Mediation Analysis: OLS vs. SUR vs. ISUR vs. 3SLS vs. SEM
Note by Hubert Gatignon
July 7, 2013, updated November 15, 2013,
April 11, 2014, May 21, 2016 and August 10, 2016

In Chap. 11 of "Statistical Analysis of Management Data" (Gatignon, 2014), tests of mediation are presented using the SUR estimation method with estimation of indirect effects through the bootstrap procedure. In this additional note, the appropriateness of SUR versus OLS or 3SLS is discussed.

Case of single exogenous (iv) variable (with single or multiple mediators)

The model is:

$$iv \rightarrow med \rightarrow dv$$

In equation:

$$med_i = \alpha_0 + \alpha_1 iv_i + \varepsilon_{1i} \quad (1)$$

$$dv_i = \beta_0 + \beta_1 med_i + \varepsilon_{2i} \quad (2)$$

If $\rho(\varepsilon_{1i}, \varepsilon_{2i}) = 0$, then OLS and SUR provide the same estimates. Note that in practice, the estimate of the correlation will not be exactly zero and the estimates may vary slightly. This is possible because 1) the generalized least squares estimator is the same as the OLS estimator and 2) the independent variable in Eq. (2) is not correlated to the error term of the equation, which does not lead to a biased estimator.

The same reasoning applies to the extended second equation where the direct effect of *iv* is included:

$$dv_i = \beta_0 + \beta_1 med_i + \beta_2 iv_i + \varepsilon_{2i} \quad (3)$$

To illustrate this, we created a dataset generated by the following data generating functions:

$$\begin{aligned} med_i &= 12 - 0.6 iv_i + \varepsilon_{1i} \quad ; \quad \sigma_{\varepsilon_1} = 1.5 \\ dv_i &= 5 + 0.75 med_i + 0.1 iv_i + \varepsilon_{2i} \quad ; \quad \sigma_{\varepsilon_2} = 3 \end{aligned} \quad (4)$$

The OLS estimates are provided in Table 1.

Table 1 – OLS estimates

```
. use "/Users/gatignon/Documents/WORK_STATA/MediationData/mediationdata_singleiv_Rho_0.dta",
clear

. regress med iv

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  1,   398) =  263.44
       Model |  582.630641     1  582.630641           Prob > F      =  0.0000
    Residual |  880.233943   398  2.21164307           R-squared     =  0.3983
-------------+------------------------------           Adj R-squared =  0.3968
       Total |  1462.86458   399  3.66632728           Root MSE      =  1.4872

------------------------------------------------------------------------------
         med |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          iv |  -.6360555   .0391883   -16.23   0.000    -.7130973   -.5590137
       _cons |   12.30566   .3964391    31.04   0.000     11.52629    13.08504
------------------------------------------------------------------------------

. regress dv med iv

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  2,   397) =   56.21
       Model |  947.534487     2  473.767243           Prob > F      =  0.0000
    Residual |  3346.36766   397  8.42913767           R-squared     =  0.2207
-------------+------------------------------           Adj R-squared =  0.2167
       Total |  4293.90214   399  10.7616595           Root MSE      =  2.9033

------------------------------------------------------------------------------
          dv |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         med |   .8976328   .0978571     9.17   0.000     .7052498    1.090016
          iv |   .1641715   .0986263     1.66   0.097    -.0297237    .3580667
       _cons |   3.390934   1.431462     2.37   0.018     .5767419    6.205127
------------------------------------------------------------------------------
```

As can be seen, all parameter estimates are close to the true parameters. The SUR estimates are not needed in this case where the error terms across equations are uncorrelated. Because, a priori, this information is not known, it can be seen in Table 2 that the SUR parameter estimates are almost identical to the OLS estimates.

Table 2 – SUR Estimates

```
. sureg (med iv) (dv med iv), corr

Seemingly unrelated regression
----------------------------------------------------------------------
Equation          Obs  Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
med               400      1    1.483437    0.3983     264.76   0.0000
dv                400      2     2.89239    0.2207     113.26   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
med          |
          iv |  -.6360555   .0390902   -16.27   0.000    -.7126708   -.5594402
       _cons |   12.30566   .3954467    31.12   0.000      11.5306    13.08072
-------------+----------------------------------------------------------------
dv           |
         med |   .8976328   .0974895     9.21   0.000     .7065569    1.088709
          iv |   .1641715   .0982558     1.67   0.095    -.0284064    .3567493
       _cons |   3.390934   1.426084     2.38   0.017     .5958621    6.186007
------------------------------------------------------------------------------

Correlation matrix of residuals:

         med       dv
med   1.0000
 dv  -0.0000   1.0000

Breusch-Pagan test of independence: chi2(1) =     0.000, Pr = 1.0000
```

The estimated correlation matrix among the error terms indicates the absence of significant correlation, which makes the SUR estimation unnecessary and the OLS estimate appropriate.

If, however, $\rho(\varepsilon_{1i},\varepsilon_{2i}) \neq 0$, then Eq. (2) or Eq. (3) cannot be estimated through OLS because the coefficients are biased due to the correlation of the variable *med* with the error term. Taking equation (2) in its mean-centered form, the bias is determined by :

$$\hat{\beta}_{ols} = (\mathbf{med}'\mathbf{med})^{-1}\mathbf{med}'\mathbf{dv} \qquad (5)$$

and

$$E\left[\hat{\beta}_{ols}\right] = E\left[(\mathbf{med}'\mathbf{med})^{-1}\mathbf{med}'\mathbf{dv}\right] = (\mathbf{med}'\mathbf{med})^{-1} E\left[\mathbf{med}'\mathbf{dv}\right] = (\mathbf{med}'\mathbf{med})^{-1} E\left[\mathbf{med}'\mathbf{dv}\right] \quad (6)$$

However

$$E\left[\mathbf{med}'\mathbf{dv}\right] = E\left[(\mathbf{iv}\boldsymbol{\alpha}+\boldsymbol{\varepsilon}_1)'(\mathbf{med}\boldsymbol{\beta}+\boldsymbol{\varepsilon}_2)\right] = (\mathbf{iv}\boldsymbol{\alpha})'(\mathbf{med}\boldsymbol{\beta})E\left[\boldsymbol{\varepsilon}_1'\boldsymbol{\varepsilon}_2\right] \neq 0 \qquad (7)$$

More specifically :

$$E\left[\boldsymbol{\varepsilon}_1'\boldsymbol{\varepsilon}_2\right] = \rho \qquad (8)$$

