Tips for achieving convergence with the Stata "sem" command

Created by David B. Braudt

"We regretfully inform you that some SEMs have difficulty converging. We figure 5% to 15% of complicated models will cause difficulty. We show you what to do and it is not difficult." (p.4 Stata sem manual).

The above admission from Stata is certainly true in Latent Curve Models (LCMs). While the what to do "is not difficult," it does not always solve the convergence problems. Below are some tricks that sometimes help the Stata sem program achieve convergence. There are many other techniques which are not discussed here, but are in the Stata manual for the sem command.

Change the iterative maximization algorithm (Stata help for algorithms)

- sem (model specification), (model options) technique(nr 5 bhhh 5)
  - Other algorithm options besides "bhhh" are: dfp and bfgs
  - In the above, "5" is the number of iterations with the preceding named algorithm. Between 5-10 iterations is usually best.

Use the final estimates of a similar model as the starting values for the new model that does not converge

- Some exercises use models that only slightly differ. If a simpler model converges and another does not, use the estimates from the simpler model as starting values for nonconvergent model
  - Two options for entering start values are: 1) enter manually or 2) capture them in a matrix and use the "from()" option.
- sem (model specification for the model that did converge), (model options)
  matrix beta = e(b)
  sem (model specification), (model options) from(beta)
  - To use starting values from a model that estimates some different parameters as well as parameters in the model failing to converge, use the skip option: from(beta, skip).

Manually set the starting values (Stata help for starting values)

Option 1: set values to zero
(z1 -> (z2, init(0))) or var((e.z1, init(0)) (e.z2, init(0)))
Option 2: 50% of variances
Set error variance of indicator to 50% of the indicator's variance
Set variance of exogenous latent variables to 50% the variance of its scaling indicator
Set error variance of latent variable to 50% the variance of its scaling indicator

When Stata does not provide a chi-square test statistic

- Try the above. If those modifications fail, use the "difficult" option.
- sem (model specification), (model options) difficult
  - This option is only useful when the saturated model fails to converge.

Note: We have found instances where the chi-square test statistic reported in Stata differs from that reported in other software packages (e.g., Mplus, lavaan).

Tips for Specific Exercises

Day 2

- Simulation Data 1
  - 4a : Requires alternating optimization algorithms (nr, bhhh, and dfp).
  - 4b : Use the starting values from 4a.
- Crime
  - 4a & b : Requires alternating optimization algorithms.
- Depression
  - 4a & b : Requires the slope constraints to be 0,1,2 instead of 1,2,3.
  - 4b : Use starting values from 4a.
- Simulation Data 2
  - 4a : Requires start values = 0 for the autoregressive parameters.
  - 4b : Use the starting values from 4a.

Day 3

- Child Weight
  - 2a : Constrain residual of first measurement = 0.
  - 3b : Use starting values from 3a.
- Math Achievement
  - 1a : Constrain residual of first measurement = 0.
  - 2a : Constrain residual of first measurement = 0.
  - 2b : Requires alternating optimization algorithms (nr, bhhh, and dfp).
  - 2c : Constrain residual of first measurement = 0.
  - 2d : Use starting values from 1b plus alternating optimization algorithms.
- Crime
  - 2b : Requires alternating optimization algorithms.
- Liberal Democracy
  - 3d : Use of starting values from 3c and constrain the residual variance of libdem76 = 0.

Day 4

- Crime
  - 3a & b, 4, 5, 6a-6c, and 6e-6g : Requires alternating optimization algorithms.

Day 5

- Math, Reading, and Covariates
  - 3b : Use starting values from Math LCM (1) and alternating optimization algorithms (all).
  - 4 : Requires alternating optimization algorithms.
  - 5 : Requires alternating optimization algorithms.
    - Note: The df are off by two because Stata will not allow the estimation of a covariance between an exogenous and endogenous variable.
  - 6 : Requires alternating optimization algorithms (nr, bhhh, and dfp).
- Liberal Democracy
  - 1b : Requires starting values from 1a to converge.
  - 1c : will not converge … Let us know if you think of something that may help.