# Two-wave models

LISREL may be useful in analyzing data from longitudinal studies. The characteristic feature of a longitudinal research design is that the same measurements are obtained from the same people at two or more occasions. The purpose of a longitudinal or panel study is to asses the changes that occur between the occasions and to attribute these changes to certain background characteristics and events existing or occurring before the first occasion and/or to various treatments and developments that occur after the first occasion.

Suppose that two variables are used on two occasions, i.e., in a two-wave longitudinal design. Assume that the two variables measure the same latent variable $\eta$ on two different occasions, i.e., $y_1$ and $y_2$ measure $\eta_1$ on the first occasion and $y_3$ and $y_4$ measure $\eta_2$ on the second occasion.

The equations defining the measurement relations are:

$$\begin{aligned} y_1 &= \eta_1 + \varepsilon_1 \\ y_2 &= \lambda_1\eta_1 + \varepsilon_2 \\ y_3 &= \eta_2 + \varepsilon_3 \\ y_4 &= \lambda_2\eta_2 + \varepsilon_4 \end{aligned}$$

The main interest is in the stability of $\eta$ over time. This can be studied by means of the structural relationship

$$\eta_2 = \beta\eta_1 + \zeta.$$

In particular, one is interested in whether $\beta$ is close to one and $\zeta$ is small.

Let $\mathbf{\Omega}$ be the covariance matrix of $(\eta_1, \eta_2)$ and let $\mathbf{\Theta}$ be the covariance matrix of $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$. If all the $\varepsilon$'s are uncorrelated, so that $\mathbf{\Theta}$ is diagonal, the covariance matrix of $(y_1, y_2, y_3, y_4)$ is

$$\mathbf{\Sigma} = \begin{bmatrix} \omega_{11}+\theta_{11} & & & \\ \lambda_1\,\omega_{11} & \lambda_1^2\omega_{11}+\theta_{22} & & \\ \omega_{21} & \lambda_1\,\omega_{21} & \omega_{22}+\theta_{33} & \\ \lambda_2\omega_{21} & \lambda_1\,\lambda_2\omega_{21} & \lambda_2\omega_{22} & \lambda_2^2\omega_{22}+\theta_{44} \end{bmatrix}.$$

The matrix $\mathbf{\Sigma}$ has 10 variances and covariances which are functions of 9 parameters. It is readily verified that all 9 parameters are identified so the model has one degree of freedom.

Often when the same variables are used repeatedly, there is a tendency for the corresponding errors (the $\varepsilon$ 's) to correlated over time because of memory or other retest effects. Hence there is a need to generalize the above model to allow for correlations between $\varepsilon_1$ and $\varepsilon_3$ and also between $\varepsilon_2$ and $\varepsilon_4$. This means that there will be two non-zero covariances $\theta_{31}$ and $\theta_{42}$ in $\mathbf{\Theta}$. The covariance matrix of the observed variables changes to

$$\mathbf{\Sigma} = \begin{bmatrix} \omega_{11}+\theta_{11} & & & \\ \lambda_1\,\omega_{11} & \lambda_1^2\omega_{11}+\theta_{22} & & \\ \omega_{21}+\theta_{31} & \lambda_1\,\omega_{21} & \omega_{22}+\theta_{33} & \\ \lambda_2\omega_{21} & \lambda_1\,\lambda_2\omega_{21}+\theta_{42} & \lambda_2\omega_{22} & \lambda_2^2\omega_{22}+\theta_{44} \end{bmatrix}.$$

This $\mathbf{\Sigma}$ has its 10 independent elements expressed in terms of 11 parameters. Hence it is clear that the model is not identified. In fact, none of the 11 parameters is identified without further conditions imposed. The loadings $\lambda_1$ and $\lambda_2$ may be multiplied by a constant and the $\omega$ 's divided by the same constant. This does not change $\sigma_{21}$, $\sigma_{32}$, $\sigma_{41}$ and $\sigma_{43}$. The change in the other $\sigma$ 's may be compensated by adjusting the $\theta$ 's additively. Hence to make the model identified one must fix one $\lambda$ or one $\omega$ at a non-zero value and one $\theta$ at some arbitrary value. However, the *correlation* between $\eta_1$ and $\eta_2$ is identified without any restrictions, since

$$Corr\left(\eta_1,\eta_2\right) = \left(\omega_{21}^2 / \omega_{11}\omega_{22}\right)^{1/2} = \left[\left(\sigma_{32}\sigma_{41}\right) / \left(\sigma_{21}\sigma_{43}\right)\right]^{1/2}.$$

