Path analysis with latent variables

Path analysis may be formulated for latent variables as well as for directly observed variables. In its most general form, there is a structural equation system of the form

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

for a set of latent variables η's and ξ's. In most applications, the system is recursive but models with non-recursive systems have also been proposed. In this section, a non-recursive system for latent variables and various alternative models are considered sequentially. The example is based on data from Duncan, Haller & Portes (1968).

The figure above is meant to represent the way in which a person's peers (e.g., best friends) influence his or her decisions (e.g., choice of occupation). The paths between respondent's ambition (REambitn) and best friend's ambition (Bfambitn) recognize that the relation must be reciprocal. As a test of this model, a sample of Michigan high-school students were paired with their best friends and measured on a number of background variables. In addition, scaled measures of occupational and educational aspiration were obtained to serve as indicators of a latent variable AMBITION.

The observed measures in the study are:

x_1 = respondent's parental aspiration (REPARASP)

x_2 = respondent's intelligence (REINTGCE)

x_3 = respondent's socioeconomic status (RESOCIEC)

x_4 = best friend's socioeconomic status (BFSOCIEC)

x_5 = best friend's intelligence (BFINTGCE)

x_6 = best friend's parental aspiration (BFPARASP)

y_1 = respondent's occupational aspiration (REOCCASP)

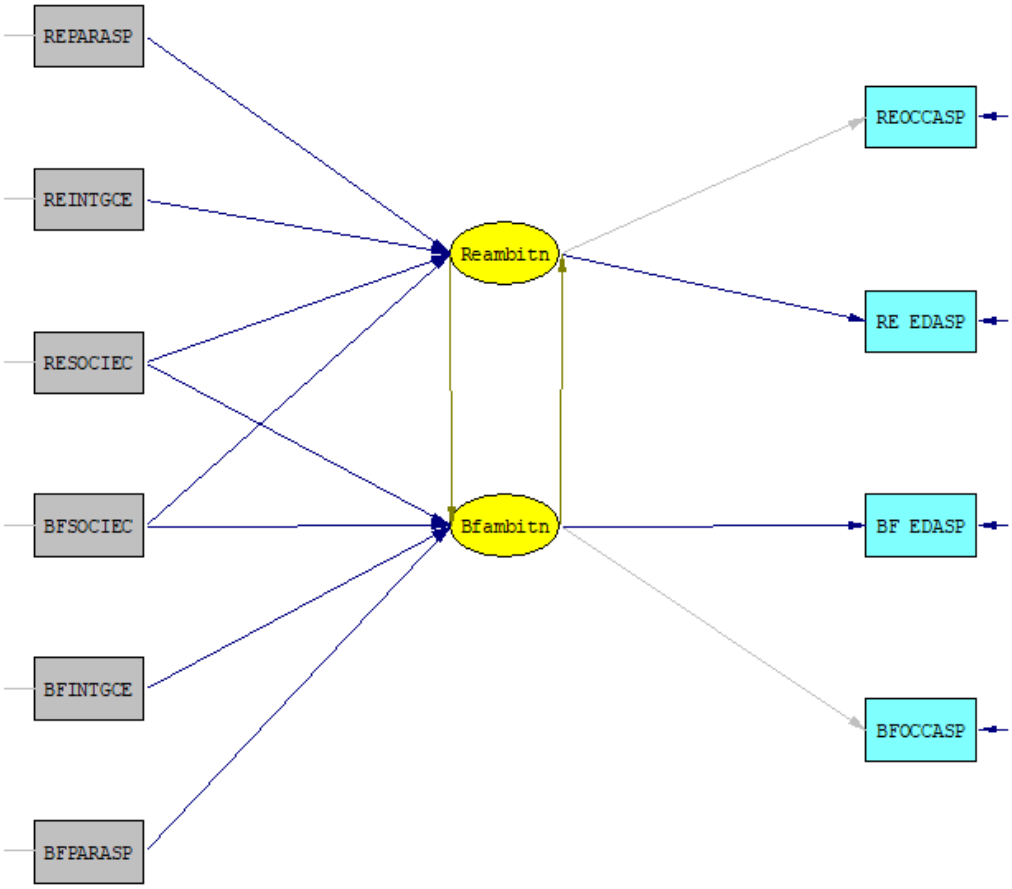
y_2 = respondent's educational aspiration (REEDASP)

y_3 = best friend's educational aspiration (BFEDASP)

y_4 = best friend's occupational aspiration (BFOCCASP)

η_1 = respondent's ambition (REAMBITN)

η_2 = best friend's ambition (BFAMBITN)



The correlation matrix based on 329 observations is shown in the table below. For correct standard errors and chi-squares, note that the covariances should be analyzed.

