Help File - hybreg

Table of Content

- Installation
  - 1. Setup Command
  - 2. Setup Menu
  - 3. Run Stata
  - 4. profile.do
- One Endogenous Variable
  - 1. IV
  - 2. GMM
    - GMM: onestep
    - GMM: twostep
    - GMM: igmm
  - 3. GEL
    - GEL: CUE
    - GEL: ET
    - GEL: EL
  - Error Cases
    - Case 1
    - Case 2
- Multiple Endogenous Variables
  - 1. IV
  - 2. GMM
    - GMM: onestep
    - GMM: twostep
    - GMM: igmm
  - 3. GEL
    - GEL: CUE
    - GEL: ET
    - GEL: EL
  - Error Cases
- Ramdon Data Generation
  - hybreg_data.dta
  - hybreg_multi.dta

Installation

1. Setup Command

This step adds `hybreg` command to Stata.

Copy the files in folder `Command_c_ado_personal` to your personal ado path, for example `C:\ado\personal`

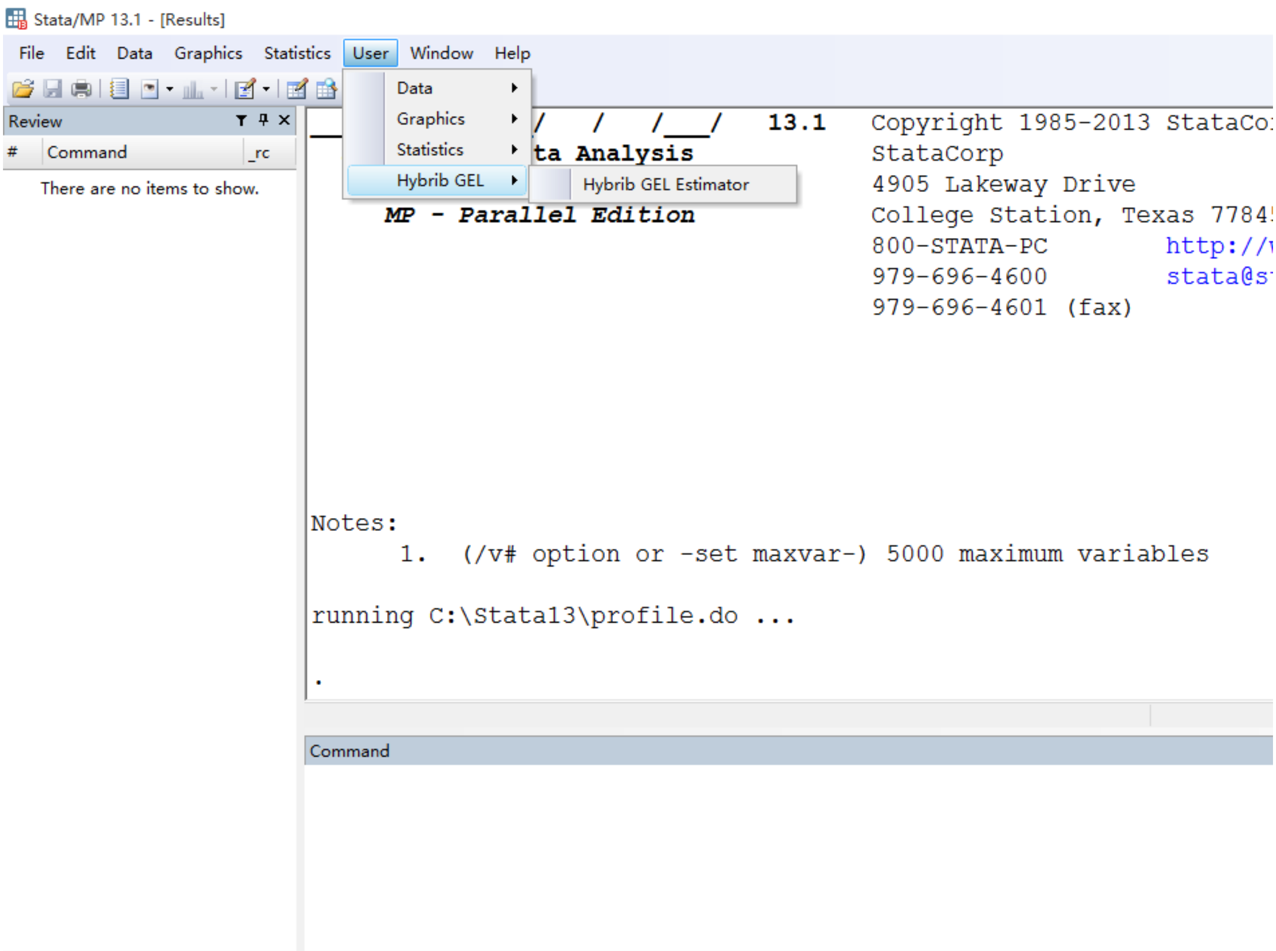
2. Setup Menu

This step adds menu items to Stata when Stata is launched.

Copy the files in folder `Menu_c_stata13` to the folder where Stata is installed, for example `C:\Stata13`

3. Run Stata

Run Stata. You may see a line: `running C:\Stata13\profile.do ...`



The `hybreg` menu is added to your Stata. You can also run `db hybreg` to launch the GUI.

4. profile.do

If you already have profile.ado predefined, you can add these lines to your profile.ado, to enable menu.

```
window menu clear
window menu append submenu "stUser" "Hybrib GEL"
window menu append item "Hybrib GEL" "Hybrib GEL Estimator" "db hybreg"
window menu refresh
```

