0.000
```

These results agree with those obtained by fitting the theoretical model correctly to the sample correlation matrix using the special statistical methods implemented in Steiger (1995) and Browne and Mels (1996).

4. References

Aish, A.M. & Jöreskog, K.G. (1990). A panel model for political efficacy and responsiveness: An application of LISREL 7 with weighted least squares. *Quality and Quantity*, **24**, 405-426.

Bishop, Y.M.M., Feinberg, S.E. & Holland, P.W. (1988). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge: MIT Press.

Brand, J.P.L. (1999). *Development, Implementation, and Evaluation of Multiple Imputation Strategies for the Statistical Analysis of Incomplete Data Sets*. Unpublished Ph.D. thesis, Erasmus University.

Browne, M.W. (1982). Covariance Structures. In D.M. Hawkins (Ed.), *Topics in Applied Multivariate Analysis*, pp. 72-141. Cambridge: Cambridge University Press.

Browne, M.W. (1984). Asymptotically Distribution-free Methods in the Analysis of Covariance Structures. *British Journal of Mathematical and Statistical Psychology*, **37**, 62-83.

Browne, M.W. & Du Toit, S.H.C. (1992). Automated Fitting of Nonstandard Models. *Multivariate Behavioral Research*, **27**, 269-300.

Browne, M.W. & Mels, G. (1996). *Path Analysis: RAMONA, SYSTAT FOR DOS; Advanced Applications, Version 6.* Evanston, IL: Systat Inc., 163-224.

Chung, S, and Cai, L. (2019). Alternative Multiple Imputation Inference for Categorical Structural Equation Modeling. *Multivariate Behavioral Research*, **54(3)**, 323–337.

Cudeck, R. (1989). The Analysis of Correlation Matrices using Covariance Structure Models. *Psychological Bulletin*, **105**, 317-327.

Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society*, Series B, **39**, 1–38.

Duncan, O.D., Haller, A.O. & Portes, A. (1971). *Peer Influences on Aspirations, a Reinterpretation*. In H.M. Blalock (Ed.), Causal Models in the Social Sciences, pp. 219-244. Aldine Atherton, Inc.

Jöreskog, K. G. (1994). On the estimation of polychoric correlations and their asymptotic covariance matrix. *Psychometrika*, **59(3)**, 381–389.

Jöreskog, K.G. (1977). Structural Equation Models in the Social Sciences: Specification, Estimation and Testing. In P.R. Krishnaiah (Ed.), *Applications of Statistics*, pp. 265-287. Amsterdam: North Holland.

Jöreskog, K.G. (1973). A General Method for Estimating a Linear Structural Equation System. In: A.S. Goldberger & O.D. Duncan (Eds.), *Structural Equation Models in the Social Sciences*. New York: Seminar Press.

Jöreskog, K.G. & Sörbom, D. (1999). *LISREL 8: User's reference guide*. Lincolnwood, IL: Scientific Software International, Inc.

Mels, G. (1988). *A General System for Path Analysis with Latent Variables*. Unpublished M.Sc. dissertation. Pretoria: University of South Africa.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, **49(1)**, 115–132.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, **44(4)**, 443–460.

Olsson, U., Drasgrow, F., and Dorans, N.J. (1992). The polyserial correlation coefficient. *Psychometrika*, **47(3)**, 337–347.

Pearson, K. (1896), Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity and Panmixia, *Philosophical Transactions of the Royal Society of London*, **187**, 253-318.

Shapiro, A. & Browne, M.W. (1990). On the Treatment of Correlation Structures as Covariance Structures. *Linear Algebra and its Applications*, **127**, 547-567.

Steiger, J.H. (1995). Structural Equation Modelling. In: *Statistica for Windows (Volume III): Statistics II, 2nd Edition*, pp. 3539-3688. Tulsa: StatSoft, Inc.

Steiger, J.H. & Hakstian, A.R. (1982). A Historical Note on the Asymptotic Distribution of Correlations. *British Journal of Mathematical and Statistical Psychology*, **36**, 157.

Van Buuren, S. (2007). Multiple Imputation of Discrete and Continuous Data by Fully Conditional Specification. *Statistical Methods in Medical Research*, **16**, 219–242.

Yuan, K.-H., & Bentler, P. M. (1997). Mean and covariance structure analysis: Theoretical and practical improvements. *Journal of the American Statistical Association*, **92(438)**, 767–774.