This means that the OLS estimates as well as the SUR estimates are biased.
The correlation between the mediator and the error term biases not only the coefficient of the mediator, as shown above, but it biases all the model coefficients. This can be seen from the generalization of Eq. (5) to the vector of coefficients for the raw data (not mean centered).
Let

$$\underset{T\mathrm{x}3}{\mathbf{X}} = \begin{bmatrix} \underset{T\mathrm{x}1}{\iota} & \underset{T\mathrm{x}1}{med} & \underset{T\mathrm{x}1}{iv} \end{bmatrix} \qquad (9)$$

where ι is a column vector of ones.
Redefining **dv** as **y**, Eq. (3) in matrix notion becomes

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_2 \qquad (10)$$

The OLS estimator is

$$\hat{\boldsymbol{\beta}}_{ols} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \qquad (11)$$

Therefore,

$$E\left[\hat{\boldsymbol{\beta}}_{ols}\right] = (\mathbf{X}'\mathbf{X})^{-1}E\left[\mathbf{X}'\mathbf{y}\right] = (\mathbf{X}'\mathbf{X})^{-1}E\left[\mathbf{X}'(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}_2)\right] \qquad (12)$$

$$E\left[\hat{\boldsymbol{\beta}}_{ols}\right] = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}E\left[\mathbf{X}'\boldsymbol{\varepsilon}_2\right] \qquad (13)$$

Even if the error terms are uncorrelated to the unit vector and to **iv**, the bias due to the correlation between med and the error term is transferred to the other parameters through the full covariance matrix $(\mathbf{X}'\mathbf{X})^{-1}$, the second part of Eq. (13).

$$(\mathbf{X}'\mathbf{X})^{-1}E\left[\mathbf{X}'\boldsymbol{\varepsilon}_2\right] = (\mathbf{X}'\mathbf{X})^{-1}E\left[\begin{pmatrix} \iota' \\ \mathbf{med}' \\ \mathbf{iv}' \end{pmatrix}\boldsymbol{\varepsilon}_2\right]$$

$$= (\mathbf{X}'\mathbf{X})^{-1}E\left[\begin{pmatrix} \iota'\boldsymbol{\varepsilon}_2 \\ \mathbf{med}'\boldsymbol{\varepsilon}_2 \\ \mathbf{iv}'\boldsymbol{\varepsilon}_2 \end{pmatrix}\right] = (\mathbf{X}'\mathbf{X})^{-1}\begin{pmatrix} 0 \\ E\left[\mathbf{med}'\boldsymbol{\varepsilon}_2\right] \\ 0 \end{pmatrix} \neq 0 \qquad (14)$$

It should be noted that the $\mathbf{X'X}$ matrix is not diagonal because, by Eq. (1) the mediator is correlated to the independent variable.

This bias can be seen by performing on OLS estimation on data generated as above but this time with a correlation of error terms across equations of 0.8. Table 3 shows that the parameters for the second equation are far from the parameters in the data generating function, in particular for the mediating variable (2.46 vs. the true parameter value of 0.76).

Table 3 – OLS estimation of correlated error model

```
. use "/Users/gatignon/Documents/WORK_STATA/MediationData/mediationdata_singleiv_Rho_08.dta",
clear

. regress med iv

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  1,   398) =  361.38
       Model |  727.109382     1  727.109382           Prob > F      =  0.0000
    Residual |  800.793577   398  2.01204416           R-squared     =  0.4759
-------------+------------------------------           Adj R-squared =  0.4746
       Total |  1527.90296   399  3.82933072           Root MSE      =  1.4185

------------------------------------------------------------------------------
         med |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          iv |  -.6632829   .0348913   -19.01   0.000    -.7318773   -.5946885
       _cons |   12.52571   .3650182    34.32   0.000     11.80811    13.24332
------------------------------------------------------------------------------

. regress dv med iv

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  2,   397) =  682.87
       Model |  5278.07946     2  2639.03973           Prob > F      =  0.0000
    Residual |  1534.25683   397  3.86462676           R-squared     =  0.7748
-------------+------------------------------           Adj R-squared =  0.7736
       Total |  6812.33628   399  17.0735245           Root MSE      =  1.9659

------------------------------------------------------------------------------
          dv |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         med |   2.455388   .0694694    35.34   0.000     2.318814    2.591962
          iv |   1.106728   .0667944    16.57   0.000     .9754126    1.238043
       _cons |  -15.27156   1.006521   -15.17   0.000    -17.25034   -13.29278
------------------------------------------------------------------------------
```

The SUR estimation does not even pick up the correlated errors as shown in Table 4.

Table 4 – SUR Estimation and correlation matrix of error terms

```
. sureg (med iv) (dv med iv), corr

Seemingly unrelated regression
----------------------------------------------------------------------
Equation          Obs  Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
med               400      1    1.414915    0.4759     363.19   0.0000
dv                400      2     1.95848    0.7748    1376.06   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
med          |
          iv |  -.6632829    .034804   -19.06   0.000    -.7314975   -.5950683
       _cons |   12.52571   .3641045    34.40   0.000     11.81208    13.23935
-------------+----------------------------------------------------------------
dv           |
         med |   2.455388   .0692084    35.48   0.000     2.319742    2.591034
          iv |   1.106728   .0665435    16.63   0.000     .9763048     1.23715
       _cons |  -15.27156   1.002739   -15.23   0.000    -17.23689   -13.30623
------------------------------------------------------------------------------

Correlation matrix of residuals:

         med       dv
med   1.0000
 dv  -0.0000   1.0000

Breusch-Pagan test of independence: chi2(1) =     0.000, Pr = 1.0000
```

This is due to the fact that the correlation is picked up by the *iv* variable (with an estimated coefficient of 1.106 vs. the true parameter value of 0.1). Indeed, running a model that does not include *iv* in the second equation provides a correlation of the error terms of 0.44 (much closer to the true value of 0.8). These results are shown in Table 5.

Table 5 – SUR Estimation without *iv* variable and correlation matrix of error terms

```
. sureg (med iv) (dv med), corr

Seemingly unrelated regression
----------------------------------------------------------------------
Equation          Obs  Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
med               400      1    1.424719    0.4686     516.72   0.0000
dv                400      1    2.647516    0.5884     442.82   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
med          |
          iv |  -.7453697   .0327903   -22.73   0.000    -.8096375   -.6811019
       _cons |    13.3681   .3438565    38.88   0.000     12.69416    14.04205
-------------+----------------------------------------------------------------
dv           |
         med |   1.291937   .0613941    21.04   0.000     1.171607    1.412267
       _cons |   2.739641   .3734964     7.34   0.000     2.007602     3.47168
------------------------------------------------------------------------------

Correlation matrix of residuals:

        med      dv
med  1.0000
 dv  0.4411  1.0000

Breusch-Pagan test of independence: chi2(1) =    77.821, Pr = 0.0000
```

One can note that the bias due to the omitted variable *iv* is small in this case because the true parameter is only 0.1.

