Problems on Structural Equation Models: Answers

John Fox

IQS Barcelona
January 2016

1.

4.1 (a) Recursive

$$\begin{aligned} y_3 &= \gamma_{31}x_1 + \gamma_{32}x_2 + \varepsilon_6 \\ y_4 &= \gamma_{41}x_1 + \gamma_{42}x_2 + \varepsilon_7 \\ y_5 &= \beta_{53}y_3 + \beta_{54}y_4 + \varepsilon_8 \end{aligned}$$

(b) Nonrecursive

$$\begin{aligned} y_3 &= \gamma_{31}x_1 + \beta_{34}y_4 + \varepsilon_6 \\ y_4 &= \gamma_{42}x_2 + \beta_{43}y_3 + \varepsilon_7 \\ y_5 &= \beta_{53}y_3 + \beta_{54}y_4 + \varepsilon_8 \end{aligned}$$

(c) Block recursive

$$\begin{aligned} y_3 &= \gamma_{31}x_1 + \beta_{34}y_4 + \varepsilon_6 \\ y_4 &= \gamma_{42}x_2 + \beta_{43}y_4 + \varepsilon_7 \\ y_5 &= \gamma_{51}x_1 + \gamma_{52}x_2 + \beta_{53}y_3 + \beta_{54}y_4 + \varepsilon_8 \end{aligned}$$

(d) Block recursive

$$\begin{aligned} y_3 &= \gamma_{31}x_1 + \gamma_{32}x_2 + \beta_{34}y_4 + \varepsilon_6 \\ y_4 &= \gamma_{41}x_1 + \gamma_{42}x_2 + \beta_{43}y_4 + \varepsilon_7 \\ y_5 &= \beta_{53}y_3 + \beta_{54}y_4 + \varepsilon_8 \end{aligned}$$

(e) Nonrecursive

$$\begin{aligned} y_2 &= \gamma_{21}x_1 + \varepsilon_4 \\ y_3 &= \gamma_{31}x_1 + \beta_{32}y_2 + \varepsilon_5 \end{aligned}$$

(f) Recursive

$$\begin{aligned} y_2 &= \gamma_{21}x_1 + \varepsilon_4 \\ y_3 &= \gamma_{31}x_1 + \beta_{32}y_2 + \varepsilon_5 \end{aligned}$$

(g) Nonrecursive

$$\begin{aligned} y_{10} &= \gamma_{10,1}x_1 + \gamma_{10,2}x_2 + \gamma_{10,3}x_3 + \gamma_{10,4}x_4 + \gamma_{10,5}x_5 + \gamma_{10,6}x_6 + \gamma_{10,7}x_7 + \gamma_{10,8}x_8 + \beta_{10,11}y_{11} + \varepsilon_{12} \\ y_{11} &= \gamma_{11,2}x_2 + \gamma_{11,3}x_3 + \gamma_{11,4}x_4 + \gamma_{11,5}x_5 + \gamma_{11,6}x_6 + \gamma_{11,7}x_7 + \gamma_{10,8}x_8 + \gamma_{11,9}x_9 + \beta_{11,10}y_{10} + \varepsilon_{13} \end{aligned}$$

(j) Recursive

$$\begin{aligned} y_5 &= \gamma_{52}x_2 + \gamma_{53}x_3 + \gamma_{54}x_4 + \varepsilon_8 \\ y_6 &= \gamma_{61}x_1 + \gamma_{62}x_2 + \gamma_{63}x_3 + \beta_{65}y_5 + \varepsilon_9 \\ y_7 &= \gamma_{71}x_1 + \gamma_{72}x_2 + \gamma_{73}x_3 + \beta_{76}y_6 + \varepsilon_{10} \end{aligned}$$

4.2 The general causal structure of the model seems reasonable to me. It's always possible in observational data to imagine that the "exogenous" variables are actually correlated with the error, but the burden should be on a critic to indicate what omitted variables are likely to make this true. The model could be more articulated, treating some of the exogenous variables as endogenous, perhaps in a block-recursvive structure, but this doesn't invalidate the model as specified. I wonder about the distribution of the endogenous variables, about the linearity of effects, and about possible interactions (e.g., of other explanatory variables with race), but without access to the original data, it's impossible to know whether these concerns are valid.

