Title stata.com

example 42g — One- and two-level mediation models (multilevel)

Description Remarks and examples References Also see

Description

To demonstrate linear mediation models, we use the following data:

```
. use http://www.stata-press.com/data/r14/gsem_multmed
(Fictional job-performance data)

. summarize

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      branch |      1,500          38    21.65593          1         75
     support |      1,500    .0084667    .5058316       -1.6        1.8
       satis |      1,500       .0212    .6087235       -1.6          2
     perform |      1,500    5.005317    .8949845    2.35022   8.084294

. notes

_dta:
  1.  Fictional data on job performance, job satisfaction, and perceived
      support from managers for 1,500 sales employees of a large department
      store in 75 locations.
  2.  Variable support is average of Likert-scale questions, each question
      scored from -2 to 2.
  3.  Variable satis is average of Likert-scale questions, each question scored
      from -2 to 2.
  4.  Variable perform is job performance measured on continuous scale.
```

See *Structural models 1: Linear regression* and *Multilevel mixed-effects models* in [SEM] **intro 5** for background.

Remarks and examples stata.com

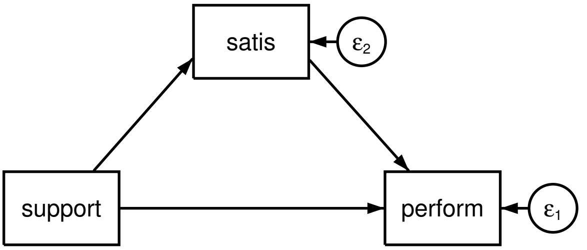
Remarks are presented under the following headings:

One-level model with sem
One-level model with gsem
Two-level model with gsem
Fitting the models with the Builder

One-level model with sem

You can fit single-level mediation models with `sem` or `gsem`. You will be better off using `sem` because then you can use `estat teffects` afterward to compute indirect and total effects.

The model we wish to fit is the simplest form of a mediation model, namely,



We are interested in the effect of managerial support on job performance, but we suspect a portion of the effect might be mediated through job satisfaction. In traditional mediation analysis, the model would be fit by a series of linear regression models as described in Baron and Kenny (1986). That approach is sufficient because the errors are not correlated. The advantage of using structural equation modeling is that you can fit a single model and estimate the indirect and total effects, and you can embed the simple mediation model in a larger model and even use latent variables to measure any piece of the mediation model.

To fit this model with the command syntax, we type

```
. sem (perform <- satis support) (satis <- support)
Endogenous variables
Observed:  perform satis
Exogenous variables
Observed:  support
Fitting target model:
Iteration 0:   log likelihood = -3779.9224
Iteration 1:   log likelihood = -3779.9224
Structural equation model                       Number of obs     =      1,500
Estimation method  = ml
Log likelihood     = -3779.9224

------------------------------------------------------------------------------
             |                 OIM
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Structural   |
  perform <- |
       satis |   .8984401   .0251903    35.67   0.000      .849068    .9478123
     support |   .6161077   .0303143    20.32   0.000     .5566927    .6755227
       _cons |   4.981054   .0150589   330.77   0.000     4.951539    5.010569
  -----------+----------------------------------------------------------------
  satis <-   |
     support |   .2288945   .0305047     7.50   0.000     .1691064    .2886826
       _cons |    .019262   .0154273     1.25   0.212    -.0109749    .0494989
-------------+----------------------------------------------------------------
var(e.perf~m)|   .3397087   .0124044                      .3162461     .364912
 var(e.satis)|   .3569007   .0130322                      .3322507    .3833795
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(0)   =      0.00, Prob > chi2 =      .
```

Notes:

1. The direct effect of managerial support on job performance is measured by `perform <- support` and is estimated to be 0.6161. The effect is small albeit highly statistically significant. The standard deviations of performance and support are 0.89 and 0.51. A one standard deviation increase in support improves performance by a third of a standard deviation.
2. The direct effect of job satisfaction on job performance is measured by `perform <- satis` and is estimated to be 0.8984. That also is a moderate effect, practically speaking, and is highly statistically significant.
3. The effect of managerial support on job satisfaction is measured by `satis <- support` and is practically small but statistically significant.
4. What is the total effect of managerial support on performance? It is the direct effect (0.6161) plus the indirect effect of support on satisfaction on performance ($0.2289 \times 0.8984 = 0.2056$), meaning the total effect is 0.8217. It would be desirable to put a standard error on that, but that's more work.