As a result of this analysis, we conclude that the mediator variable in Eq. (2) needs to be instrumented with an instrument that does not contain the correlated error term but that is highly correlated with this mediator.

One obvious candidate for instrument is the exogenous variable *iv* itself. Then Eq. (2) can be estimated by OLS or more efficiently by SUR :

$$\begin{cases} med_i = \alpha_0 + \alpha_1 iv_i + \varepsilon_{1i} \\ dv_i = \beta_0 + \beta_1 \hat{med}_i + \varepsilon_{2i} \end{cases} \quad (15)$$

This analysis is performed and the results are shown in Table 6.

Table 6 – Instrumental estimation

```
. regress med iv

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  1,   398) =  361.38
       Model |  727.109382     1  727.109382           Prob > F      =  0.0000
    Residual |  800.793577   398  2.01204416           R-squared     =  0.4759
-------------+------------------------------           Adj R-squared =  0.4746
       Total |  1527.90296   399  3.82933072           Root MSE      =  1.4185

------------------------------------------------------------------------------
         med |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          iv |  -.6632829   .0348913   -19.01   0.000    -.7318773   -.5946885
       _cons |   12.52571   .3650182    34.32   0.000     11.80811    13.24332
------------------------------------------------------------------------------

. predict double medhat
(option xb assumed; fitted values)

. sureg (med iv) (dv medhat), corr

Seemingly unrelated regression
----------------------------------------------------------------------
Equation          Obs  Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
med               400      1    1.414915    0.4759     363.19   0.0000
dv                400      1    3.988165    0.0661      28.30   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
med          |
          iv |  -.6632829    .034804   -19.06   0.000    -.7314975   -.5950683
       _cons |   12.52571   .3641045    34.40   0.000     11.81208    13.23935
-------------+----------------------------------------------------------------
dv           |
      medhat |   .7868274   .1479017     5.32   0.000     .4969453    1.076709
       _cons |    5.62835   .8690341     6.48   0.000     3.925075    7.331626
------------------------------------------------------------------------------

Correlation matrix of residuals:

        med      dv
med  1.0000
 dv  0.8711  1.0000

Breusch-Pagan test of independence: chi2(1) =   303.539, Pr = 0.0000


. mat list e(Sigma)

symmetric e(Sigma)[2,2]
           med         dv
med  2.0019839
 dv   4.915647  15.905462
```

The parameter estimates are extremely close to their true values, including in the second equation. The covariance of the error terms of 4.916 corresponds to a correlation of 0.87, which is significant and close to the true value of 0.8.

A problem arises, however, when attempting to estimate the model with the direct effect, i.e., Eq. (3). The reason is the perfect collinearity between **iv** and $\hat{\mathbf{med}}$.
Therefore, omitting *iv* in Eq. (3) leads also to a bias if $\beta_1 \neq 0$.

The analysis above highlights the need to incorporate covariates on the mediator and dependent equations. These covariates then act as exogenous variables. We discuss this case next.

Case of several exogenous (iv) variables

The case of several exogenous variables is similar. However, now it is possible to estimate a more complete model as long as one of the exogenous variables (i.e., one of the direct effects) is omitted from the last (*dv*) equation.

A new dataset was generated according to the equations in (16) with no correlation between the error terms:

$$\begin{aligned} med_i &= 12 - 0.6 iv_{1i} + 0.3 iv_{2i} + \varepsilon_{1i} \quad ; \quad \sigma_{\varepsilon_1} = 1.5 \\ dv_i &= 5 + 0.75 med_i + 0.1 iv_{1i} + \varepsilon_{2i} \quad ; \quad \sigma_{\varepsilon_2} = 3 \end{aligned} \tag{16}$$

The OLS and SUR estimations provide almost similar results, as the correlation is practically zero (i.e., -0.002) in Table 7.

Table 7 – Estimation of model with covariate (exogenous variable)

```
. use "/Users/gatignon/Documents/WORK_STATA/MediationData/mediationdata_MultipleIV_Rho_0.dta",
clear

. regress med iv1 iv2

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  2,   397) =  135.21
       Model |  676.580788     2  338.290394           Prob > F      =  0.0000
    Residual |  993.273804   397  2.50194913           R-squared     =  0.4052
-------------+------------------------------           Adj R-squared =  0.4022
       Total |  1669.85459   399  4.18509923           Root MSE      =  1.5818

------------------------------------------------------------------------------
         med |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         iv1 |  -.5840254   .0398751   -14.65   0.000    -.6624181   -.5056327
         iv2 |   .2797617   .0400665     6.98   0.000     .2009928    .3585307
       _cons |   12.00387   .5830706    20.59   0.000     10.85758    13.15016
------------------------------------------------------------------------------

. regress dv med iv1

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  2,   397) =   41.84
       Model |  903.523301     2  451.761651           Prob > F      =  0.0000
    Residual |  4287.03229   397    10.79857           R-squared     =  0.1741
-------------+------------------------------           Adj R-squared =  0.1699
       Total |   5190.5556   399  13.0089113           Root MSE      =  3.2861

------------------------------------------------------------------------------
          dv |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         med |     .75889   .0984002     7.71   0.000     .5654394    .9523406
         iv1 |    .043062   .1013104     0.43   0.671    -.1561099    .2422338
       _cons |   5.563675   1.695621     3.28   0.001     2.230156    8.897194
------------------------------------------------------------------------------

. sureg (med iv1 iv2) (dv med iv1), corr

Seemingly unrelated regression
----------------------------------------------------------------------
Equation          Obs  Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
med               400      2    1.575812    0.4052     272.47   0.0000
dv                400      2    3.273778    0.1741      84.90   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
med          |
         iv1 |  -.5840251   .0397252   -14.70   0.000    -.6618852   -.5061651
         iv2 |   .2797703   .0399159     7.01   0.000     .2015366     .358004
       _cons |   12.00378   .5808794    20.66   0.000     10.86528    13.14229
-------------+----------------------------------------------------------------
dv           |
         med |   .7626875   .0980303     7.78   0.000     .5705515    .9548234
         iv1 |   .0453152   .1009297     0.45   0.653    -.1525034    .2431337
       _cons |   5.507104   1.689248     3.26   0.001     2.196238     8.81797
------------------------------------------------------------------------------

Correlation matrix of residuals:

         med       dv
med   1.0000
 dv  -0.0021   1.0000

Breusch-Pagan test of independence: chi2(1) =     0.002, Pr = 0.9673

. mat list e(Sigma)

symmetric e(Sigma)[2,2]
            med          dv
med   2.4831845
```

```
dv  -.01058788   10.717581
```

As in the simple case of a single *iv*, the parameters are reproduced relatively accurately.