4.5 Here, too, the general causal structure of the model seems reasonable, but the assumption that the errors are uncorrelated does not: I expect that the omitted causes of the endogenous variables are similar. As well, all of the endogenous variables are counts, which are probably positively skewed. Although one can't know without checking the data, assuming normally distributed errors here is probably not a good idea.

2.

(a) IV estimating equations:

$$\begin{aligned} s_{z_1y} &= s_{z_1x_1}\widehat{\beta}_1 + s_{z_1x_2}\widehat{\beta}_2 \\ s_{z_2y} &= s_{z_2x_1}\widehat{\beta}_1 + s_{z_2x_2}\widehat{\beta}_2 \\ s_{z_3y} &= s_{z_3x_1}\widehat{\beta}_1 + s_{z_3x_2}\widehat{\beta}_2 \end{aligned}$$

(b) We have three estimating equations but only two unknown parameters; generally, the estimating equations will be inconsistent — there will be no pair of values $\widehat{\beta}_1, \widehat{\beta}_2$ that satisfies all three simultaneously.

(c) There are several ways to proceed. One simple way would be to arbitrarily get rid of one of the IVs. After all, we have more IVs than we need, and any two will give us consistent estimates of β_1 and β_2. (Looking ahead, we could estimate the equation by a method like 2SLS.)

3.

4.8 **(a)** overidentified

(b) just-identified

(c) just-identified

(d) underidentified

(e) underidentified

(f) just-identified

(g) just-identified

(j) overidentified

4.10 (a) OLS

(b) IV (or equivalently, 2SLS or FIML)

(c) IV (or equivalently, 2SLS or FIML), taking account of the block-recursive structure (so x_3 and x_4 can be used as IVs in estimating for the equation for y_5)

(d) The model cannot be estimated (although the structural equation for y_5 could be estimated by 2SLS).

(e) The model cannot be estimated (although the structural equation for y_2 could be estimated by OLS).

(f) OLS

(g) IV (or equivalently, 2SLS or FIML)