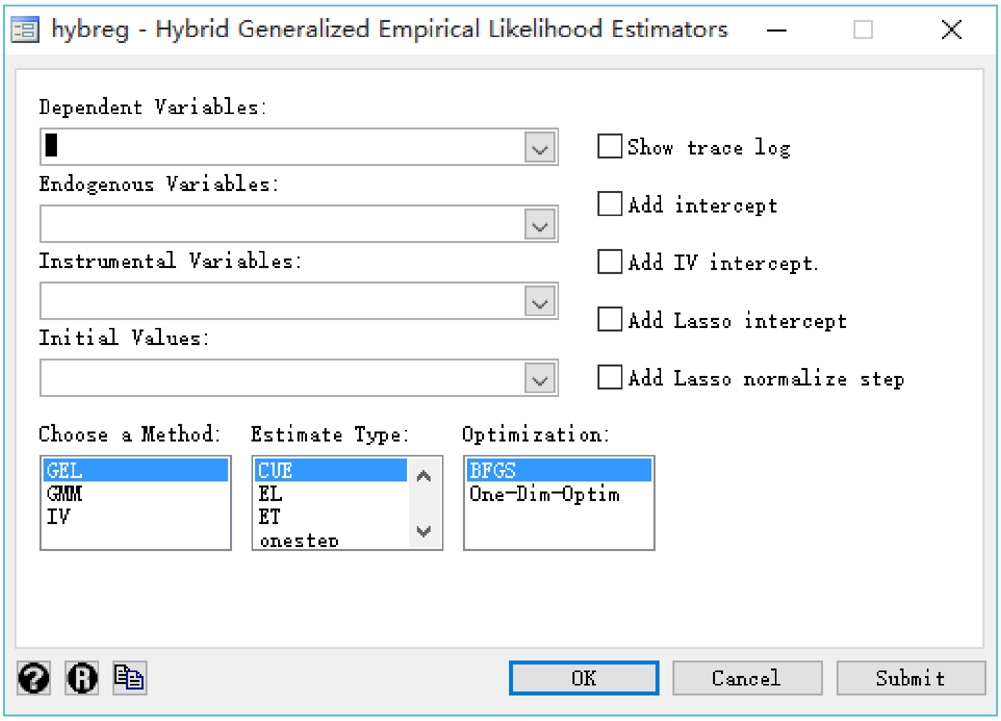
One Endogenous Variable

First we clear the data, stars are the comments in Stata

```
clear
```

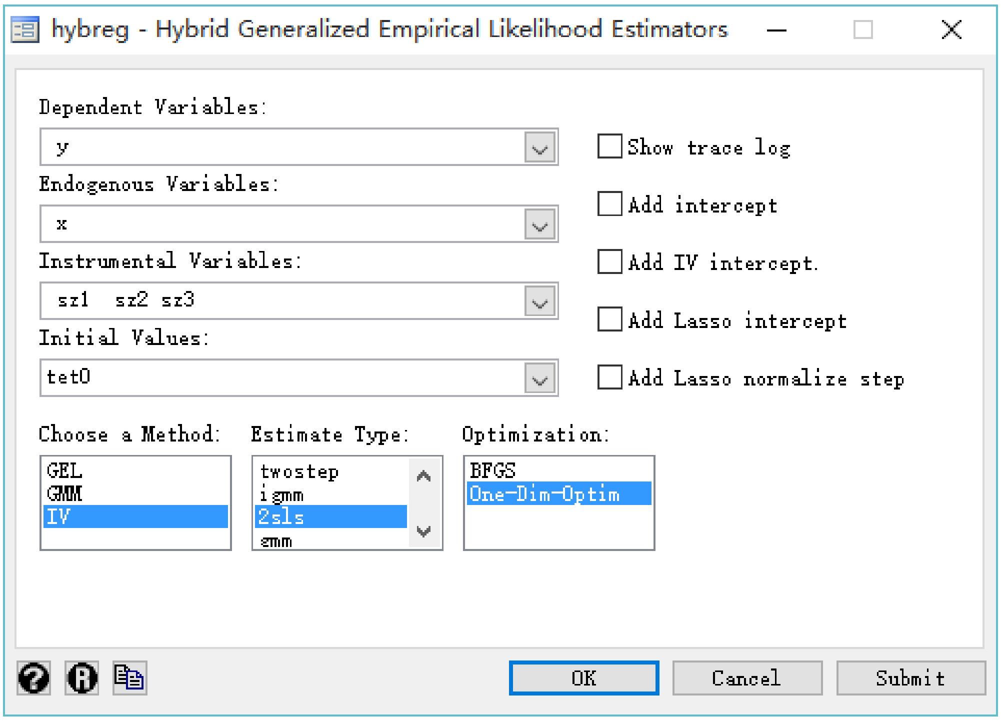
Command line is hard to use, the recommanded way to use hybreg is the GUI graphic interface. GUI will be able to generate the command line for you.

```
sysuse hybreg_Data

help hybreg

db hybreg
```



1. IV

This is the IV estimator. You can choose `2sls` , `liml` or `gmm` for the type of estimator.




```
hybreg y ( x = sz1 sz2 sz3 ), trace(0) method(IV) types(2sls) intercept(0)
tet0(tet0) lars_normalize(0) lars_intercept(0) optfct(optim)
```

Stata output :