Another dataset was generated with the same parameter values as in Eq. (16) but with a correlation of error terms across equation for 0.8.

The OLS or SUR estimations clearly show biased parameter values for the second equation (Table 8). For example the coefficient of the mediator is 2.14 (OLS) and 1.95 (SUR) instead of the true value of 0.76.

Table 8 – OLS and SUR estimations

```
. use
"/Users/gatignon/Documents/WORK_STATA/MediationData/mediationdata_MultipleIV_Rho_08.dta",
clear

. regress med iv1 iv2

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  2,   397) =  142.33
       Model |  646.207781     2  323.10389            Prob > F      =  0.0000
    Residual |  901.250972   397  2.27015358           R-squared     =  0.4176
-------------+------------------------------           Adj R-squared =  0.4147
       Total |  1547.45875   399  3.87834274           Root MSE      =  1.5067

------------------------------------------------------------------------------
         med |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         iv1 |  -.5751035    .039274   -14.64   0.000    -.6523146   -.4978925
         iv2 |   .3489572   .0384289     9.08   0.000     .2734077    .4245067
       _cons |   11.16968   .5361259    20.83   0.000     10.11568    12.22368
------------------------------------------------------------------------------

. regress dv med iv1

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  2,   397) =  663.03
       Model |  5067.22472     2  2533.61236           Prob > F      =  0.0000
    Residual |  1517.04826   397  3.82128024           R-squared     =  0.7696
-------------+------------------------------           Adj R-squared =  0.7684
       Total |  6584.27298   399  16.5019373           Root MSE      =  1.9548

------------------------------------------------------------------------------
          dv |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         med |    2.14051   .0592518    36.13   0.000     2.024023    2.256997
         iv1 |   .9607916   .0606844    15.83   0.000     .8414886    1.080095
       _cons |  -15.93324   .9985147   -15.96   0.000    -17.89627    -13.9702
------------------------------------------------------------------------------

. sureg (med iv1 iv2) (dv med iv1), corr

Seemingly unrelated regression
----------------------------------------------------------------------
Equation          Obs  Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
med               400      2    1.503605    0.4156     312.38   0.0000
dv                400      2    1.971756    0.7638    1146.28   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
med          |
         iv1 |  -.5773227   .0391252   -14.76   0.000    -.6540067   -.5006387
         iv2 |   .3937164   .0377711    10.42   0.000     .3196865    .4677464
       _cons |   10.74818   .5308602    20.25   0.000     9.707716    11.78865
-------------+----------------------------------------------------------------
dv           |
         med |   1.953465   .0582376    33.54   0.000     1.839321    2.067608
         iv1 |   .8564574   .0602171    14.22   0.000      .738434    .9744807
       _cons |  -13.22935   .9849646   -13.43   0.000    -15.15984   -11.29885
------------------------------------------------------------------------------

Correlation matrix of residuals:

        med      dv
med  1.0000
 dv  0.1789  1.0000

Breusch-Pagan test of independence: chi2(1) =    12.799, Pr = 0.0003


. mat list e(Sigma)

symmetric e(Sigma)[2,2]
           med          dv
med  2.2531274
```

```
 dv  .52290004  3.7926206
```

When applying the instrumental approach, however, the parameters are recovered much more accurately. In Table 9, the mediator variable is predicted first using the two exogenous variables as predictors and the predicted value of the mediator is used in the second equation as an instrument for the mediator variable. SUR is used to account for the correlated errors across equations.

Table 9 – Instrumental estimation

```
. regress med iv1 iv2

      Source |       SS       df       MS              Number of obs =     400
-------------+------------------------------           F(  2,   397) =  142.33
       Model |  646.207781     2  323.10389           Prob > F      =  0.0000
    Residual |  901.250972   397  2.27015358           R-squared     =  0.4176
-------------+------------------------------           Adj R-squared =  0.4147
       Total |  1547.45875   399  3.87834274           Root MSE      =  1.5067

------------------------------------------------------------------------------
         med |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         iv1 |  -.5751035    .039274   -14.64   0.000    -.6523146   -.4978925
         iv2 |   .3489572   .0384289     9.08   0.000     .2734077    .4245067
       _cons |   11.16968   .5361259    20.83   0.000     10.11568    12.22368
------------------------------------------------------------------------------

. predict medhat
(option xb assumed; fitted values)

.
. sureg (med iv1 iv2) (dv medhat iv1), corr

Seemingly unrelated regression
----------------------------------------------------------------------
Equation          Obs  Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
med               400      2    1.501042    0.4176     286.80   0.0000
dv                400      2    3.971176    0.0419      17.51   0.0002
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
med          |
         iv1 |  -.5751035   .0391264   -14.70   0.000      -.65179   -.4984171
         iv2 |   .3489573   .0382845     9.11   0.000     .2739211    .4239935
       _cons |   11.16968   .5341116    20.91   0.000     10.12284    12.21652
-------------+----------------------------------------------------------------
dv           |
      medhat |   1.023148   .2902531     3.53   0.000     .4542622    1.592033
         iv1 |   .3375245   .1921007     1.76   0.079    -.0389859    .7140348
       _cons |   .2191332   4.323431     0.05   0.960    -8.254636    8.692903
------------------------------------------------------------------------------

Correlation matrix of residuals:

        med      dv
med  1.0000
 dv  0.8968  1.0000

Breusch-Pagan test of independence: chi2(1) =   321.701, Pr = 0.0000

. mat list e(Sigma)

symmetric e(Sigma)[2,2]
```

```
              med          dv
med   2.2531274
 dv   5.3457417   15.770238
```

These steps are performed automatically with a three stage least squares estimation and the results are identical as shown in Table 10.