(j) OLS

4.11 Because this is a just-identified model, IV, 2SLS, and FIML estimates will coincide. Estimates by FIML from the `sem` function:

```
> library(sem)
> source("Rindfuss.R") # input covariances
> Rindfuss

            FatherOcc   Black    Sibs    Farm   South Parents Catholic Smoking
FatherOcc    456.6769  0.0000  0.0000  0.0000  0.0000  0.0000   0.0000  0.0000
Black         -0.9201  0.0894  0.0000  0.0000  0.0000  0.0000   0.0000  0.0000
Sibs         -15.8253  0.1416  9.2112  0.0000  0.0000  0.0000   0.0000  0.0000
Farm          -3.2442  0.0124  0.3908  0.2209  0.0000  0.0000   0.0000  0.0000
South         -1.3205  0.0451  0.2181  0.0491  0.2294  0.0000   0.0000  0.0000
Parents       -0.4631  0.0174 -0.0458 -0.0055  0.0132  0.1498   0.0000  0.0000
Catholic       0.4768 -0.0191  0.0179 -0.0295 -0.0589 -0.0085   0.1772  0.0000
Smoking       -0.3143  0.0031  0.0291 -0.0096 -0.0018  0.0089  -0.0014  0.1170
Miscarriage    0.2356  0.0031  0.0018 -0.0045 -0.0039  0.0021  -0.0003  0.0009
Education     18.6603 -0.1567 -2.3493 -0.2052 -0.2385 -0.1434  -0.0119 -0.1380
FirstBirth    16.2133 -0.2305 -1.4237 -0.2262 -0.3458  0.1752   0.1683 -0.1702
            Miscarriage Education FirstBirth
FatherOcc        0.0000    0.0000     0.0000
Black            0.0000    0.0000     0.0000
Sibs             0.0000    0.0000     0.0000
Farm             0.0000    0.0000     0.0000
South            0.0000    0.0000     0.0000
Parents          0.0000    0.0000     0.0000
Catholic         0.0000    0.0000     0.0000
Smoking          0.0000    0.0000     0.0000
Miscarriage      0.0888    0.0000     0.0000
Education        0.0267    5.5696     0.0000
FirstBirth       0.2626    3.6580    16.6382
```

```
> Rindfuss.mod <- specifyEquations()
> Education = gamma10.1*FatherOcc + gamma10.2*Black + gamma10.3*Sibs
>            + gamma10.4*Farm + gamma10.5*South + gamma10.6*Parents
>            + gamma10.7*Catholic + gamma10.8*Smoking + beta10.11*FirstBirth
> FirstBirth = gamma11.2*Black + gamma11.3*Sibs + gamma11.4*Farm
>            + gamma11.5*South + gamma11.6*Parents + gamma11.7*Catholic
```

```
>             + gamma11.8*Smoking + gamma11.9*Miscarriage + beta11.10*Education
> V(Education) = sigma10.10
> V(FirstBirth) = sigma11.11
> C(Education, FirstBirth) = sigma10.11
```

```
> Rindfuss.sem <- sem(Rindfuss.mod, S=Rindfuss, N=1766,
+     fixed.x=c("FatherOcc", "Black", "Sibs", "Farm", "South",
+         "Parents", "Catholic", "Smoking", "Miscarriage"))
> summary(Rindfuss.sem)

 Model Chisquare =  1.770604e-08   Df =  0 Pr(>Chisq) = NA
 AIC =  42
 BIC =  1.770604e-08

 Normalized Residuals
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-6.079e-05  0.000e+00  0.000e+00 -6.827e-07  0.000e+00  7.069e-05

 R-square for Endogenous Variables
 Education FirstBirth
    0.3287     0.2379

 Parameter Estimates
           Estimate    Std Error  z value     Pr(>|z|)
gamma10.1   0.02744160 0.00272080  10.0858571  6.380588e-24
gamma10.2  -0.60434091 0.19601105  -3.0831982  2.047887e-03
gamma10.3  -0.16839989 0.01637242 -10.2855806  8.184729e-25
gamma10.4  -0.11695289 0.10931218  -1.0698981  2.846652e-01
gamma10.5  -0.53864909 0.11699799  -4.6039176  4.146165e-06
gamma10.6  -0.88684193 0.14973897  -5.9225859  3.169182e-09
gamma10.7  -0.51774211 0.11763161  -4.4013860  1.075616e-05
gamma10.8  -0.88105478 0.15680937  -5.6186361  1.924707e-08
beta10.11   0.08486083 0.05344050   1.5879497  1.122977e-01
gamma11.2  -1.43425783 0.33504480  -4.2807942  1.862275e-05
gamma11.3   0.10336705 0.04186307   2.4691705  1.354267e-02
gamma11.4  -0.08506349 0.20604793  -0.4128335  6.797286e-01
gamma11.5  -0.30836638 0.21727719  -1.4192304  1.558319e-01
gamma11.6   2.24255222 0.25387010   8.8334635  1.014835e-18
gamma11.7   0.83142593 0.22459270   3.7019276  2.139677e-04
gamma11.8  -0.64618994 0.29594574  -2.1834744  2.900089e-02
gamma11.9   2.69145990 0.28863119   9.3249101  1.110778e-20
beta11.10   0.83874364 0.14529629   5.7726432  7.803759e-09
sigma10.10  3.73892279 0.18762783  19.9273362  2.357823e-88
sigma11.11 12.67942187 0.50300729  25.2072329 3.337186e-140
sigma10.11 -1.92672838 0.86828146  -2.2190136  2.648580e-02

gamma10.1  Education <--- FatherOcc
gamma10.2  Education <--- Black
gamma10.3  Education <--- Sibs
gamma10.4  Education <--- Farm
gamma10.5  Education <--- South
gamma10.6  Education <--- Parents
gamma10.7  Education <--- Catholic
```

```
gamma10.8  Education <--- Smoking
beta10.11  Education <--- FirstBirth
gamma11.2  FirstBirth <--- Black
gamma11.3  FirstBirth <--- Sibs
gamma11.4  FirstBirth <--- Farm
gamma11.5  FirstBirth <--- South
gamma11.6  FirstBirth <--- Parents
gamma11.7  FirstBirth <--- Catholic
gamma11.8  FirstBirth <--- Smoking
gamma11.9  FirstBirth <--- Miscarriage
beta11.10  FirstBirth <--- Education
sigma10.10 Education <--> Education
sigma11.11 FirstBirth <--> FirstBirth
sigma10.11 FirstBirth <--> Education

 Iterations =  54
```