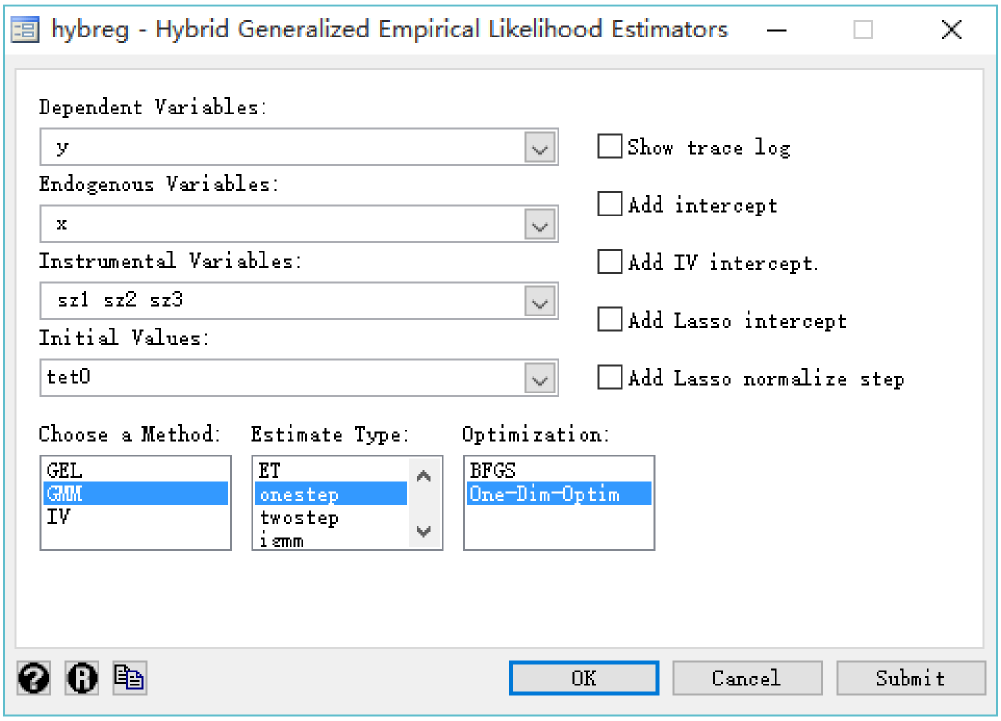
```
Lasso step:
Selected Instrumental Variables:  sz1
End of Lasso step.


Instrumental variables (2SLS) regression               Number of obs =     100
                                                       Wald chi2(1)  =       .
                                                       Prob > chi2   =       .
                                                       R-squared     =       .
                                                       Root MSE      =   4.279

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   .8614375   .4464976     1.93   0.054    -.0136816    1.736557
------------------------------------------------------------------------------
Instrumented:  x
Instruments:   sz1
```

other IV results:

```
hybreg y ( x = sz1 sz2  sz3 ), trace(0) method(IV) types(liml) intercept(0)
tet0(tet0) lars_normalize(0) lars_intercept(0) optfct(optim)

Lasso step:
Selected Instrumental Variables:  sz1
End of Lasso step.


Instrumental variables (LIML) regression               Number of obs =      100
                                                       Wald chi2(1)  =        .
                                                       Prob > chi2   =        .
                                                       R-squared     =        .
                                                       Root MSE      =    4.279

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   .8614375   .4464976     1.93   0.054    -.0136816    1.736557
------------------------------------------------------------------------------
Instrumented:  x
Instruments:   sz1


hybreg y ( x = sz1 sz2  sz3 ), trace(0) method(IV) types(gmm) intercept(0)
tet0(tet0) lars_normalize(0) lars_intercept(0) optfct(optim)

Instrumental variables (GMM) regression                Number of obs =      100
                                                       Wald chi2(1)  =        .
                                                       Prob > chi2   =        .
                                                       R-squared     =        .
GMM weight matrix: Robust                              Root MSE      =    4.279

------------------------------------------------------------------------------
             |               Robust
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   .8614375   .7173258     1.20   0.230    -.5444952     2.26737
------------------------------------------------------------------------------
Instrumented:  x
Instruments:   sz1
```

2. GMM

This is the GMM estimator, You can choose `onestep` , `twostep` or `igmm` for the type of estimator.



