Example 44g — Interval regression

Description Remarks and examples Also see

Description

Interval regression is demonstrated using `intregxmpl.dta`:

```
. use https://www.stata-press.com/data/r19/intregxmpl
(Wages of women)
```

See *Structural models 1: Linear regression* in [SEM] **Intro 5** for background.

Remarks and examples

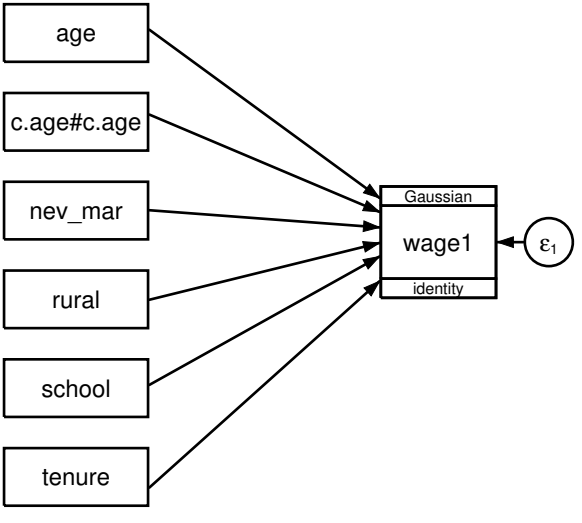
Remarks are presented under the following headings:

Fitting interval regression models

The first example in [R] **intreg** is

```
. intreg wage1 wage2 age c.age#c.age nev_mar rural school tenure
```

This model corresponds to



Interval measure information does not appear in the path diagram by default. It can be added to the path diagram by customizing the appearance of `wage1` in the Builder. The Builder reports the interval measure information for `wage1` in the Details pane.

To fit this model with gsem, we type

```
. gsem wage1 <- age c.age#c.age nev_mar rural school tenure,
> family(gaussian, udepvar(wage2))
Refining starting values:
Grid node 0:  Log likelihood = -856.59446
Fitting full model:
Iteration 0:  Log likelihood = -856.59446
Iteration 1:  Log likelihood = -856.33321
Iteration 2:  Log likelihood = -856.33293
Iteration 3:  Log likelihood = -856.33293
Generalized structural equation model                 Number of obs     = 488
Lower response: wage1                                       Uncensored =   0
Upper response: wage2                                  Left-censored =  14
Family:         Gaussian                              Right-censored =   6
Link:           Identity                               Interval-cens. = 468
Log likelihood = -856.33293
```

	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
wage1						
age	.7914438	.4433604	1.79	0.074	-.0775265	1.660414
c.age#c.age	-.0132624	.0073028	-1.82	0.069	-.0275757	.0010509
nev_mar	-.2075022	.8119581	-0.26	0.798	-1.798911	1.383906
rural	-3.043044	.7757324	-3.92	0.000	-4.563452	-1.522637
school	1.334721	.1357873	9.83	0.000	1.068583	1.600859
tenure	.8000664	.1045077	7.66	0.000	.5952351	1.004898
_cons	-12.70238	6.367117	-1.99	0.046	-25.1817	-.2230584
var(e.wage1)	53.28454	3.693076			46.51635	61.0375

Notes:

1. Just like intreg, gsem requires two dependent variables for fitting interval regression models. The udepvar() suboption in family(gaussian) allows you to specify the dependent variable containing the upper-limit values for the interval regression. Consequently, the dependent variable participating in the path specification necessarily contains the lower-limit values.
2. Reported coefficients match those reported by intreg.
3. Reported standard errors (SEs) match those reported by intreg.
4. gsem reports the point estimate of e.wage1 as 53.28454. This is an estimate of σ^2, the error variance. intreg reports an estimated σ as 7.299626. And $\sqrt{53.28454} = 7.299626$.

Fitting the model with the Builder

Use the diagram in *Fitting interval regression models* above for reference.

1. Open the dataset.

   In the Command window, type

   ```
   . use https://www.stata-press.com/data/r19/intregxmpl
   ```

2. Open a new Builder diagram.

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

3. Put the Builder in `gsem` mode by clicking on the GSEM button.

4. Create the independent variables.

   Select the Add observed variables set tool, , and then click in the diagram about one-fourth 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. type the following variable names in the *Variables* control: `age`, `c.age#c.age`, `nev_mar`, `rural`, `school`, and `tenure`;

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

   d. click on **OK**.

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

5. Create the interval responses.

   a. Select the Add generalized response variable tool, .

   b. Click about one-third of the way in from the right side of the diagram, to the right of the `nev_mar` rectangle.

   c. In the Contextual Toolbar, select `Gaussian, Identity` in the *Family/Link* control (it may already be selected).

   d. In the Contextual Toolbar, use the *Variable* control to select the variable `wage1`.

   e. In the Contextual Toolbar, click on the **Properties...** button.

   f. In the resulting *Variable properties* dialog box, click on the **Censoring...** button in the **Variable** tab.

   g. In the resulting *Censoring* dialog box, select the *Interval-measured, depvar is lower boundary* radio button. In the resulting *Interval-measured* box below, use the *Upper bound* control to select the variable `wage2`.

   h. Click on **OK** in the *Censoring* dialog box, and then click on **OK** in the *Variable properties* dialog box. The Details pane will now show that `wage1` is the lower bound and `wage2` is the upper bound of our interval measure.

6. Create paths from the independent variables to the dependent variable.

   a. Select the Add path tool, .

   b. Click in the right side of the `age` rectangle (it will highlight when you hover over it), and drag a path to the left side of the `wage1` rectangle (it will highlight when you can release to connect the path).

c. Continuing with the ⟶ tool, create the following paths by clicking first 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:

```
c.age#c.age -> wage1
nev_mar -> wage1
rural -> wage1
school -> wage1
tenure -> wage1
```

7. Estimate.

   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 in the Builder by typing

```
. webgetsem gsem_intreg
```

Also see

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

[SEM] **Example 43g** — Tobit regression

[SEM] **Example 45g** — Heckman selection model

[SEM] **Example 46g** — Endogenous treatment-effects model

[SEM] **Intro 5** — Tour of models

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