Title

example 7 — Nonrecursive structural model

Description   Remarks and examples   References   Also see

Description

To demonstrate a nonrecursive structural model with all variables observed, we use data from Duncan, Haller, and Portes (1968):

```
. use http://www.stata-press.com/data/r14/sem_sm1
(Structural model with all observed values)
. ssd describe
  Summary statistics data from
  http://www.stata-press.com/data/r14/sem_sm1.dta
    obs:             329                 Structural model with all obse..
   vars:              10                 25 May 2014 10:13
                                         (_dta has notes)
  ------------------------------------------------------------------------
  variable name                  variable label
  ------------------------------------------------------------------------
  r_intel                        respondent's intelligence
  r_parasp                       respondent's parental aspiration
  r_ses                          respondent's family socioeconomic status
  r_occasp                       respondent's occupational aspiration
  r_educasp                      respondent's educational aspiration
  f_intel                        friend's intelligence
  f_parasp                       friend's parental aspiration
  f_ses                          friend's family socioeconomic status
  f_occasp                       friend's occupational aspiration
  f_educasp                      friend's educational aspiration
  ------------------------------------------------------------------------
. notes
_dta:
  1.  Summary statistics data from Duncan, O.D., Haller, A.O., and Portes, A.,
      1968, "Peer Influences on Aspirations:  A Reinterpretation", _American
      Journal of Sociology_ 74, 119-137.
  2.  The data contain 329 boys with information on five variables and the same
      information for each boy's best friend.
```

If you typed ssd status, you would learn that this dataset contains the correlation matrix only. Variances (standard deviations) and means are undefined. Thus we need to use this dataset cautiously. It is always better if you enter the variances and means if you have them.

That these data are the correlations only will not matter for how we will use them.

Remarks and examples

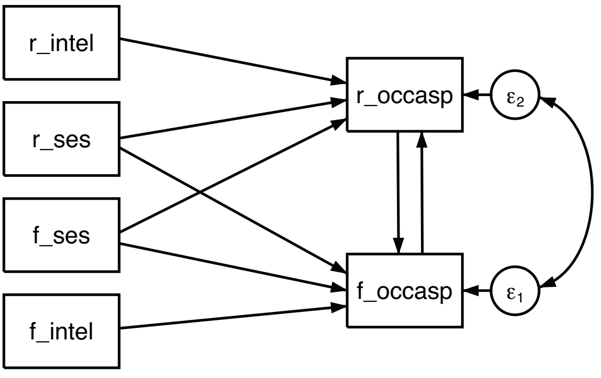
stata.com

See *Structural models 8: Dependencies between response variables* in [SEM] **intro 5** for background.

Remarks are presented under the following headings:

Fitting the model

In the referenced paper above, the authors fit the following model:



```
. sem (r_occasp <- f_occasp r_intel r_ses f_ses)
>     (f_occasp <- r_occasp f_intel f_ses r_ses),
>                         cov(e.r_occasp*e.f_occasp) standardized
Endogenous variables
Observed:  r_occasp f_occasp
Exogenous variables
Observed:  r_intel r_ses f_ses f_intel
Fitting target model:
Iteration 0:   log likelihood = -2617.0489
Iteration 1:   log likelihood = -2617.0489
Structural equation model                       Number of obs     =        329
Estimation method  = ml
Log likelihood     = -2617.0489
```

Standardized	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
r_occ~p <-						
f_occasp	.2773441	.1281904	2.16	0.031	.0260956	.5285926
r_intel	.2854766	.05	5.71	0.000	.1874783	.3834748
r_ses	.1570082	.0520841	3.01	0.003	.0549252	.2590912
f_ses	.0973327	.060153	1.62	0.106	-.020565	.2152304
f_occ~p <-						
r_occasp	.2118102	.156297	1.36	0.175	-.0945264	.5181467
r_ses	.0794194	.0587732	1.35	0.177	-.0357739	.1946127
f_ses	.1681772	.0537199	3.13	0.002	.062888	.2734663
f_intel	.3693682	.0525924	7.02	0.000	.2662891	.4724474
var(e.r_oc~p)	.6889244	.0399973			.6148268	.7719519
var(e.f_oc~p)	.6378539	.039965			.5641425	.7211964
cov(e.r_oc~p, e.f_occasp)	-.2325666	.2180087	-1.07	0.286	-.6598558	.1947227

```
LR test of model vs. saturated: chi2(0)   =      0.00, Prob > chi2 =      .
```

Notes:

1. We specified the `standardized` option, but in this case that did not matter much because these data are based on the correlation coefficients only, so standardized values are equal to unstandardized values. The exception is the correlation between the latent endogenous variables, as reflected in the correlation of their errors, and we wanted to show that results match those in the original paper.
2. Nearly all results match those in the original paper. The authors normalized the errors to have a variance of 1; `sem` normalizes the paths from the errors to have coefficient 1. While you can apply most normalizing constraints any way you wish, `sem` restricts errors to have path coefficients of 1 and this cannot be modified. You could, however, prove to yourself that `sem` would produce the same variances as the authors produced by typing

```
. sem, coeflegend
. display sqrt(_b[var(e.r_occasp):_cons])
. display sqrt(_b[var(e.f_occasp):_cons])
```

because the coefficients would be the standard deviations of the errors estimated without the variance-1 constraint. Thus all results match. We replayed results by using the `coeflegend` option so that we would know what to type to refer to the two error variances, namely, `_b[var(e.r_occasp):_cons]` and `_b[var(e.f_occasp):_cons]`.

Fitting the model with the Builder

Use the diagram above for reference.

1. Open the dataset.

   In the Command window, type

   ```
   . use http://www.stata-press.com/data/r14/sem_sm1
   ```

2. Open a new Builder diagram.

   Select menu item **Statistics > SEM (structural equation modeling) > Model building and estimation**.

3. Create the four independent variables.

   Select the Add Observed Variables Set tool, , and then click in the diagram about one-third of the way in from the left and one-fourth of the way up from the bottom.

   In the resulting dialog box,

   a. select the *Select variables* radio button (it may already be selected);

   b. use the *Variables* control to select the four variables in this order: `r_intel`, `r_ses`, `f_ses`, and `f_intel`;

   c. select `Vertical` in the *Orientation* control;

   d. click on **OK**.

   If you wish, move the set of variables by clicking on any variable and dragging it.

4. Create the two dependent variables.

   Select the Add Observed Variables Set tool, , and then click in the diagram about two-thirds of the way in from the left, vertically aligned with the top of the `f_intel` rectangle.

   In the resulting dialog box,

   a. select the *Select variables* radio button (it may already be selected);

   b. use the *Variables* control to select the two variable names `r_occasp` and `f_occasp`;

   c. select `Vertical` in the *Orientation* control;

   d. select the *Distances* tab;

   e. select `.5` (inch) from the *Distance between variables* control;

   f. click on **OK**.

   If you wish, move the set of variables by clicking on any variable and dragging it.

5. Create paths from the independent variables to the dependent variables.

   a. Select the Add Path tool, .

   b. Click in the right side of the `r_intel` rectangle (it will highlight when you hover over it), and drag a path to the left side of the `r_occasp` rectangle (it will highlight when you can release to connect the path).

c. Continuing with the tool, create the following paths by first clicking in the right side of the rectangle for the independent variable and dragging it to the left side of the rectangle for the dependent variable.

```
r_ses -> r_occasp
r_ses -> f_occasp
f_ses -> r_occasp
f_ses -> f_occasp
f_intel -> f_occasp
```

6. Create paths for the feedback loop between endogenous variables.

   a. Continue with the Add Path tool, .

   b. Click in the bottom of the `r_occasp` rectangle, slightly to the left of the center, and drag a path to the `f_occasp` rectangle.

   c. Click in the top of the `f_occasp` rectangle, slightly to the right of the center, and drag a path to the `r_occasp` rectangle.

7. Correlate the error terms.

   a. Select the Add Covariance tool, .

   b. Click in the error term for `r_occasp` (the circle labeled ϵ_1), and drag a covariance to the error term for `f_occasp` (the circle labeled ϵ_2).

8. Clean up.

   If you do not like where a path has been connected to its variable, use the Select tool, , to click on the path, and then simply click on where it connects to a rectangle and drag the endpoint. Similarly, you can change where the covariance connects to the error terms by clicking on the covariance and dragging the endpoint. You can also change the bow of the covariance by clicking on the covariance and dragging the control point that extends from one end of the selected covariance.

9. Estimate.

   Click on the **Estimate** button, , in the Standard Toolbar, and then click on **OK** in the resulting *SEM estimation options* dialog box.

10. Show standardized estimates.

    From the SEM Builder menu, select **View > Standardized Estimates**.

Tips: When you draw paths that should be exactly horizontal or vertical, such as the two between `r_occasp` and `f_occasp`, holding the *Shift* key as you drag the path will guarantee that the line is perfectly vertical. Also, when drawing paths from the independent variables to the dependent variables, you may find it more convenient to change the automation settings as described in the tips of [SEM] **example 9**. However, this does not work for the feedback loop between the dependent variables.