The model may therefore be used to estimate this correlation coefficient and to test whether this is one. The maximum likelihood estimate of the correlation coefficient is $\left[\left(s_{32}s_{41}\right)/\left(s_{21}s_{43}\right)\right]^{1/2}$. To make further use of the model it is necessary to make some assumption about the nature of the variables. For example, if it can be assumed that the two variables on each occasion are tau-equivalent, we can set both $\lambda_1$ and $\lambda_2$ equal to one. Then the model can be estimated and tested with one degree of freedom. If $\lambda_1 = \lambda_2$ the model is just identified.

While the above model is not identified as it stands, it becomes so as soon as there is information about one or more background variables affecting $\eta_1$ or $\eta_2$ or both.

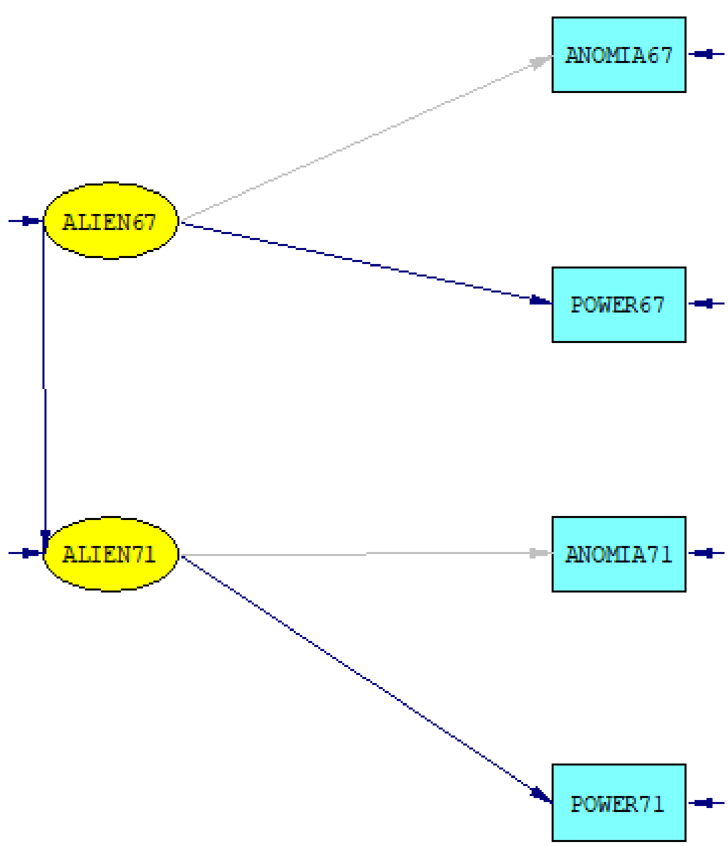
In this example (Wheaton, et al. (1977)) data on attitude scales were collected from 932 persons in two rural regions in Illinois at three points in time: 1966, 1967, and 1971. The variables used for the present example are the Anomia subscale and the Powerlessness subscale, taken to be indicators of Alienation. This example uses data from 1967 and 1971 only. The background variables are the respondent's education (years of schooling completed) and Duncan's Socioeconomic Index (SEI). These are taken to be indicators of the respondent's socioeconomic status (SES). The sample covariance matrix of the six observed variables is given in the table below.

