Description

Below we demonstrate *sem*'s `reliability()` option with the following data:

```stata
. use https://www.stata-press.com/data/r17/sem_rel
(measurement error with known reliabilities)
. summarize
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>1,234</td>
<td>701.081</td>
<td>71.79378</td>
<td>487</td>
<td>943</td>
</tr>
<tr>
<td>x1</td>
<td>1,234</td>
<td>100.278</td>
<td>14.1552</td>
<td>51</td>
<td>149</td>
</tr>
<tr>
<td>x2</td>
<td>1,234</td>
<td>100.2066</td>
<td>14.50912</td>
<td>55</td>
<td>150</td>
</tr>
</tbody>
</table>

`. notes
_dta:
1. Fictional data.
2. Variables x1 and x2 each contain a test score designed to measure X. The
test is scored to have mean 100.
3. Variables x1 and x2 are both known to have reliability 0.5.
4. Variable y is the outcome, believed to be related to X.

See [SEM] sem and gsem option reliability() for background.

Remarks and examples

Remarks are presented under the following headings:

- Baseline model (reliability ignored)
- Model with reliability
- Model with two measurement variables and reliability
Baseline model (reliability ignored)

```
.seml (y <- x1)
Endogenous variables
Observed: y
Exogenous variables
Observed: x1
Fitting target model:
Iteration 0: log likelihood = -11629.745
Iteration 1: log likelihood = -11629.745
Structural equation model
Number of obs = 1,234
Estimation method: ml
Log likelihood = -11629.745

<table>
<thead>
<tr>
<th></th>
<th>OIM</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>std. err.</td>
<td>z</td>
<td>P&gt;</td>
<td>z</td>
<td></td>
</tr>
<tr>
<td>Structural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x1</td>
<td>3.54976</td>
<td>.1031254</td>
<td>34.42</td>
<td>0.000</td>
<td>3.347637</td>
<td>3.751882</td>
</tr>
<tr>
<td>_cons</td>
<td>345.1184</td>
<td>10.44365</td>
<td>33.05</td>
<td>0.000</td>
<td>324.6492</td>
<td>365.5876</td>
</tr>
<tr>
<td>var(e.y)</td>
<td>2627.401</td>
<td>105.7752</td>
<td></td>
<td></td>
<td>2428.053</td>
<td>2843.115</td>
</tr>
</tbody>
</table>

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

Notes:
1. In these data, variable x1 is measured with error.
2. If we ignore that, we obtain a path coefficient for y<-x1 of 3.55.
3. We also ran this model for y<-x2. We obtained a path coefficient of 3.48.
Model with reliability

```
. sem (x1<-X) (y<-X), reliability(x1 .5)
```

Endogenous variables
Measurement: x1 y
Exogenous variables
Latent: X

Fitting target model:
Iteration 0: log likelihood = -11745.845
Iteration 1: log likelihood = -11661.626
Iteration 2: log likelihood = -11631.469
Iteration 3: log likelihood = -11629.755
Iteration 4: log likelihood = -11629.745
Iteration 5: log likelihood = -11629.745

Structural equation model
Estimation method: ml
Log likelihood = -11629.745

( 1) [x1]X = 1
( 2) [\]var(e.x1) = 100.1036

<table>
<thead>
<tr>
<th></th>
<th>OIM</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement</td>
<td></td>
<td>Coefficient</td>
<td>std. err.</td>
<td>z</td>
<td>P&gt;</td>
</tr>
<tr>
<td>x1</td>
<td>X</td>
<td>1 (constrained)</td>
<td>100.278</td>
<td>.4027933</td>
<td>248.96</td>
</tr>
<tr>
<td>_cons</td>
<td></td>
<td>7.09952</td>
<td>.352463</td>
<td>20.14</td>
<td>0.000</td>
</tr>
<tr>
<td>y</td>
<td>X</td>
<td>701.081</td>
<td>2.042929</td>
<td>343.17</td>
<td>0.000</td>
</tr>
<tr>
<td>_cons</td>
<td></td>
<td>104.631</td>
<td>207.3381</td>
<td>2.152334</td>
<td>5086.411</td>
</tr>
<tr>
<td>var(e.x1)</td>
<td></td>
<td>100.1036</td>
<td>(constrained)</td>
<td>2.152334</td>
<td>5086.411</td>
</tr>
<tr>
<td>var(e.y)</td>
<td></td>
<td>104.631</td>
<td>207.3381</td>
<td>2.152334</td>
<td>5086.411</td>
</tr>
<tr>
<td>var(X)</td>
<td></td>
<td>100.1036</td>
<td>8.060038</td>
<td>85.48963</td>
<td>117.2157</td>
</tr>
</tbody>
</table>

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

Notes:
1. We wish to estimate the effect of y<-x1 when x1 is measured with error (0.50 reliability). To do that, we introduce latent variable X and write our model as (x1<-X) (y<-X).
2. When we ignored the measurement error of x1, we obtained a path coefficient for y<-x1 of 3.55. Taking into account the measurement error, we obtain a coefficient of 7.1.
Model with two measurement variables and reliability

```
.semm(x1 x2<-X) (y<-X), reliability(x1 .5  x2 .5)
```

Endogenous variables
Measurement: x1 x2 y

Exogenous variables
Latent: X

Fitting target model:
Iteration 0: log likelihood = -16258.636
Iteration 1: log likelihood = -16258.401
Iteration 2: log likelihood = -16258.4

Structural equation model

|                     | OIM                          | Coefficient | std. err. | z      | P>|z|  | [95% conf. interval] |
|---------------------|------------------------------|-------------|-----------|--------|------|----------------------|
| **Measurement**     |                              |             |           |        |      |                      |
| x1                  |                              | 1 (constrained) |          |        |      |                      |
| _cons               |                              | 100.278     | .4037851  | 248.34 | 0.000| 99.48655 101.0694    |
| x2                  |                              | 1.030101    | .0417346  | 24.68  | 0.000| .9483029 1.1119      |
| _cons               |                              | 100.2066    | .4149165  | 241.51 | 0.000| 99.39342 101.0199    |
| y                   |                              | 7.031299    | .2484176  | 28.30  | 0.000| 6.544409 7.518188    |
| _cons               |                              | 701.081     | 2.042928  | 343.17 | 0.000| 697.077 705.0851     |
| var(e.x1)           |                              | 100.1036    |           |        |      |                      |
| var(e.x2)           |                              | 105.1719    |           |        |      |                      |
| var(e.y)            |                              | 152.329     | 105.26    |        |      | 39.31868 590.1553    |
| var(X)              |                              | 101.0907    | 7.343656  |        |      | 87.67509 116.5591    |

LR test of model vs. saturated: chi2(2) = 0.59 Prob > chi2 = 0.7430

Notes:

1. We wish to estimate the effect of y<-X. We have two measures of X—x1 and x2—both measured with error (0.50 reliability).
2. In the previous section, we used just x1. We obtained path coefficient 7.1 with standard error 0.4. Using both x1 and x2, we obtain path coefficient 7.0 and standard error 0.2.
3. We at StataCorp created these fictional data. The true coefficient is 7.

Also see

[SEM] `sem` and `gsem` option `reliability()` — Fraction of variance not due to measurement error
[SEM] `Example 1` — Single-factor measurement model