The reciprocal effects between education and age at first birth are primarily what are of interest here. The coefficient for the direct effect of age at first birth on education at marriage is (surprisingly) positive, but non-significant. The coefficient for the direct effect of education on age at first birth is positive (as expected) and highly statistically significant. The estimated correlation between the errors is

$$r_{10,11} = \frac{-1.927}{\sqrt{3.739 \times 12.679}} = -0.280$$

This isn't a very large correlation, but I would have expected it to be positive (and the negative disturbance covariance is statistically significant), so there is a suggestion that the model may be misspecified.

4.14 Because this is a recursive model, OLS and FIML estimates coincide. Estimates by FIML from the `sem` function:

```
> source("Lincoln.R")
> Lincoln

           UnionStaff Employment logWorkers Unionized  Strikes Strikers
UnionStaff   0.007744   0.000000   0.000000  0.000000 0.000000 0.000000
Employment   0.000635   0.000400   0.000000  0.000000 0.000000 0.000000
logWorkers   0.052401   0.005077   1.065024  0.000000 0.000000 0.000000
Unionized    0.006624   0.001471   0.066069  0.037636 0.000000 0.000000
Strikes      0.054564   0.012024   0.823108  0.137249 1.809025 0.000000
Strikers     0.084675   0.015990   1.131609  0.171958 2.025220 2.496400
PersonDays   0.103616   0.019572   1.325756  0.184820 1.969703 2.567911
           PersonDays
UnionStaff   0.000000
Employment   0.000000
logWorkers   0.000000
Unionized    0.000000
Strikes      0.000000
Strikers     0.000000
PersonDays   2.989441
```

```
> Lincoln.mod <- specifyEquations()
> Strikes = gamma52*Employment + gamma53*logWorkers + gamma54*Unionized
> Strikers = gamma61*UnionStaff + gamma62*Employment + gamma63*logWorkers
>            + beta65*Strikes
> PersonDays = gamma71*UnionStaff + gamma72*Employment + gamma73*logWorkers
>             + beta76*Strikers
```

```
NOTE: adding 3 variances to the model
```

```
> Lincoln.sem <- sem(Lincoln.mod, S=Lincoln, N=78,
+     fixed.x=c("UnionStaff", "Employment", "logWorkers", "Unionized"))
> summary(Lincoln.sem)

 Model Chisquare =  7.446498   Df =  4 Pr(>Chisq) = 0.1140918
 AIC =  35.4465
 BIC =  -9.980337

 Normalized Residuals
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-0.10520  0.00000  0.00000  0.02334  0.01595  0.24860

 R-square for Endogenous Variables
   Strikes   Strikers PersonDays
    0.5174     0.9564     0.9239

 Parameter Estimates
              Estimate   Std Error  z value   Pr(>|z|)
gamma52       15.2681485 5.80957526  2.628101  8.586311e-03
gamma53        0.5732958 0.11036615  5.194490  2.052823e-07
gamma54        2.0435878 0.61497411  3.323047  8.903993e-04
gamma61        2.7045417 0.54488194  4.963537  6.922085e-07
gamma62        5.8458431 2.16005186  2.706344  6.802857e-03
gamma63        0.1996484 0.05050657  3.952919  7.720372e-05
beta65         0.9082389 0.03766839 24.111431 1.896764e-128
gamma71        2.3615937 0.81105338  2.911761  3.593974e-03
gamma72       11.8012534 3.22542845  3.658817  2.533819e-04
gamma73        0.2794016 0.07773259  3.594395  3.251468e-04
beta76         0.7463017 0.05633287 13.248069  4.629613e-40
V[Strikes]     0.8730760 0.14070894  6.204837  5.475368e-10
V[Strikers]    0.1086105 0.01750417  6.204837  5.475368e-10
V[PersonDays]  0.2269132 0.03657037  6.204837  5.475368e-10

gamma52       Strikes <--- Employment
gamma53       Strikes <--- logWorkers
gamma54       Strikes <--- Unionized
gamma61       Strikers <--- UnionStaff
gamma62       Strikers <--- Employment
gamma63       Strikers <--- logWorkers
beta65        Strikers <--- Strikes
gamma71       PersonDays <--- UnionStaff
gamma72       PersonDays <--- Employment
gamma73       PersonDays <--- logWorkers
```

```
beta76        PersonDays <--- Strikers
V[Strikes]    Strikes <--> Strikes
V[Strikers]   Strikers <--> Strikers
V[PersonDays] PersonDays <--> PersonDays

 Iterations =  0
```