We can use `estat teffects` after estimation to obtain the total effect and its standard error:

```
. estat teffects
```

Direct effects

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
perform <-						
satis	.8984401	.0251903	35.67	0.000	.849068	.9478123
support	.6161077	.0303143	20.32	0.000	.5566927	.6755227
satis <-						
support	.2288945	.0305047	7.50	0.000	.1691064	.2886826

Indirect effects

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
perform <-						
satis	0	(no path)				
support	.205648	.0280066	7.34	0.000	.150756	.26054
satis <-						
support	0	(no path)				

Total effects

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
perform <-						
satis	.8984401	.0251903	35.67	0.000	.849068	.9478123
support	.8217557	.0404579	20.31	0.000	.7424597	.9010516
satis <-						
support	.2288945	.0305047	7.50	0.000	.1691064	.2886826

One-level model with gsem

We can fit the same model with `gsem`. The command is the same except that we substitute `gsem` for `sem`, and results are identical:

```
. gsem (perform <- satis support) (satis <- support)
Iteration 0:   log likelihood = -2674.3421
Iteration 1:   log likelihood = -2674.3421  (backed up)
Generalized structural equation model          Number of obs     =      1,500
Response       : perform
Family         : Gaussian
Link           : identity
Response       : satis
Family         : Gaussian
Link           : identity
Log likelihood = -2674.3421
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
perform <-						
satis	.8984401	.0251903	35.67	0.000	.849068	.9478123
support	.6161077	.0303143	20.32	0.000	.5566927	.6755227
_cons	4.981054	.0150589	330.77	0.000	4.951539	5.010569
satis <-						
support	.2288945	.0305047	7.50	0.000	.1691064	.2886826
_cons	.019262	.0154273	1.25	0.212	-.0109749	.0494989
var(e.perf~m)	.3397087	.0124044			.3162461	.364912
var(e.satis)	.3569007	.0130322			.3322507	.3833795

After `gsem`, however, we cannot use `estat teffects`:

```
. estat teffects
subcommand estat teffects is unrecognized
r(321);
```

We can, however, calculate the indirect and total effects for ourselves and obtain the standard error by using `nlcom`. Referring back to note 4 of the previous section, the formula for the indirect effect and total effects are

$$\text{indirect effect} = \beta_1\beta_4$$
$$\text{total effect} = \beta_2 + \beta_1\beta_4$$

where

β_1 = path coefficient for `perform <- satis`

β_4 = path coefficient for `satis    <- support`

β_2 = path coefficient for `perform <- support`

It turns out that we can access the coefficients by typing

β_1 = `_b[perform:satis]`

β_4 = `_b[satis:support]`

β_2 = `_b[perform:support]`

which is most easily revealed by typing

```
. gsem, coeflegend
  (output omitted)
```

Thus we can obtain the indirect effect by typing

```
. nlcom _b[perform:satis]*_b[satis:support]
       _nl_1:  _b[perform:satis]*_b[satis:support]
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_nl_1	.205648	.0280066	7.34	0.000	.150756	.26054