GMM: onestep

```
hybreg y ( x = sz1 sz2  sz3 ), trace(0) method(GMM) types(onestep) intercept(0)
tet0(tet0) lars_normalize(0) lars_intercept(0) optfct(optimize)

Lasso step:
Selected Instrumental Variables:  sz1
End of Lasso step.


Step 1
Iteration 0:   GMM criterion Q(b) =   .6815524
Iteration 1:   GMM criterion Q(b) =  1.071e-26
Iteration 2:   GMM criterion Q(b) =  5.835e-33

GMM estimation

Number of parameters =   1
Number of moments    =   1
Initial weight matrix: Unadjusted                     Number of obs   =      100

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        /b_x |   .8614375   .7173258     1.20   0.230    -.5444952     2.26737
------------------------------------------------------------------------------
Instruments for equation 1: sz1
```

GMM: twostep

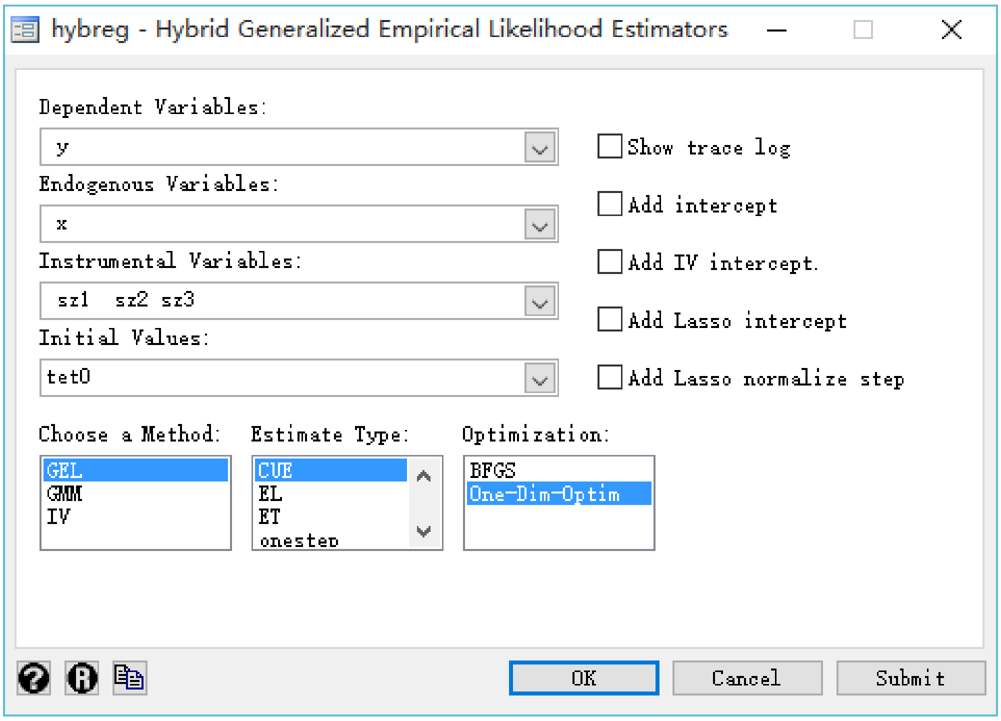
```
hybreg y ( x = sz1 sz2  sz3 ), trace(0) method(GMM) types(twostep) intercept(0)
tet0(tet0) lars_normalize(0) lars_intercept(0) optfct(optimize)

Lasso step:
Selected Instrumental Variables:  sz1
End of Lasso step.


Step 1
Iteration 0:   GMM criterion Q(b) =   .6815524
Iteration 1:   GMM criterion Q(b) =  1.071e-26
Iteration 2:   GMM criterion Q(b) =  5.835e-33

Step 2
Iteration 0:   GMM criterion Q(b) =  1.235e-34
Iteration 1:   GMM criterion Q(b) =  1.235e-34

GMM estimation

Number of parameters =   1
Number of moments    =   1
Initial weight matrix: Unadjusted                     Number of obs   =     100
GMM weight matrix:     Robust

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        /b_x |   .8614375   .7173258     1.20   0.230    -.5444952     2.26737
------------------------------------------------------------------------------
Instruments for equation 1: sz1
```