Because this model is overidentified, it was possible that it would not do a good job of reproducing the observed correlational structure of the data, but the overidentification test (labelled `Model Chisquare` in the output) is nonsignificant. Despite the small sample size, all of the specified structural parameters are statistically significant. Not surprisingly, given the nature of the three endogenous variables (numbers of strikes, strikers, and person-days lost to strikes), the coefficients $\widehat{\beta}_{65}$ and $\widehat{\beta}_{76}$ are particularly highly significant.

4.

(a) Label the variables as follows

x_1	Education
x_2	SEI
y_1	Anomia67
y_2	Powerless67
y_3	Anomia71
y_4	Powerless71
ξ_1	SES
η_1	Alienation67
η_2	Alienation71

structural submodel:

$$\begin{aligned} \eta_1 &= \gamma_{11}\xi_1 + \zeta_1 \\ \eta_2 &= \gamma_{21}\xi_1 + \beta_{21}\eta_1 + \zeta_2 \end{aligned}$$

measurement submodel:

$$\begin{aligned} x_1 &= \xi_1 + \delta_1 \\ x_2 &= \lambda^x_{21}\xi_1 + \delta_2 \\ y_1 &= \eta_1 + \varepsilon_1 \\ y_2 &= \lambda^y_{21}\eta_1 + \varepsilon_2 \\ y_3 &= \eta_2 + \varepsilon_3 \\ y_4 &= \lambda^y_{42}\eta_2 + \varepsilon_4 \end{aligned}$$

note that $Cov(\varepsilon_1, \varepsilon_3)$ and $Cov(\varepsilon_2, \varepsilon_4)$ are not prespecified to be 0.

(b) Allowing correlated measurement errors across waves of the panel study appears to make sense. Specifying that the structural disturbances ζ_1 and ζ_2 are uncorrelated across waves almost surely does not (and will likely lead to the direct effect of Alienation67 on Alienation71 being over-estimated). Having only a single exogenous cause of alienation makes for a very sparse model and, in my opinion, calls into question the exogeneity of SES. Isn't SES likely to be correlated with other possible causes of alienation? — race, ethnicity, and age come immediately to mind. Finally, it doesn't make sense to me to think of SES as a *cause*, rather than a consequence, of Education and SEI. (The resulting model would not be identified, I believe.)

(c) Parameters: $\gamma_{11}, \gamma_{21}, \beta_{21}, \lambda^x_{21}, \lambda^y_{21}, \lambda^y_{42}, \phi_{11}, \psi_{11}, \psi_{22}, \theta^\delta_{11}, \theta^\delta_{22}, \theta^\varepsilon_{11}, \theta^\varepsilon_{22}, \theta^\varepsilon_{33}, \theta^\varepsilon_{44}, \theta^\varepsilon_{13}, \theta^\varepsilon_{24}$ (17 in all)

Number of observed variances and covariances $= (6)(6+1)/2 = 21$.

Given that the model is identified, it is overidentified since there are more observed variances and covariances than parameters to estimate.