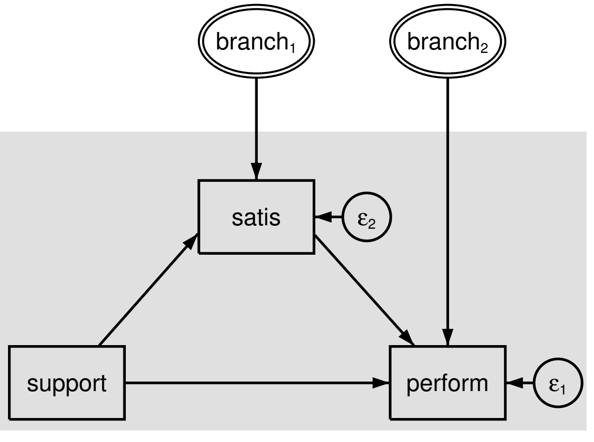
and we can obtain the total effect by typing

```
. nlcom _b[perform:support]+_b[perform:satis]*_b[satis:support]
       _nl_1:  _b[perform:support]+_b[perform:satis]*_b[satis:support]
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_nl_1	.8217557	.0404579	20.31	0.000	.7424597	.9010516

Two-level model with gsem

It may be easier to use `sem` rather than `gsem` for fitting single-level models, but if you want to fit multilevel models, you must use `gsem`. A variation on the model we just fit is



In this model, we include a random intercept in each equation at the branch (individual store) level. The model above is one of many variations on two-level mediation models; see Krull and MacKinnon (2001) for an introduction to multilevel mediation models, and see Preacher, Zyphur, and Zhang (2010) for a discussion of fitting these models with structural equation modeling.

To fit this model with the command syntax, we type

```
. gsem (perform <- satis support M1[branch]) (satis <- support M2[branch]),
> cov(M1[branch]*M2[branch]@0)
Fitting fixed-effects model:
Iteration 0:   log likelihood = -2674.3421
Iteration 1:   log likelihood = -2674.3421
Refining starting values:
Grid node 0:   log likelihood = -2132.1613
Fitting full model:
Iteration 0:   log likelihood = -2132.1613  (not concave)
Iteration 1:   log likelihood = -1801.3155
Iteration 2:   log likelihood = -1769.6421
Iteration 3:   log likelihood = -1705.1282
Iteration 4:   log likelihood =  -1703.746
Iteration 5:   log likelihood = -1703.7141
Iteration 6:   log likelihood =  -1703.714
Generalized structural equation model           Number of obs     =      1,500
Response       : perform
Family         : Gaussian
Link           : identity
Response       : satis
Family         : Gaussian
Link           : identity
Log likelihood =  -1703.714
 ( 1)  [perform]M1[branch] = 1
 ( 2)  [satis]M2[branch] = 1
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
perform <-						
satis	.604264	.0336398	17.96	0.000	.5383313	.6701968
support	.6981525	.0250432	27.88	0.000	.6490687	.7472364
M1[branch]	1	(constrained)				
_cons	4.986596	.0489465	101.88	0.000	4.890663	5.082529
satis <-						
support	.2692633	.0179649	14.99	0.000	.2340528	.3044739
M2[branch]	1	(constrained)				
_cons	.0189202	.0570868	0.33	0.740	-.0929678	.1308083
var(M1[branch])	.1695962	.0302866			.119511	.2406713
var(M2[branch])	.2384738	.0399154			.1717781	.3310652
var(e.perf~m)	.201053	.0075451			.1867957	.2163985
var(e.satis)	.1188436	.0044523			.1104299	.1278983

Notes:

1. In *One-level model with sem* above, we measured the direct effects on job performance of job satisfaction and managerial support as 0.8984 and 0.6161. Now the direct effects are 0.6043 and 0.6982.

2. We can calculate the indirect and total effects just as we did in the previous section, which we will do below. We mentioned earlier that there are other variations of two-level mediation models, and how you calculate total effects depends on the model chosen.

   In this case, the indirect effect is

   ```
   . nlcom _b[perform:satis]*_b[satis:support]
          _nl_1:  _b[perform:satis]*_b[satis:support]
   ```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_nl_1	.1627062	.0141382	11.51	0.000	.1349958	.1904165

   and the total effect is

   ```
   . nlcom _b[perform:support]+_b[perform:satis]*_b[satis:support]
          _nl_1:  _b[perform:support]+_b[perform:satis]*_b[satis:support]
   ```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_nl_1	.8608587	.0257501	33.43	0.000	.8103894	.911328

Fitting the models with the Builder

Use the diagram in *One-level model with sem* above for reference.