Table 10 – Three-stage least squares estimates

```
. reg3 (med iv1 iv2) (dv med iv1)

Three-stage least-squares regression
----------------------------------------------------------------------
Equation          Obs  Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
med               400      2    1.501042    0.4176     286.80   0.0000
dv                400      2    2.681402    0.5632      38.41   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
med          |
         iv1 |  -.5751035   .0391264   -14.70   0.000      -.65179   -.4984171
         iv2 |   .3489572   .0382845     9.11   0.000      .273921    .4239934
       _cons |   11.16968   .5341116    20.91   0.000     10.12284    12.21652
-------------+----------------------------------------------------------------
dv           |
         med |   1.023147   .1959836     5.22   0.000     .6390266    1.407268
         iv1 |   .3375242   .1297095     2.60   0.009     .0832983    .5917501
       _cons |     .21914   2.919251     0.08   0.940    -5.502486    5.940766
------------------------------------------------------------------------------
Endogenous variables:  med dv
Exogenous variables:   iv1 iv2
------------------------------------------------------------------------------

. mat list e(Sigma)

symmetric e(Sigma)[2,2]
           med          dv
med  2.2531274
 dv  3.0404604   7.1899174
```

Note that the correlation estimate is slightly different with values of 0.89 and 0.76 (as obtained from the estimated covariance matrix Sigma) respectively for the estimates shown in Table 9 and in Table 10.

A final remark can be useful concerning the example found in http://www.ats.ucla.edu/stat/stata/faq/mulmediation.htm. The SUR estimates are reported as shown in Table 11.

Table 11 – SUR estimates of education data

```
. use http://www.ats.ucla.edu/stat/data/hsb2, clear
(highschool and beyond (200 cases))

. sureg (read math)(write math)(science read write math), corr

Seemingly unrelated regression
----------------------------------------------------------------------
Equation          Obs  Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
read              200      1    7.662848    0.4386     156.26   0.0000
write             200      1    7.437294    0.3812     123.23   0.0000
science           200      3    6.983853    0.4999     199.96   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
read         |
        math |    .724807   .0579824    12.50   0.000     .6111636    .8384504
       _cons |   14.07254   3.100201     4.54   0.000     7.996255    20.14882
-------------+----------------------------------------------------------------
write        |
        math |   .6247082   .0562757    11.10   0.000     .5144099    .7350065
       _cons |   19.88724   3.008947     6.61   0.000     13.98981    25.78467
-------------+----------------------------------------------------------------
science      |
        read |   .3015317   .0679912     4.43   0.000     .1682715     .434792
       write |   .2065257   .0700532     2.95   0.003     .0692239    .3438274
        math |   .3190094   .0759047     4.20   0.000      .170239    .4677798
       _cons |   8.407353   3.160709     2.66   0.008     2.212476    14.60223
------------------------------------------------------------------------------

Correlation matrix of residuals:

            read    write  science
   read   1.0000
  write   0.3187   1.0000
science  -0.0000   0.0000   1.0000

Breusch-Pagan test of independence: chi2(3) =    20.318, Pr = 0.0001

. mat list e(Sigma)

symmetric e(Sigma)[3,3]
               read        write      science
   read   58.719247
  write   18.164815   55.313342
science  -1.839e-13   3.930e-13    48.77421
```

However, it is not possible to apply the instrumental methods because the third equation is not identified. In this particularly example, the structure of the error terms covariance is bock diagonal where the error term in the third equation is uncorrelated with the error terms in the other two equations, as shown in the correlation matrix of residuals.

Consequently, the SUR estimation shown in Table 11 is not biased. This would not be generally the case, however, with correlated error terms.

In summary, this shows that identification is necessary in recursive systems when the equation errors are correlated and that instrumental variables can recover the true parameter estimates. However, this requires that covariates are included in the analysis. When the equation errors are not correlated, separate OLS estimation of each equation is possible.

SUR vs. ISUR

The iterative SUR estimation re-estimates the covariance matrix of error terms until the parameter estimates converge. These estimates are the maximum likelihood estimates.

We use the same model as in Table 9, where the SUR estimates using the instrumental variable approach are provided. The ISUR estimates are shown on the top part of Table 12. The log-likelihood is also provided upon request with the command "**display e(ll)**", or using the command "**estat ic**" that gives, in addition to the log-likelihood, Akaike's AIC critererion as well as the BIC criterion. These measures that be used to compare non-nested models since they correct the log-likelihood for the number of parameters estimated. The coefficient estimates are exactly the same as the SUR estimates. This is often the case in simple models that converge rapidly. However, the ISUR parameters (maximum likelihood) are asymptotically more efficient.

Table 12 – ISUR estimates and estimates using SEM

```
. *ISUR
. sureg (med iv1 iv2) (dv medhat iv1) ///
>          , isure corr nolog

Seemingly unrelated regression, iterated
--------------------------------------------------------------------------
Equation             Obs   Parms        RMSE    "R-sq"       chi2        P
--------------------------------------------------------------------------
med                  400       2    1.501042    0.4176     286.80   0.0000
dv                   400       2    3.971176    0.0419      17.51   0.0002
--------------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
med          |
         iv1 |  -.5751035   .0391264   -14.70   0.000      -.65179   -.4984171
         iv2 |   .3489573   .0382845     9.11   0.000     .2739211    .4239935
       _cons |   11.16968   .5341116    20.91   0.000     10.12284    12.21652
-------------+----------------------------------------------------------------
dv           |
      medhat |   1.023148   .2902531     3.53   0.000     .4542622    1.592033
         iv1 |   .3375245   .1921007     1.76   0.079    -.0389859    .7140348
       _cons |   .2191332   4.323431     0.05   0.960    -8.254636    8.692903
------------------------------------------------------------------------------

Correlation matrix of residuals:

        med      dv
med  1.0000
 dv  0.8968  1.0000

Breusch-Pagan test of independence: chi2(1) =   321.701, Pr = 0.0000

. display e(ll)
-1523.0545

. estat ic

Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |        Obs  ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
           . |        400         .  -1523.055       6    3058.109   3082.058
-----------------------------------------------------------------------------
               Note: N=Obs used in calculating BIC; see [R] BIC note.

. *SEM
. sem (med <- iv1 iv2) (dv <- medhat iv1) ///
> , cov(e.med*e.dv )

Endogenous variables

Observed:  med dv

Exogenous variables

Observed:  iv1 iv2 medhat

Fitting target model:
```

```
Iteration 0:   log likelihood =   2327.341
Iteration 1:   log likelihood =  2327.4698  (backed up)

Structural equation model                       Number of obs     =        400
Estimation method  = ml
Log likelihood     =  2327.4698

--------------------------------------------------------------------------------
               |                 OIM
               |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------+----------------------------------------------------------------
Structural     |
  med <-       |
           iv1 |  -.5751035   .0402632   -14.28   0.000    -.6540179   -.4961892
           iv2 |   .3489572   .0376988     9.26   0.000      .275069    .4228454
         _cons |   11.16968   .5238414    21.32   0.000     10.14297    12.19639
  -------------+----------------------------------------------------------------
  dv <-        |
           iv1 |   .3375242   .1864021     1.81   0.070    -.0278172    .7028657
        medhat |   1.023147   .2850764     3.59   0.000     .4644079    1.581887
         _cons |   .2191389   4.202358     0.05   0.958    -8.017331    8.455608
---------------+----------------------------------------------------------------
     var(e.med)|   2.253127   .1593219                      1.961535    2.588067
      var(e.dv)|   15.77024   1.115133                      13.72931    18.11456
---------------+----------------------------------------------------------------
cov(e.med,e.dv)|   5.345742   .4003458    13.35   0.000     4.561078    6.130405
--------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(2)   =      1.93, Prob > chi2 = 0.3800
```

SUR vs. SEM

The question we address in this section concerns the comparison of results obtained from SUR versus those obtained via SEM.