(d)
```
> source("Wheaton.R")
> S.Wheaton

            Anomia67 Powerless67 Anomia71 Powerless71 Education     SEI
Anomia67      11.834       6.947    6.819       4.783    -3.839 -21.899
Powerless67    0.000       9.364    5.091       5.028    -3.889 -18.831
Anomia71       0.000       0.000   12.532       7.495    -3.841 -21.748
Powerless71    0.000       0.000    0.000       9.986    -3.625 -18.775
Education      0.000       0.000    0.000       0.000     9.610  35.522
SEI            0.000       0.000    0.000       0.000     0.000 450.288
```

```
> model.wh.1 <- specifyEquations()
> Anomia67 = 1*Alienation67
> Powerless67 = lamby21*Alienation67
> Anomia71 = 1*Alienation71
> Powerless71 = lamby42*Alienation71
> Education = 1*SES
> SEI = lambx21*SES
> Alienation67 = gam11*SES
> Alienation71 = gam21*SES + beta21*Alienation67
> V(SES) = phi11
> C(Anomia67, Anomia71) = the13
> C(Powerless67, Powerless71) = the24
```

```
NOTE: adding 8 variances to the model
```

```
> sem.wheaton.1 <- sem(model.wh.1, S=S.Wheaton, N=932)
> summary(sem.wheaton.1)

 Model Chisquare =  4.730179   Df =  4 Pr(>Chisq) = 0.3161195
 AIC =  38.73018
 BIC =  -22.61915

 Normalized Residuals
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.5932000 -0.0196300  0.0000003 -0.0197800  0.0152100  0.5209000

 R-square for Endogenous Variables
Alienation67     Anomia67  Powerless67 Alienation71     Anomia71  Powerless71
      0.3171       0.5998       0.7260       0.4972       0.6486       0.6922
   Education          SEI
      0.7082       0.4118

 Parameter Estimates
                Estimate    Std Error   z value    Pr(>|z|)
lamby21           0.9787328  0.06160671  15.886789 7.823418e-57
lamby42           0.9220701  0.05952854  15.489548 4.081488e-54
lambx21           5.2194882  0.42239523  12.356883 4.471642e-35
gam11            -0.5750067  0.05643232 -10.189315 2.213183e-24
gam21            -0.2267657  0.05235455  -4.331347 1.481999e-05
beta21            0.6070476  0.05104806  11.891688 1.307369e-32
```

```
phi11              6.8056479  0.65003492  10.469665 1.190634e-25
the13              1.6246860  0.31404091   5.173485 2.297676e-07
the24              0.3390552  0.26144912   1.296831 1.946895e-01
V[Alienation67]    4.8466819  0.46810163  10.353909 4.017385e-25
V[Anomia67]        4.7357695  0.45382926  10.435135 1.713703e-25
V[Powerless67]     2.5661049  0.40374213   6.355802 2.073425e-10
V[Alienation71]    4.0876091  0.40476977  10.098603 5.603483e-24
V[Anomia71]        4.4039125  0.51585999   8.537031 1.377151e-17
V[Powerless71]     3.0731817  0.43491032   7.066242 1.591853e-12
V[Education]       2.8043514  0.50787899   5.521692 3.357504e-08
V[SEI]           264.8813851 18.15558376  14.589527 3.274586e-48

lamby21          Powerless67 <--- Alienation67
lamby42          Powerless71 <--- Alienation71
lambx21          SEI <--- SES
gam11            Alienation67 <--- SES
gam21            Alienation71 <--- SES
beta21           Alienation71 <--- Alienation67
phi11            SES <--> SES
the13            Anomia71 <--> Anomia67
the24            Powerless71 <--> Powerless67
V[Alienation67] Alienation67 <--> Alienation67
V[Anomia67]      Anomia67 <--> Anomia67
V[Powerless67]   Powerless67 <--> Powerless67
V[Alienation71] Alienation71 <--> Alienation71
V[Anomia71]      Anomia71 <--> Anomia71
V[Powerless71]   Powerless71 <--> Powerless71
V[Education]     Education <--> Education
V[SEI]           SEI <--> SEI

 Iterations =  95
```

With the exception of the measurement-error covariance of the two powerlessness measures, all of the parameter estimates are highly statistically significant. The estimates of the structural parameters appear reasonable (although, as indicated above, the coefficient $\widehat{\beta}_{21}$ is probably an over-estimate). The overidentification test and fit indices suggest that the model does a good job of reproducing the correlational structure of the data.