1. Open the dataset.

   In the Command window, type

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

2. Open a new Builder diagram.

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

3. Create a regression component for the `perform` outcome.

   Select the Add Regression Component tool, , and then click in the center of the diagram.

   In the resulting dialog box,

   a. select `perform` in the *Dependent variable* control;

   b. select `support` with the *Independent variables* control;

   c. select `Left` in the *Independent variables' direction* control;

   d. click on **OK**.

   e. Use the Select tool, , to select only the `perform` rectangle, and drag it to the right to increase the distance between the rectangles. (You can hold the *Shift* key while dragging to ensure that the movement is directly to the right.)

4. Create the mediating variable.

   a. Select the Add Observed Variable tool, , and then click in the diagram above the path from `support` to `perform`.

   b. In the Contextual Toolbar, select `satis` with the *Variable* control.

5. Create the paths to and from the mediating variable.

   a. Select the Add Path tool, .

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

   c. Continuing with the tool, draw a path from the lower right of the `satis` rectangle to the upper left of the `perform` rectangle.

6. Clean up the direction of the error term.

   We want the error for each of the endogenous variables to be to the right of the rectangle. The error for `satis` may have been created in another direction. If so,

   a. choose the Select tool, ;

   b. click in the `satis` rectangle;

   c. click on one of the **Error Rotation** buttons, , in the Contextual Toolbar until the error is to the right of the rectangle.

7. Clean up the location of the paths.

   If you do not like where the paths have been connected to the rectangles, use the Select tool, , to click on the path, and then simply click on where it connects to a rectangle and drag the endpoint.

8. Estimate.

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

9. To fit the model in *Two-level model with gsem*, continue with the previous diagram, and put the builder in `gsem` mode by clicking on the button.

10. Create the multilevel latent variable corresponding to the random intercept for `satis`.

    a. Select the Add Multilevel Latent Variable tool, , and click above the rectangle for `satis`.

    b. In the Contextual Toolbar, click on the ... button.

    c. Select the nesting level and nesting variable by selecting `2` from the *Nesting depth* control and selecting `branch` > *Observations* in the next control.

    d. Specify `M1` as the *Base name*.

    e. Click on **OK**.

11. Create the multilevel latent variable corresponding to the random intercept for `perform`.

    a. Select the Add Multilevel Latent Variable tool, , and click above the rectangle for `satis` and to the right of the `branch`$_1$ double oval.

    b. In the Contextual Toolbar, click on the ... button.

    c. Select the nesting level and nesting variable by selecting `2` from the *Nesting depth* control and selecting `branch` > *Observations* in the next control.

    d. Specify `M2` as the *Base name*.

    e. Click on **OK**.

12. Draw paths from the multilevel latent variables to their corresponding endogenous variables.
    a. Select the Add Path tool, .
    b. Click in the bottom of the `branch`$_1$ double oval, and drag a path to the top of the `satis` rectangle.
    c. Continuing with the tool, click in the bottom of the `branch`$_2$ double oval, and drag a path to the top of the `perform` rectangle.
13. Estimate again.

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

You can open a completed diagram for the first model in the Builder by typing

```
. webgetsem sem_med
```

You can open a completed diagram for the second model in the Builder by typing

```
. webgetsem gsem_mlmed
```

References

Baron, R. M., and D. A. Kenny. 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51: 1173–1182.

Krull, J. L., and D. P. MacKinnon. 2001. Multilevel modeling of individual and group level mediated effects. *Multivariate Behavorial Research* 36: 249–277.

Preacher, K. J., M. J. Zyphur, and Z. Zhang. 2010. A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods* 15: 209–233.

Also see

[SEM] **example 38g** — Random-intercept and random-slope models (multilevel)

[SEM] **gsem** — Generalized structural equation model estimation command

[SEM] **intro 5** — Tour of models