Table: Correlations for background and aspiration measures for 329 respondents and their best friends

Respondent										
REINTGCE	1.0000									
REPARASP	-/1839	1.0000								
RESOCIEC	0.2220	0.0489	1.0000							
REOCCASP	0.4105	0.2137	0.3240	1.0000						
REEDASP	0.4043	0.2742	0.4047	0.6247	1.0000					
Best friend										
BFINTGCE	0.3355	0.0782	0.2302	0.2995	0.2863	1.0000				
BFPARASP	0.1021	0.1147	0.0931	0.0760	0.0702	0.2087	1.0000			
BFSOCIEC	0.1861	0.0186	0.2707	0.2930	0.2407	0.2950	-0.0438	1.0000		
BFOCCASP	0.2598	0.0839	0.2786	0.4216	0.3275	0.5007	0.1988	0.3607	1.0000	
BFEDASP	0.2903	0.1124	0.3054	0.3269	0.3669	0.5191	0.2784	0.4105	0.6404	1.0000

The specifications for the LISREL model are as follows.

Let $\xi_i \equiv x_i$ or $\mathbf{\Lambda}_x(6\times 6) = \mathbf{I}$ and $\mathbf{\Theta}_\delta = \mathbf{0}$ (specified by the FI option). And let $\lambda_2 = \lambda_{21}^{(y)}, \lambda_3 = \lambda_{32}^{(y)}, \beta_1 = \beta_{12}, \beta_2 = \beta_{21}, \quad \gamma_1 = \gamma_{11}, \gamma_2 = \gamma_{12}, \quad \gamma_3 = \gamma_{13}, \gamma_4 = \gamma_{14}, \quad \gamma_5 = \gamma_{23}, \gamma_6 = \gamma_{24}, \gamma_7 = \gamma_{25}$, and $\gamma_8 = \gamma_{26}$.

Then the structural equations are

$$\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} 0 & \beta_1 \\ \beta_2 & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 & 0 & 0 \\ 0 & 0 & \gamma_5 & \lambda_6 & \lambda_7 & \lambda_8 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix},$$

and the equations relating the η 's to the *y*'s are

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_2 & 0 \\ 0 & \lambda_3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}.$$

In $\mathbf{\Lambda}_y$ the scales for η_1 and η_2 are fixed the same as in y_1 and y_4, respectively. Since y_1 and y_4 are on the same scale, η_1 and η_2 are also on the same scale, as is necessary to make meaningful comparisons between the respondent and the best friend. Since $\boldsymbol{\xi} \equiv \mathbf{x}$ and there are no constraints on $\mathbf{\Phi}$, $\mathbf{\Phi} = \mathbf{\Sigma}_{xx}$, which is estimated as $\mathbf{S}_{xx}$. The matrix $\mathbf{\Psi}(2\times 2)$ is

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

with $\psi_{ii} = Var(\zeta_i)$ for *i* = 1,2, and $\psi_{21} = Cov(\zeta_1, \zeta_2)$. Since $\boldsymbol{\xi} \equiv \mathbf{x}$, the structural equations are equivalent to

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

with reduced form

$$\boldsymbol{\eta} = (\mathbf{I} - \mathbf{B})^{-1}\mathbf{\Gamma}\mathbf{x} + (\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\zeta}.$$

In addition,

$$\mathbf{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} \left(\mathbf{I} - \mathbf{B}\right)^{-1} \boldsymbol{\Gamma} \mathbf{x} + \boldsymbol{\Lambda}_y \left(\mathbf{I} - \mathbf{B}\right)^{-1} \boldsymbol{\zeta} + \boldsymbol{\varepsilon} = \mathbf{P}\mathbf{x} + \mathbf{z},$$

$$\mathbf{z} = \boldsymbol{\Lambda}_y \left(\mathbf{I} - \mathbf{B}\right)^{-1} \boldsymbol{\zeta} + \boldsymbol{\varepsilon}.$$

Identification:

Because $\boldsymbol{\Psi}\left(2 \times 2\right)$ is unconstrained, it follows that, for any given $\mathbf{B}$ and $\boldsymbol{\Gamma}$, there is a one-to-one correspondence between $\boldsymbol{\Psi}$ and the covariance matrix of $\boldsymbol{\eta}$, $\boldsymbol{\Omega}$ say, where

$$\boldsymbol{\Omega} = \begin{bmatrix} \omega_{11} & \\ \omega_{21} & \omega 22 \end{bmatrix}.$$