**Table: Covariance matrix for variables in the stability of alienation example**

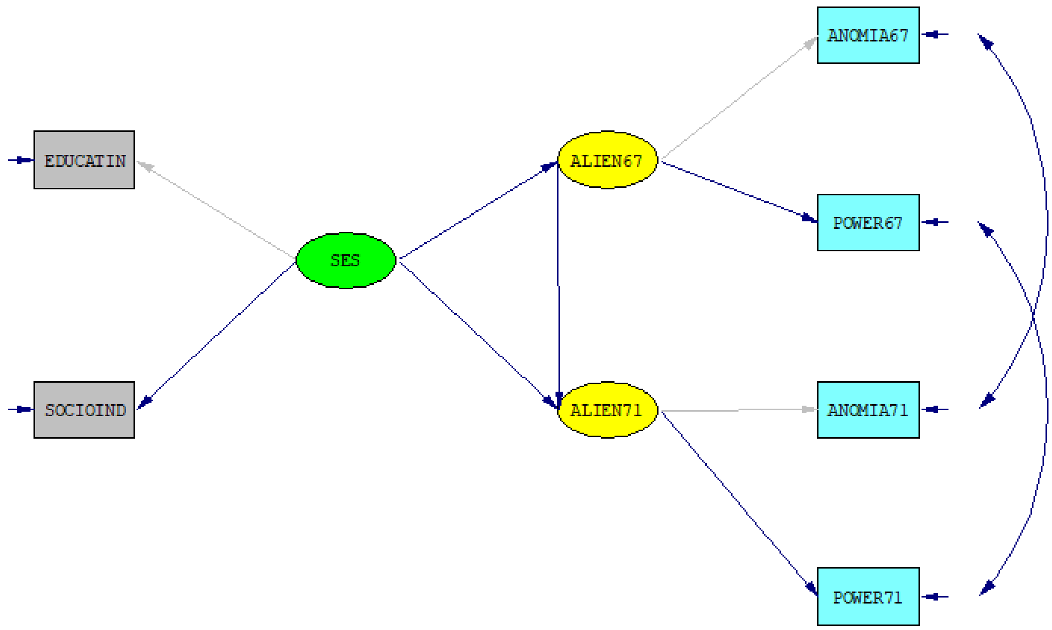
| | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $x_1$ | $x_2$ |
|---|---|---|---|---|---|---|
| ANOMIA67 | 11.834 | | | | | |
| POWERL67 | 6.947 | 9.364 | | | | |
| ANOMIA71 | 6.819 | 5.091 | 12.532 | | | |
| POWERL71 | 4.783 | 5.028 | 7.495 | 9.986 | | |
| EDUCATIN | -3.839 | -3.889 | -3.841 | -3.625 | 9.610 | |
| SOCIOIND* | -2.190 | -1.883 | -2.175 | -1.878 | 3.552 | 4.503 |

Four models will be considered:

- **Model A:**



- **Model D:**




- **Model C is equal to D with** $\theta_{42}^{(\varepsilon)} = 0$**.**
- **Model B is equal to C with** $\theta_{31}^{(\varepsilon)} = 0$ **.**

The variables in the model are:

$y_1$ = ANOMIA 67
$y_2$ = POWERLESSNESS 67
$y_3$ = ANOMIA 71
$y_4$ = POWERLESSNESS 71
$x_1$ = EDUCATION
$x_2$ = SEI
$\xi$ = SES
$\eta_1$ = ALIENATION 67
$\eta_2$ = ALIENATION 71

In the first model we use only the $y$'s and $\eta$'s with all $\varepsilon$'s uncorrelated. This is a Submodel 3A with $\mathbf{B} = \mathbf{0}$. The command file (**EX64A.LIS** in the **LISREL Examples** folder) is:

```
Stability of Alienation, Model A (uncorrelated error terms)
DA NI=6 NO=932
LA
ANOMIA67 POWER67 ANOMIA71 POWER71 EDUCATIN SOCIOIND
CM FI=EX64.COV
```

```
SE
1 2 3 4 /
MO NY=4 NE=2 BE=SD TE=SY
LE
ALIEN67 ALIEN71
FR LY(2,1) LY(4,2)
VA 1 LY(1,1) LY(3,2)
OU
```

The file **EX64.COV** contains a covariance matrix for 6 variables but only four of them are used in Model A. No SE command is necessary, however, because the four variables in the model are the *first four* variables in the file. The specification TE = SY on the MO command does the same thing as if TE is the default. The only difference is that when TE = SY, the entire lower half of $\mathbf{\Theta}_\varepsilon$ is stored. The off-diagonal elements of $\mathbf{\Theta}_\varepsilon$ are still fixed zeroes but, with TE = SY, any such element can be declared free. Also, with this specification, one can get modification indices for the off-diagonal elements of $\mathbf{\Theta}_\varepsilon$ which is important in this example.