First, we will use the same model specification as above again, i.e., using the instrumental variable approach. The parameter estimated with SEM are shown at the bottom of Table 12. Again, the estimated coefficients are identical up to the sixth digit after the decimal point. Note that we requested the covariance between the error terms of the two equations to be estimated (this issue is discussed next). This covariance matrix is identical to the one obtained with the SUR or ISUR estimation methods. However, the log-likelihood that results from this estimation is different from the one obtained with the ISUR estimation (2327.4698 vs. -1523.0545). These two log-likelihood are not comparable because the likelihood functions is different with the two models. In the case of SEM, the likelihood function is relative to the likelihood under the assumption of the saturated model, which explains the positive value provided in the output for the SEM estimation. (See Gatignon, 2014, pages 79-82).

Not knowing a priori if the error terms of each equation are correlated or not, we have to assume that the model is identified, which is the only way to get proper parameter estimates of the covariances among the error terms.

We will now use a model specification slightly more complex with one independent variable (iv) that is the focus of the study, two mediating variables (med1 and med2) leading to two mediator equations, and two control variables (cv1 and cv2). Therefore, the model has three exogenous variables (iv, cv1 and cv2), and three endogenous variables (med1, med2 and dv). The data used for illustrating the difference between SUR with Instruments (or 3SLS) versus SEM contains a significant correlation (0.8) between the error terms of the second mediator equation and the dv equation). Note that SEM is used here only to reflect structural relationships, as no measurement model is specified and each variable is assumed to be observed without error, just as in the regression examples above; the estimation of an SEM model with both measurement and structural parameters is covered in Chapter 11 of Gatignon (2014).

First we apply the SUR estimation with Instrumental variables in the dv equation for which the code is shown in shown in Table 13:

Table 13 – SUR estimates with instrumented mediator variables-code

```
regress med1 iv cv1 cv2
predict med1hat
regress med2 iv cv1 cv2
predict med2hat
sureg (med1 iv cv1 cv2) (med2 iv cv1 cv2) (dv iv med1hat med2hat), corr
mat list e(Sigma)
```

The results are shown in Table 14:

Table 14 – SUR estimates with instrumented mediator variables-results

```
. regress med1 iv cv1 cv2

      Source |       SS           df       MS      Number of obs   =       400
-------------+----------------------------------   F(3, 396)       =    119.18
       Model |  760.269157         3  253.423052   Prob > F        =    0.0000
    Residual |  842.072787       396  2.12644643   R-squared       =    0.4745
-------------+----------------------------------   Adj R-squared   =    0.4705
       Total |  1602.34194       399   4.0158946   Root MSE        =    1.4582

------------------------------------------------------------------------------
        med1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          iv |   .4353785   .0360196    12.09   0.000     .3645649    .5061921
         cv1 |  -.5158117   .0387398   -13.31   0.000     -.591973   -.4396504
         cv2 |  -.0141616   .0335059    -0.42   0.673    -.0800332      .05171
       _cons |   10.87417   .6392709    17.01   0.000      9.61738    12.13096
------------------------------------------------------------------------------

. predict med1hat
(option xb assumed; fitted values)

. regress med2 iv cv1 cv2

      Source |       SS           df       MS      Number of obs   =       400
-------------+----------------------------------   F(3, 396)       =     49.32
       Model |  570.653089         3  190.217696   Prob > F        =    0.0000
    Residual |  1527.25292       396  3.85669928   R-squared       =    0.2720
-------------+----------------------------------   Adj R-squared   =    0.2665
       Total |  2097.90601       399  5.25790979   Root MSE        =    1.9638

------------------------------------------------------------------------------
        med2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          iv |   .5143321   .0485087    10.60   0.000     .4189653    .6096989
         cv1 |   .0337605    .052172     0.65   0.518    -.0688083    .1363293
         cv2 |   .2741464   .0451234     6.08   0.000     .1854351    .3628576
       _cons |   7.657081   .8609255     8.89   0.000     5.964525    9.349637
------------------------------------------------------------------------------

. predict med2hat
(option xb assumed; fitted values)

. sureg (med1 iv cv1 cv2) (med2 iv cv1 cv2) (dv iv med1hat med2hat), corr

Seemingly unrelated regression
----------------------------------------------------------------------
Equation             Obs   Parms        RMSE    "R-sq"       chi2        P
```

```
----------------------------------------------------------------------
med1                400       3    1.450925    0.4745     361.14   0.0000
med2                400       3    1.954004    0.2720     149.46   0.0000
dv                  400       3    2.388436    0.1846      90.53   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
med1         |
          iv |   .4353785   .0358391    12.15   0.000     .3651352    .5056218
         cv1 |  -.5158117   .0385456   -13.38   0.000    -.5913596   -.4402637
         cv2 |  -.0141617   .0333379    -0.42   0.671    -.0795028    .0511794
       _cons |   10.87417   .6360665    17.10   0.000     9.627501    12.12084
-------------+----------------------------------------------------------------
med2         |
          iv |   .5143321   .0482656    10.66   0.000     .4197334    .6089309
         cv1 |   .0337605   .0519105     0.65   0.515    -.0679822    .1355032
         cv2 |   .2741463   .0448972     6.11   0.000     .1861495    .3621431
       _cons |   7.657081   .8566101     8.94   0.000     5.978157    9.336006
-------------+----------------------------------------------------------------
dv           |
          iv |   .2742002   .1412007     1.94   0.052    -.0025482    .5509485
     med1hat |   .8005514   .1250243     6.40   0.000     .5555082    1.045595
     med2hat |  -.7266239   .2013998    -3.61   0.000     -1.12136   -.3318876
       _cons |   9.386481   2.471676     3.80   0.000     4.542085    14.23088
------------------------------------------------------------------------------

Correlation matrix of residuals:

         med1     med2       dv
med1   1.0000
med2  -0.0320   1.0000
  dv   0.4595   0.5215   1.0000

Breusch-Pagan test of independence: chi2(3) =   193.637, Pr = 0.0000

. mat list e(Sigma)

symmetric e(Sigma)[3,3]
            med1        med2          dv
med1    2.105182
med2  -.09064448   3.8181323
  dv   1.5923296   2.4337537   5.7046267
```