Also, $\boldsymbol{\Sigma}_{yy} = \boldsymbol{\Lambda}_y \, \boldsymbol{\Omega} \, \boldsymbol{\Lambda}_y^{'} + \boldsymbol{\Theta}_{\varepsilon}$

$$= \begin{bmatrix} \omega_{11} + \theta_{11}^{(\varepsilon)} & & & \\ \lambda_2 \omega_{11} & \lambda_2^2 \omega_{11} + \theta_{22}^{(\varepsilon)} & & \\ \lambda_3 \omega_{21} & \lambda_2 \lambda_3 \omega_{21} & \lambda_3^2 \omega_{22} + \theta_{33}^{(\varepsilon)} & \\ \omega_{21} & \lambda_2 \omega_{21} & \lambda_3 \omega_{22} & \omega_{22} + \theta_{44}^{(\varepsilon)} \end{bmatrix}.$$

In addition,

$$\left(\mathbf{I} - \mathbf{B}\right)^{-1} = \left(1 - \beta_1 \beta_2\right)^{-1} \begin{bmatrix} 1 & \beta_1 \\ \beta_2 & 1 \end{bmatrix}$$

$$\boldsymbol{\Pi} = \left(1 - \beta_1 \beta_2\right)^{-1} \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 + \beta_1 \gamma_5 & \gamma_4 + \beta_1 \gamma_6 & \beta_1 \gamma_7 & \beta_1 \gamma_8 \\ \beta_2 \gamma_1 & \beta_2 \gamma_2 & \gamma_5 + \beta_2 \gamma_3 & \gamma_6 + \beta_2 \gamma_4 & \gamma_7 & \gamma_8 \end{bmatrix}.$$

The first and last rows of P are identical to the first and second row of $\boldsymbol{\Pi} = \left(\mathbf{I} - \mathbf{B}\right)^{-1} \boldsymbol{\Gamma}$, respectively. The second row of $\mathbf{P}$ is λ_2 times the first row of $\boldsymbol{\Pi}$ and the third row of $\mathbf{P}$ is λ_3 times the second row of $\boldsymbol{\Pi}$. Hence it is clear that λ_2 and λ_3 are identified and that

$$\lambda_2 = \frac{p_{2i}}{p_{1i}} \quad \text{and}$$

$$\lambda_3 = \frac{p_{3i}}{p_{4i}}, \quad i = 1, 2, ..., 6.$$

Since $\mathbf{\Pi}$ consists of two rows of **P**, $\mathbf{\Pi}$ is identified. From $\mathbf{\Pi}$ it follows that $\gamma_1, \gamma_2, \gamma_7, \gamma_8, \beta_1,$ and β_2 are determined as

$$\gamma_1 = \pi_{11}, \quad \gamma_2 = \pi_{12}, \quad \gamma_7 = \pi_{25}, \quad \gamma_8 = \pi_{26},$$
$$\beta_1 = (\pi_{15} / \pi_{25}) = (\pi_{16} / \pi_{26}), \quad \beta_2 = (\pi_{21} / \pi_{11}) = (\pi_{22} / \pi_{12}).$$

Then $\gamma_3, \gamma_4, \gamma_5$ and γ_6 are obtained from $\pi_{13}, \pi_{14}, \pi_{23},$ and π_{24}. With λ_2 and λ_3 determined we can now obtain ω_{11}, ω_{21}, and ω_{22} from the off-diagonal elements of $\mathbf{\Sigma}_{yy}$. Finally, the $\theta_{ii}^{(\varepsilon)}$ (i = 1, 2l 3l 4) can be determined from the diagonal elements of $\mathbf{\Sigma}_{yy}$.

This analysis shows that all parameters are identified. Altogether there are 19 parameters (2 β's, 8 γ's, 3 ω's, 2 λ's, and 4 $\theta^{(\varepsilon)}$'s) if $\mathbf{x}$ is fixed. When $\mathbf{x}$ is random there will be an additional 21 parameters in $\mathbf{\Phi} = \mathbf{\Sigma}_{xx}$. In both cases the degrees of freedom will be 15.

The LISREL command file for the analysis is (**EX55A.LIS** in the **LISREL Examples** folder):

```
Peer Influences on Ambition: Model with BE(2,1) = BE(1,2) and PS(2,1) = 0
DA NI=10 NO=329
LA FI=EX55.LAB
KM FI=EX55.COR
SELECT
4 5 10 9 2 1 3 8 6 7
MO NY=4 NE=2 NX=6 FIXED-X BE=FU
LE
Reambitn Bfambitn
FR LY(2,1) LY(3,2) BE(1,2)
FI GA(5)-GA(8)
VA 1 LY(1) LY(8)
EQ BE(1,2) BE(2,1)
OU SE TV EF SS
```

The labels and correlations are read from the files **EX55.LAB** and **EX55.COR**, respectively. The variables in the correlation matrix are in the order x_2, x_1, x_3, y_1, y_2, x_5, x_6, x_4, y_4, and y_3. The SE command is necessary to order the variables to correspond to the model. The MO command specifies FIXED-X which takes care of NK = 6, LX = ID, TD = ZE, and PH = $\mathbf{S}_{xx}$. The other parameter matrices are default except PS = DI and BE = FU. The FR command specifies the free elements in LY and BE and the FI command specifies the four fixed zeros in GA. Note that these four zeros re consecutive elements in GA and we specify these by their linear index. The VA command sets the scale for the two η's. Finally, the EQ command specifies that the two β 's are equal.