The results are summarized in the second column of the table below. The overall $\chi^2$ is 61.17 with 1 degree of freedom. The model suffers from two kinds of specification errors: there is bias in $\beta$ due to omitted variables and the error terms are correlated for the same variables. The modification indices for $\mathbf{\Theta}_\varepsilon$ indicate that $\varepsilon_1$, $\varepsilon_2$ and $\varepsilon_3$, $\varepsilon_4$ should be correlated between sets but not within sets.

```
Modification Indices and Expected Change

 No Non-Zero Modification Indices for LAMBDA-Y

 No Non-Zero Modification Indices for BETA

 No Non-Zero Modification Indices for PSI

         Modification Indices for THETA-EPS

            ANOMIA67    POWER67   ANOMIA71    POWER71
            --------   --------   --------   --------
 ANOMIA67       - -
  POWER67       - -        - -
 ANOMIA71     59.14      59.14        - -
  POWER71     59.14      59.14        - -        - -

         Expected Change for THETA-EPS

            ANOMIA67    POWER67   ANOMIA71    POWER71
            --------   --------   --------   --------
 ANOMIA67       - -
  POWER67       - -        - -
 ANOMIA71      2.22      -1.88        - -
  POWER71     -1.81       1.54        - -        - -

 Maximum Modification Index is   59.14 for Element ( 4, 2) of THETA-EPS
```

However, as the model has only one degree of freedom, only one parameter can be relaxed. The modification indices show that any of the four correlations can be relaxed yielding a model with perfect fit. This is an example of a case when the modification indices reveal several equivalent models.

**Table: Maximum likelihood estimates for models A – D (standard errors in parentheses)**

| Parameter | Model A | Model B | Model C | Model D |
|---|---|---|---|---|
| $\lambda_1$ | 0.85 (0.04) | 0.89 (0.04) | 1.03 (0.05) | 0.98 (0.06) |
| $\lambda_2$ | 0.82 (0.04) | 0.85 (0.04) | 0.97 (0.05) | 0.92 (0.06) |
| $\lambda_3$ | | 0.53 (0.04) | 0.52 (0.04) | 0.52 (0.04) |
| $\beta$ | 0.79 (0.04) | 0.71 (0.05) | 0.62 (0.05) | 0.61 (0.05) |
| $\gamma_1$ | | -0.61 (0.06) | -0.55 (0.05) | -0.58 (0.06) |
| $\gamma_2$ | | -0.17 (0.05) | -0.21 (0.05) | -0.23 (0.05) |
| $\phi$ | | 6.67 (0.64) | 6.88 (0.66) | 6.80 (0.65) |
| $\psi_{11}$ | 8.20 (0.62) | 5.31 (0.47) | 4.71 (0.43) | 4.85 (0.47) |
| $\psi_{22}$ | 4.09 (0.43) | 3.74 (0.39) | 3.87 (0.34) | 4.09 (0.40) |
| $\theta_{11}^{(\varepsilon)}$ | 3.63 (0.37) | 4.02 (0.34) | 5.07 (0.37) | 4.74 (0.45) |
| $\theta_{22}^{(\varepsilon)}$ | 3.48 (0.29) | 3.19 (0.27) | 2.22 (0.32) | 2.57 (0.40) |
| $\theta_{33}^{(\varepsilon)}$ | 3.34 (0.40) | 3.70 (0.37) | 4.81 (0.40) | 4.40 (0.52) |
| $\theta_{44}^{(\varepsilon)}$ | 3.88 (0.30 | 3.63 (0.29) | 2.68 (0.33) | 3.07 (0.43) |
| $\theta_{31}^{(\varepsilon)}$ | | | 1.89 (0.24) | 1.62 (0.31) |
| $\theta_{42}^{(\varepsilon)}$ | | | | 0.34 (0.26) |
| $\theta_{11}^{(\delta)}$ | | 2.95 (0.50) | 2.73 (0.52) | 2.81 (0.51) |
| $\theta_{22}^{(\delta)}$ | | 2.61 (0.18) | 2.67 (0.18) | 2.65 (0.18) |
| $\chi^2$ | 61.17 | 71.55 | 6.34 | 4.74 |
| *df* | 1 | 6 | 5 | 4 |

To deal with the omitted variables bias, one must include the education measures in the model. Consider first Model B.