As shown earlier in this note, the estimated coefficients are the same as when using 3SLS, as they are shown in Table 15:

Table 15 – 3SLS estimates

```
. reg3 (med1 iv cv1 cv2) (med2 iv cv1 cv2) (dv iv med1 med2)

Three-stage least-squares regression
----------------------------------------------------------------------
Equation          Obs   Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
med1              400       3    1.450925    0.4745     361.14   0.0000
med2              400       3    1.954004    0.2720     149.46   0.0000
dv                400       3    3.187871   -0.4527      50.82   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
med1         |
          iv |   .4353785   .0358391    12.15   0.000     .3651352    .5056218
         cv1 |  -.5158117   .0385456   -13.38   0.000    -.5913596   -.4402637
         cv2 |  -.0141616   .0333379    -0.42   0.671    -.0795027    .0511795
       _cons |   10.87417   .6360665    17.10   0.000     9.627501    12.12084
-------------+----------------------------------------------------------------
med2         |
          iv |   .5143321   .0482656    10.66   0.000     .4197334    .6089309
         cv1 |   .0337605   .0519105     0.65   0.515    -.0679822    .1355032
         cv2 |   .2741464   .0448972     6.11   0.000     .1861495    .3621432
       _cons |   7.657081   .8566101     8.94   0.000     5.978156    9.336006
-------------+----------------------------------------------------------------
dv           |
          iv |   .2741999   .1884621     1.45   0.146    -.0951791    .6435789
        med1 |   .8005515   .1668713     4.80   0.000     .4734897    1.127613
        med2 |  -.7266235   .2688104    -2.70   0.007    -1.253482   -.1997647
       _cons |   9.386476   3.298973     2.85   0.004     2.920608    15.85234
------------------------------------------------------------------------------
Endogenous variables:  med1 med2 dv
Exogenous variables:   iv cv1 cv2
------------------------------------------------------------------------------

. mat list e(Sigma)

symmetric e(Sigma)[3,3]
            med1        med2          dv
med1    2.105182
med2  -.09064448   3.8181323
  dv  -.15884142   5.2806639   10.162523
```

When we use the SEM procedure, we obtain the estimates shown in Table 16:

Table 16 – SEM model and estimated parameters

```
. sem (med1 <- iv cv1 cv2) (med2 <- iv cv1 cv2) (dv <- iv med1 med2)

Endogenous variables

Observed:  med1 med2 dv

Exogenous variables

Observed:  iv cv1 cv2

Fitting target model:

Iteration 0:   log likelihood = -4921.2392
Iteration 1:   log likelihood = -4921.2392

Structural equation model                       Number of obs      =       400
Estimation method  = ml
Log likelihood     = -4921.2392

------------------------------------------------------------------------------
             |                 OIM
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Structural   |
  med1 <-    |
          iv |   .4353785   .0358391    12.15   0.000     .3651352    .5056218
         cv1 |  -.5158117   .0385456   -13.38   0.000    -.5913596   -.4402637
```

```
         cv2 |  -.0141616   .0333379    -0.42   0.671    -.0795027    .0511795
       _cons |   10.87417   .6360665    17.10   0.000     9.627501    12.12084
  -----------+----------------------------------------------------------------
  med2 <-    |
          iv |   .5143321   .0482656    10.66   0.000     .4197334    .6089309
         cv1 |   .0337605   .0519105     0.65   0.515    -.0679822    .1355032
         cv2 |   .2741464   .0448972     6.11   0.000     .1861495    .3621432
       _cons |   7.657081   .8566101     8.94   0.000     5.978156    9.336006
  -----------+----------------------------------------------------------------
  dv <-      |
        med1 |   .8359019   .0535579    15.61   0.000     .7309302    .9408735
        med2 |   .5369659    .045776    11.73   0.000     .4472465    .6266852
          iv |  -.3826769   .0581491    -6.58   0.000    -.4966471   -.2687068
       _cons |  -4.441153   .7549511    -5.88   0.000     -5.92083   -2.961476
-------------+----------------------------------------------------------------
  var(e.med1)|   2.105182   .1488588                      1.832739    2.418125
  var(e.med2)|   3.818132   .2699827                      3.324007    4.385711
    var(e.dv)|   3.495547   .2471725                      3.043169    4.015173
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(3)   =     80.88, Prob > chi2 = 0.0000

. estat gof

----------------------------------------------------------------------------
Fit statistic        |      Value   Description
---------------------+------------------------------------------------------
Likelihood ratio     |
          chi2_ms(3) |     80.876   model vs. saturated
            p > chi2 |      0.000
         chi2_bs(12) |    742.732   baseline vs. saturated
            p > chi2 |      0.000
----------------------------------------------------------------------------
```

The parameters for the effects of the mediators and the direct effect if iv on the dv are clearly different from those obtained using SUR with the instrumented mediators, and so are the variance of the error term of the third equation.

What is not clear is what assumption is used for the covariances among these three error terms. We then run an SEM model where the covariances are constrained to 0 and another one where they are specifically estimated. Tables 17 shows the results where the covariances are constrained to 0:

Table 17 – SEM model with covariances of error terms constrained to 0

```
. sem (med1 <- iv cv1 cv2) (med2 <- iv cv1 cv2) (dv <- iv med1 med2) ///
> , cov(e.med1*e.med2@0) cov(e.med1*e.dv@0) cov(e.med2*e.dv@0)

Endogenous variables

Observed:  med1 med2 dv

Exogenous variables

Observed:  iv cv1 cv2

Fitting target model:

Iteration 0:   log likelihood = -4921.2392
Iteration 1:   log likelihood = -4921.2392

Structural equation model                       Number of obs     =        400
Estimation method  = ml
Log likelihood     = -4921.2392

------------------------------------------------------------------------------
```

```
             |                 OIM
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Structural   |
  med1 <-    |
          iv |   .4353785   .0358391    12.15   0.000     .3651352    .5056218
         cv1 |  -.5158117   .0385456   -13.38   0.000    -.5913596   -.4402637
         cv2 |  -.0141616   .0333379    -0.42   0.671    -.0795027    .0511795
       _cons |   10.87417   .6360665    17.10   0.000     9.627501    12.12084
  -----------+----------------------------------------------------------------
  med2 <-    |
          iv |   .5143321   .0482656    10.66   0.000     .4197334    .6089309
         cv1 |   .0337605   .0519105     0.65   0.515    -.0679822    .1355032
         cv2 |   .2741464   .0448972     6.11   0.000     .1861495    .3621432
       _cons |   7.657081   .8566101     8.94   0.000     5.978156    9.336006
  -----------+----------------------------------------------------------------
  dv <-      |
        med1 |   .8359019   .0535579    15.61   0.000     .7309302    .9408735
        med2 |   .5369659    .045776    11.73   0.000     .4472465    .6266852
          iv |  -.3826769   .0581491    -6.58   0.000    -.4966471   -.2687068
       _cons |  -4.441153   .7549511    -5.88   0.000     -5.92083   -2.961476
-------------+----------------------------------------------------------------
  var(e.med1)|   2.105182   .1488588                      1.832739    2.418125
  var(e.med2)|   3.818132   .2699827                      3.324007    4.385711
    var(e.dv)|   3.495547   .2471725                      3.043169    4.015173
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(3)   =     80.88, Prob > chi2 = 0.0000

. estat gof

----------------------------------------------------------------------------
Fit statistic        |      Value   Description
---------------------+------------------------------------------------------
Likelihood ratio     |
          chi2_ms(3) |     80.876   model vs. saturated
            p > chi2 |      0.000
         chi2_bs(12) |    742.732   baseline vs. saturated
            p > chi2 |      0.000
----------------------------------------------------------------------------
```

These results are identical to those when nothing is specified about the covariances of the error terms, which shows that failure to specifically ask for their estimation assumes they are independent, which is not appropriate if this is not the case. Consequently the model should be specified with these covariances as parameters to be estimated. Table 18 shows the results where the covariances of the error terms are estimated:

Table 18 – SEM model with estimated covariances of error terms

```
. sem (med1 <- iv cv1 cv2) (med2 <- iv cv1 cv2) (dv <- iv med1 med2) ///
> , cov(e.med1*e.med2) cov(e.med1*e.dv) cov(e.med2*e.dv )

Endogenous variables

Observed:  med1 med2 dv

Exogenous variables

Observed:  iv cv1 cv2

Fitting target model:

Iteration 0:   log likelihood = -4880.8012
Iteration 1:   log likelihood = -4880.8012

Structural equation model                       Number of obs     =        400
Estimation method  = ml
Log likelihood     = -4880.8012
```

```
--------------------------------------------------------------------------------
                   |                 OIM
                   |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------------+------------------------------------------------------------
Structural         |
  med1 <-          |
                iv |   .4353785   .0358391    12.15   0.000     .3651352    .5056218
               cv1 |  -.5158117   .0385456   -13.38   0.000    -.5913596   -.4402637
               cv2 |  -.0141616   .0333379    -0.42   0.671    -.0795027    .0511795
             _cons |   10.87417   .6360665    17.10   0.000     9.627501    12.12084
  -----------------+------------------------------------------------------------
  med2 <-          |
                iv |   .5143321   .0482656    10.66   0.000     .4197334    .6089309
               cv1 |   .0337605   .0519105     0.65   0.515    -.0679822    .1355032
               cv2 |   .2741464   .0448972     6.11   0.000     .1861495    .3621432
             _cons |   7.657081   .8566101     8.94   0.000     5.978156    9.336006
  -----------------+------------------------------------------------------------
  dv <-            |
              med1 |   .8005515   .1668713     4.80   0.000     .4734897    1.127613
              med2 |  -.7266235   .2688104    -2.70   0.007    -1.253482   -.1997647
                iv |   .2741999   .1884621     1.45   0.146    -.0951791    .6435789
             _cons |   9.386476   3.298973     2.85   0.004     2.920608    15.85234
-------------------+------------------------------------------------------------
      var(e.med1)|   2.105182   .1488588                      1.832739    2.418125
      var(e.med2)|   3.818132   .2699827                      3.324007    4.385711
        var(e.dv)|   10.16252   2.918504                      5.788298    17.84236
-------------------+------------------------------------------------------------
cov(e.med1,e.med2)|  -.0906445   .1418281    -0.64   0.523    -.3686224    .1873334
  cov(e.med1,e.dv)|  -.1588414   .4172006    -0.38   0.703    -.9765395    .6588567
  cov(e.med2,e.dv)|   5.280664   1.101823     4.79   0.000     3.121131    7.440197
--------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(0)   =      0.00, Prob > chi2 =      .

. estat gof

----------------------------------------------------------------------------
Fit statistic        |      Value   Description
---------------------+------------------------------------------------------
Likelihood ratio     |
         chi2_ms(0) |      0.000   model vs. saturated
           p > chi2 |          .
        chi2_bs(12) |    742.732   baseline vs. saturated
           p > chi2 |      0.000
----------------------------------------------------------------------------
```

The coefficients obtained with this model specification are exactly the same as with the SUR using instrumented mediators and as in the 3SLS estimation. The variances and covariances are also exactly the same as those obtained in 3SLS.

In conclusion, SEM gives the same results but it is important to remember to free the covariances among the error terms to be estimated. One issue, however, that appears in these examples, concerns the measures of goodness of fit in the SEM approach. SEM estimation is an analysis of covariance structure where the data consists in the variances and covariances among the exogenous and endogenous variables. Consequently, in Structural Equation Models that do not include multiple item measures, the degrees of freedom are small at best. Our model where we estimate the covariance is in fact saturated and has no degrees of freedom at all. The significance of the models obtained from least squares estimations can be determined based on the usual degrees of freedom that are based on the number of observations, even if the R-squared are not interpretable for generalized least squares estimated models since they vary between minus infinity and 1 (Judge et al. 1985).

References

Gatignon, Hubert (2014), *Statistical Analysis of Management Data*, Springer Science+Business Media, LLC, New York, NY.

Judge, George G., William E. Griffiths, R. Carter Hill, Helmut Lütkepohl, and Tsoung-Chao Lee (1985), *The Theory and Practice of Econometrics,* 2nd Edition, New York, NY: John Wiley and Sons, Inc..

"Stata FAQ: How Can I Analyze Multiple Mediators in Stata?" (n.d.), <http://www.ats.ucla.edu/stat/stata/faq/mulmediation.htm> (Jul. 7, 2013).