You can open a completed diagram in the Builder by typing

```
. webgetsem sem_sm1
```

Checking stability with estat stable

```
. estat stable

Stability analysis of simultaneous equation systems

   Eigenvalue stability condition
```

Eigenvalue	Modulus
.2423722	.242372
-.2423722	.242372

```
   stability index =  .2423722
   All the eigenvalues lie inside the unit circle.
   SEM satisfies stability condition.
```

Notes:

1. `estat stable` is for use on nonrecursive models. Recursive models are, by design, stable.
2. Stability concerns whether the parameters of the model are such that the model would blow up if it were operated over and over again. If the results are found not to be stable, then that casts questions about the validity of the model.
3. The stability is the maximum of the moduli, and the moduli are the absolute values of the eigenvalues. Usually, the two eigenvalues are not identical, but it is a property of this model that they are.
4. If the stability index is less than 1, then the reported estimates yield a stable model.

In the next section, we use `estat teffects` to estimate total effects. That is appropriate only if the model is stable, as we find that it is.

Reporting total, direct, and indirect effects with estat teffects

```
. estat teffects
```

Direct effects

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
r_occ~p <-						
r_occasp	0	(no path)				
f_occasp	.2773441	.1287622	2.15	0.031	.0249748	.5297134
r_intel	.2854766	.0522001	5.47	0.000	.1831662	.3877869
r_ses	.1570082	.052733	2.98	0.003	.0536534	.260363
f_ses	.0973327	.0603699	1.61	0.107	-.0209901	.2156555
f_intel	0	(no path)				
f_occ~p <-						
r_occasp	.2118102	.1563958	1.35	0.176	-.09472	.5183404
f_occasp	0	(no path)				
r_intel	0	(no path)				
r_ses	.0794194	.0589095	1.35	0.178	-.0360411	.1948799
f_ses	.1681772	.0543854	3.09	0.002	.0615838	.2747705
f_intel	.3693682	.0557939	6.62	0.000	.2600142	.4787223

Indirect effects

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
r_occ~p <-						
r_occasp	.0624106	.0460825	1.35	0.176	-.0279096	.1527307
f_occasp	.0173092	.0080361	2.15	0.031	.0015587	.0330597
r_intel	.0178168	.0159383	1.12	0.264	-.0134217	.0490552
r_ses	.0332001	.0204531	1.62	0.105	-.0068872	.0732875
f_ses	.0556285	.0292043	1.90	0.057	-.0016109	.112868
f_intel	.1088356	.052243	2.08	0.037	.0064411	.21123
f_occ~p <-						
r_occasp	.0132192	.0097608	1.35	0.176	-.0059115	.0323499
f_occasp	.0624106	.0289753	2.15	0.031	.0056201	.1192011
r_intel	.0642406	.0490164	1.31	0.190	-.0318298	.160311
r_ses	.0402881	.0315496	1.28	0.202	-.021548	.1021242
f_ses	.0323987	.0262124	1.24	0.216	-.0189765	.083774
f_intel	.0230525	.0202112	1.14	0.254	-.0165607	.0626657

Total effects

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
r_occ~p <-						
r_occasp	.0624106	.0460825	1.35	0.176	-.0279096	.1527307
f_occasp	.2946533	.1367983	2.15	0.031	.0265335	.5627731
r_intel	.3032933	.0509684	5.95	0.000	.2033971	.4031896
r_ses	.1902083	.050319	3.78	0.000	.091585	.2888317
f_ses	.1529612	.050844	3.01	0.003	.0533089	.2526136
f_intel	.1088356	.052243	2.08	0.037	.0064411	.21123
f_occ~p <-						
r_occasp	.2250294	.1661566	1.35	0.176	-.1006315	.5506903
f_occasp	.0624106	.0289753	2.15	0.031	.0056201	.1192011
r_intel	.0642406	.0490164	1.31	0.190	-.0318298	.160311
r_ses	.1197074	.0483919	2.47	0.013	.0248611	.2145537
f_ses	.2005759	.0488967	4.10	0.000	.10474	.2964118
f_intel	.3924207	.0502422	7.81	0.000	.2939478	.4908936

Note:

1. In the path diagram we drew for this model, you can see that the intelligence of the respondent, `r_intel`, has both direct and indirect effects on the occupational aspiration of the respondent, `r_occasp`. The tables above reveal that

$$0.3033 = 0.2855 + 0.0178$$

where 0.2855 is the direct effect and 0.0178 is the indirect effect.

References

Acock, A. C. 2013. *Discovering Structural Equation Modeling Using Stata*. Rev. ed. College Station, TX: Stata Press.

Duncan, O. D., A. O. Haller, and A. Portes. 1968. Peer influences on aspirations: A reinterpretation. *American Journal of Sociology* 74: 119–137.

Also see

[SEM] **example 8** — Testing that coefficients are equal, and constraining them

[SEM] **sem** — Structural equation model estimation command

[SEM] **estat stable** — Check stability of nonrecursive system

[SEM] **estat teffects** — Decomposition of effects into total, direct, and indirect