This model is specified as:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \lambda_1 & 0 \\ 0 & 1 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda_3 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}$$

$$\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \beta & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} \xi + \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}$$

It is assumed that $\zeta_1$ and $\zeta_2$ are uncorrelated. The scales for $\eta_1$, $\eta_2$ and $\xi$ have been chosen to be the same as for $y_1$, $y_3$, and $x_1$, respectively. In Model B all four $\varepsilon$-terms are uncorrelated, whereas in Model C $\varepsilon_1$ and $\varepsilon_3$ are correlated, and in Model D $\varepsilon_2$ and $\varepsilon_4$ are correlated also.

Consider first the identification of Model B. Let $\phi = Var(\xi)$. We have six observed variables with 21 variances and covariances. Model B has 15 parameters (3 $\lambda$'s, 1 $\beta$, 2 $\gamma$'s, 1 $\phi$, 2 $\psi$'s, and 6 $\theta$'s), so that, *if* all these are identified, the model will have 6 degrees of freedom. The structural equations are:

$$\eta_1 = \gamma_1 \xi + \zeta_1$$
$$\eta_2 = \beta \eta_1 + \gamma_2 \xi + \zeta_2$$

with reduced form

$$\eta_1 = \gamma_1 \xi + \zeta_1$$
$$\eta_2 = (\lambda_2 + \beta \gamma_1)\xi + (\zeta_2 + \beta \zeta_1) = \pi \xi + \upsilon.$$

Hence

$$Cov(y_1, x_1) = Cov(\eta_1, x_1) = \gamma_1 \phi$$
$$Cov(y_2, x_1) = \lambda_1 Cov(\eta_1, x_1) = \lambda_1 \gamma_1 \phi$$
$$Cov(y_3, x_1) = Cov(\eta_2, x_1) = \pi \phi$$
$$Cov(y_4, x_1) = \lambda_2 Cov(\eta_2, x_1) = \lambda_2 \pi \phi$$

Of we use $x_2$ instead of $x_1$ in these equations, all four right sides will be multiplied by $\lambda_3$. Thus, $\lambda_3$ is overdetermined, since

$$\lambda_3 = Cov(y_i, x_2) / Cov(y_i, x_1) \quad i = 1, 2, ..., 4.$$

With $\lambda_3$ determined, $\phi$ is determined by

$$Cov(x_1, x_2) = \lambda_3 \phi.$$

With $\phi$ determined, the above equations determine $\gamma_1$, $\lambda_1$, $\pi$ and $\lambda_2$, respectively. Furthermore,

$$Cov(y_1, y_2) = \lambda_1 Var(\eta_1) = \lambda_1 \left(\gamma_1^2 \phi + \psi_{11}\right),$$

which determines $\psi_{11}$, and

$$Cov(y_3, y_4) = \lambda_2 Var(\eta_2) = \lambda_2 \left(\pi^2 \phi + Var(\upsilon)\right),$$

which determines

$$Var(\upsilon) = \psi_{22} + \beta^2 \psi_{11}.$$

For given $\lambda_1$, $\lambda_2$, $\gamma_1$, $\pi$, $\phi$, and $\lambda\psi_{11}$ the four equations

$$\begin{aligned}
Cov(y_1, y_3) &= \gamma_1 \pi \phi + \beta \psi_{11} \\
Cov(y_1, y_4) &= \lambda_2 \left(\gamma_1 \pi \phi + \beta \psi_{11}\right) \\
Cov(y_2, y_3) &= \lambda_1 \left(\gamma_1 \pi \phi + \beta \psi_{11}\right) \\
Cov(y_2, y_4) &= \lambda_1 \lambda_2 \left(\gamma_1 \pi \phi + \beta \psi_{11}\right)
\end{aligned}$$

show that $\beta$ is overdetermined. Then, with $\beta$ determined, $\gamma_2 = \pi - \beta\gamma_1$ and $\psi_{22}$ are obtained. The error variances $\theta_{ii}^{(\delta)}$ are determined from $Var(y_i)$, $i$ = 1, 2, 3, 4 and $\theta_{ii}^{(\delta)}$ from $Var(x_i)$, $i$ = 1, 2. Hence it is clear that Model B is identified and has six independent restrictions on $\mathbf{\Sigma}$.