(e)
```
> model.wh.2 <- specifyEquations()
> Anomia67 = 1*Alienation67
> Powerless67 = lamby*Alienation67
> Anomia71 = 1*Alienation71
> Powerless71 = lamby*Alienation71  # same loading
> Education = 1*SES
> SEI = lambx21*SES
> Alienation67 = gam11*SES
> Alienation71 = gam21*SES + beta21*Alienation67
> V(SES) = phi11
> V(Anomia67) = the11
> V(Anomia71) = the11 # same variance
> V(Powerless67) = the22
> V(Powerless71) = the22 # same variance
> C(Anomia67, Anomia71) = the13
```

```
> C(Powerless67, Powerless71) = the24
```

```
NOTE: adding 4 variances to the model
```

```
> sem.wheaton.2 <- sem(model.wh.2, S=S.Wheaton, N=932)
> summary(sem.wheaton.2)

 Model Chisquare =  6.058066   Df =  7 Pr(>Chisq) = 0.532986
 AIC =  34.05807
 BIC =  -41.80326

 Normalized Residuals
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-0.645200 -0.106100  0.000694 -0.019590  0.136100  0.422300

 R-square for Endogenous Variables
Alienation67     Anomia67  Powerless67 Alienation71     Anomia71  Powerless71
      0.3196       0.6101       0.7031       0.4964       0.6310       0.7213
   Education          SEI
      0.7071       0.4124

 Parameter Estimates
                Estimate    Std Error   z value    Pr(>|z|)
lamby             0.9546511  0.05230153  18.252831 1.964587e-74
lambx21           5.2272976  0.42284037  12.362343 4.177995e-35
gam11            -0.5830133  0.05594769 -10.420686 1.995089e-25
gam21            -0.2187281  0.05133924  -4.260448 2.040176e-05
beta21            0.5956030  0.04717088  12.626496 1.508464e-36
phi11             6.7954762  0.64929634  10.465909 1.238822e-25
the11             4.6191536  0.40476505  11.411938 3.645171e-30
the22             2.7813043  0.35076147   7.929333 2.203268e-15
the13             1.6481526  0.31213408   5.280271 1.289928e-07
the24             0.3142721  0.26053328   1.206265 2.277154e-01
V[Alienation67]   4.9183066  0.44928584  10.946943 6.872776e-28
V[Alienation71]   3.9786511  0.36665309  10.851269 1.966763e-27
V[Education]      2.8145190  0.50717555   5.549398 2.866550e-08
V[SEI]          264.6038362 18.15923626  14.571309 4.276203e-48

lamby           Powerless67 <--- Alienation67
lambx21         SEI <--- SES
gam11           Alienation67 <--- SES
gam21           Alienation71 <--- SES
beta21          Alienation71 <--- Alienation67
phi11           SES <--> SES
the11           Anomia67 <--> Anomia67
the22           Powerless67 <--> Powerless67
the13           Anomia71 <--> Anomia67
the24           Powerless71 <--> Powerless67
V[Alienation67] Alienation67 <--> Alienation67
V[Alienation71] Alienation71 <--> Alienation71
V[Education]    Education <--> Education
V[SEI]          SEI <--> SEI
```

```
 Iterations =  86

> anova(sem.wheaton.1, sem.wheaton.2)  # LR test

LR Test for Difference Between Models

              Model Df Model Chisq Df LR Chisq Pr(>Chisq)
sem.wheaton.1        4      4.7302
sem.wheaton.2        7      6.0581  3   1.3279     0.7225
```

The nonsignificant likelihood-ratio test comparing the two models suggests that constraining the measurement-model parameters to be equal across waves is consistent with the data.

Note: I decided to use a **knitr** `.Rnw` document rather than an R Markdown `.Rmd` document for these problems because LaTeX math is more easily incorporated in the former and typesetting is better. The general idea is similar: There are code blocks with live R commands that are executed when the document is compiled, and RStudio integrates support for **knitr** and `.Rnw` documents.