The overall goodness-of-fit measure is

```
                Goodness-of-Fit Statistics

 Degrees of Freedom for (C1)-(C2)                    17
 Maximum Likelihood Ratio Chi-Square (C1)            26.981 (P = 0.0584)
```

A test of the hypothesis $\psi_{21} = 0$ gives $\chi^2 = 0.19$ with one degree of freedom and a test of $\beta_1 = \beta_2$, given $\psi_{21} = 0$, gives $\chi^2 = 0.01$ with one degree of freedom. Hence, it is clear that these hypotheses cannot be rejected. The overall goodness-of-fit of the model with $\psi_{21} = 0$ and $\beta_1 = \beta_2$ is given by $\chi^2 = 26.90$ with 17 degrees of freedom. The probability level is 0.06.

The TSLS estimates and the maximum likelihood estimates with their standard errors are given in the table below.

Table: Estimates for the model with ψ_{21} = 0 and $\beta_1 = \beta_2$

Parameter	TSLS estimates	Unscaled solution (ML)	Standardized solution (ML)
λ_1	1.000	1.000	0.767
λ_2	1.122	1.061 (0.089)	0.813
λ_3	1.120	1.074 (0.081)	0.828
λ_4	1.000	1.000	0.771
β_1	0.210	0.180 (0.039)	0.181
β_2	0.210	0.180 (0.039)	0.179
γ_1	0.156	0.164 (0.039)	0.214
γ_2	0.242	0.254 (0.042)	0.331
γ_3	0.208	0.221 (0.042)	0.288
γ_4	0.072	0.077 (0.041)	0.101
γ_5	0.058	0.068 (0.039)	0.089
γ_6	0.208	0.218 (0.039)	0.283
γ_7	0.314	0.331 (0.041)	0.429
γ_8	0.150	0.152 (0.036)	0.197
ψ_{11}	0.266	0.281 (0.046)	0.478
ψ_{22}	0.220	0.229 (0.039)	0.385
$\theta_{11}^{(\varepsilon)}$	0.443	0.412 (0.051)	0.412
$\theta_{22}^{(\varepsilon)}$	0.299	0.338 (0.052)	0.338
$\theta_{33}^{(\varepsilon)}$	0.283	0.313 (0.046)	0.313
$\theta_{44}^{(\varepsilon)}$	0.428	0.404 (0.046)	0.404

The standardized solution in which η_1 and η_2 are scaled to unit variance is also given. It is seen that the corresponding parameters for the respondent and his best friend are very close. There are good reasons to

suggest that the whole model should be completely symmetric between the respondent and his best friend so that not only $\beta_1 = \beta_2$ but also

$$\lambda_2 = \lambda_3, \gamma_1 = \gamma_8, \gamma_2 = \gamma_7, \gamma_3 = \gamma_6, \gamma_4 = \gamma_5, \psi_{11} = \psi_{22},$$
$$\theta_{11}^{(\varepsilon)} = \theta_{44}^{(\varepsilon)}, \theta_{22}^{(\varepsilon)} = \theta_{33}^{(\varepsilon)}.$$

The overall χ^2 for this model is 30.76 with 25 degrees of freedom. This has a probability level of 0.20. Thus, this model is more parsimonious and has a better fit than the other models.

The command file for this model (**EX55B.LIS**) is the same as the previous one but with the following commands added before the OU command:

```
EQ BE(1,2) BE(2,1)
EQ LY(2,1) LY(3,2)
EQ GA(1,1) GA(2,6)
EQ GA(1,2) GA(2,5)
EQ GA(1,3) GA(2,4)
EQ GA(1,4) GA(2,3)
EQ PS(1) PS(2)
EQ TE(1) TE(4)
EQ TE(2) TE(3)
```

The resulting stability index for the model, and the total and indirect effects of $\boldsymbol{\eta}$ on $\boldsymbol{\eta}$ are given below.

```
  Total Effects of ETA on ETA

            REAMBITN   BFAMBITN
            --------   --------
REAMBITN       0.033      0.185
              (0.015)    (0.043)
               2.215      4.292

BFAMBITN       0.185      0.033
              (0.043)    (0.015)
               4.292      2.215

   Largest Eigenvalue of B*B' (Stability Index) is   0.032

        Indirect Effects of ETA on ETA

            REAMBITN   BFAMBITN
            --------   --------
REAMBITN       0.033      0.006
              (0.015)    (0.004)
               2.215      1.492

BFAMBITN       0.006      0.033
              (0.004)    (0.015)
               1.492      2.215
```