In Model D there are two more parameters, namely $\theta_{31}^{(\varepsilon)}$ and $\theta_{42}^{(\varepsilon)}$. By substitution in the four equations above, it can be shown that Model D is also identified and has four degrees of freedom.

The command file for Model B is (**EX64B.LIS**):

```
Stability of Alienation, Model B (Uncorrelated Errors)
DA NI=6 NO=932
LA
ANOMIA67 POWER67 ANOMIA71 POWER71 EDUCATIN SOCIOIND
CM FI=EX64.COV
MO NY=4 NX=2 NE=2 NK=1 BE=SD TE=SY
LE
ALIEN67 ALIEN71
LK
SES
FR LY(2,1) LY(4,2) LX(2,1)
VA 1 LY(1,1) LY(3,2) LX(1,1)
OU
```

The model includes all eight parameter matrices but only two need to be declared on the MO command: **B** is subdiagonal and $\mathbf{\Psi}$ is diagonal. As before, in order to see the modification indices for the off-diagonal elements of $\mathbf{\Theta}_\varepsilon$ we also include the specification TE = SY. The free parameters $\lambda_1$, $\lambda_2$ and $\lambda_3$ in $\mathbf{\Lambda}_y$ and $\mathbf{\Lambda}_x$ must be declared free by a FR command. One element in each column of $\mathbf{\Lambda}_y$ and $\mathbf{\Lambda}_x$ is assigned the value one to fix the scales for $\eta_1$, $\eta_2$, and $\xi$. Note that neither $\xi$ nor $\eta_1$ or $\eta_2$ are standardized in this example.

The value of $\chi^2$ for this model is 71.55 with six degrees of freedom. This is not considered an acceptable fit. As in Model A, the modification indices for $\theta_{31}^{(\varepsilon)}$ and $\theta_{42}^{(\varepsilon)}$ are large:

```
      Modification Indices for THETA-EPS

            ANOMIA67    POWER67   ANOMIA71    POWER71
            --------   --------   --------   --------
 ANOMIA67       - -
  POWER67       - -        - -
 ANOMIA71     63.71      49.75        - -
  POWER71     49.83      37.26        - -        - -
```

As in many other longitudinal studies, where the same measures are repeated over time, there is a tendency for the measurement errors in these measures to correlate over time due to memory or other retest effects. This suggests that the most likely improvement of the model is obtained by freeing the elements $\theta_{31}^{(\varepsilon)}$ and $\theta_{42}^{(\varepsilon)}$ of $\mathbf{\Theta}_\varepsilon$. The largest modification index is 63.71 for element $\theta_{31}^{(\varepsilon)}$ of $\mathbf{\Theta}_\varepsilon$, predicting a drop in $\chi^2$ of about 63.71 if $\theta_{31}^{(\varepsilon)}$ is relaxed. This can be verified by running the model again adding TE(3,1) on the FR command. This is Model C (**EX64C.LIS**). The $\chi^2$ for this modified model is 6.34 with 5 degrees of freedom. The drop in $\chi^2$ from Model B to Model C is 65.14 with one degree of freedom, which is about what the modification index predicted. Model C fits quite well. For Model C, the largest modification index, 1.59, now occurs for the element $\theta_{42}^{(\varepsilon)}$ but this is not significant. Thus, in this example, it seems that there is strong autocorrelation in the measurement error of ANOMIA only. The memory or retest effect in POWERLESSNESS seems to be much weaker. All results for Models B, C, and D are given in the table of results shown previously.

For this example (Model C) it may be instructive to examine the sections of the output called COVARIANCE and TOTAL AND INDIRECT EFFECTS.

The total effect of SES on Alienation 71 is almost equal to the direct effect of SES on Alienation 67, although the direct effect of SES on Alienation 71 is much smaller. The effects of SES on Alienation are negative, indicating that Alienation decreases when SES increases. Also shown in the section of TOTAL EFFECTS are the total effects of SES on the observed *y*-measures and also the total effects of $\eta_1$ and $\eta_2$ on these observed measures. Although, according to the model, SES does not have a direct effect on any observed *y*, there are negative indirect effects via $\eta_1$ and $\eta_2$. Similarly, although $\eta_1$ does not have a direct effect on $y_3$ and $y_4$, $\eta_1$ affects $y_3$ and $y_4$ indirectly via $\eta_2$.