GMM: igmm

```
hybreg y ( x = sz1 sz2  sz3 ), trace(0) method(GMM) types(igmm) intercept(0)
tet0(tet0) lars_normalize(0) lars_intercept(0) optfct(optimize)

Lasso step:
Selected Instrumental Variables:  sz1
End of Lasso step.


Step 1
Iteration 0:   GMM criterion Q(b) =   .6815524
Iteration 1:   GMM criterion Q(b) =  1.071e-26
Iteration 2:   GMM criterion Q(b) =  5.835e-33

Step 2
Iteration 0:   GMM criterion Q(b) =  1.235e-34
Iteration 1:   GMM criterion Q(b) =  1.235e-34
iterative GMM weight matrix converged
iterative GMM parameter vector converged

GMM estimation

Number of parameters =   1
Number of moments    =   1
Initial weight matrix: Unadjusted                     Number of obs   =      100
GMM weight matrix:     Robust

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        /b_x |   .8614375   .7173258     1.20   0.230    -.5444952     2.26737
------------------------------------------------------------------------------
Instruments for equation 1: sz1
```

3. GEL

This is the GEL estimator, You can choose `CUE` , `EL` or `ET` for the type of estimator.


hybreg - Hybrid Generalized Empirical Likelihood Estimators
Dependent Variables:
y
Endogenous Variables:
x
Instrumental Variables:
sz1 sz2 sz3
Initial Values:
tet0
Show trace log
Add intercept
Add IV intercept.
Add Lasso intercept
Add Lasso normalize step
Choose a Method:
GEL
GMM
IV
Estimate Type:
CUE
EL
ET
onestep
Optimization:
BFGS
One-Dim-Optim
OK
Cancel
Submit

GEL: CUE

```
hybreg y ( x = sz1 sz2  sz3 ), trace(0) method(GEL) types(CUE) intercept(0)
tet0(tet0) lars_normalize(0) lars_intercept(0) optfct(optimize)

Type of GEL: CUE

Coefficients:

------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+-----------------------------------------------
          x |    .8613381    .7173808    1.200671     .229879
------------------------------------------------------------



Lambda:

------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+-----------------------------------------------
         h1 |   -2.92e-06           0           .           .
------------------------------------------------------------



Over-identifying restrictions tests: degrees of freedom is 0

------------------------------------
   Variable |  statistics     p-value
------------+-----------------------
    LR test |    1.59e-08           .
    LM test |    3.87e-08           .
    J test  |    1.92e-08           .
------------------------------------



Convergence code for the coefficients: There is no convergence code for optimize
```

GEL: ET

```
hybreg y ( x = sz1 sz2  sz3 ), trace(0) method(GEL) types(ET) intercept(0)
tet0(tet0) lars_normalize(0) lars_intercept(0) optfct(optimize)

Lasso step:
Selected Instrumental Variables:  sz1
End of Lasso step.

Type of GEL: ET

Coefficients:

------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+-----------------------------------------------
          x |     .859578     .718029    1.197136    .2312537
------------------------------------------------------------



Lambda:

------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+-----------------------------------------------
         h1 |    2.65e-14           0           .           .
------------------------------------------------------------



Over-identifying restrictions tests: degrees of freedom is 0

-------------------------------------
   Variable |  statistics     p-value
------------+------------------------
    LR test |   -7.55e-15           .
    LM test |    3.19e-24           .
    J test  |    6.71e-06           .
-------------------------------------



Convergence code for the coefficients: There is no convergence code for optimize
```

GEL: EL

```
hybreg y ( x = sz1 sz2  sz3 ), trace(0) method(GEL) types(EL) intercept(0)
tet0(tet0) lars_normalize(0) lars_intercept(0) optfct(optimize)


Lasso step:
Selected Instrumental Variables:  sz1
End of Lasso step.

Type of GEL: EL

Coefficients:

------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+-----------------------------------------------
          x |    .8613522    .7173686    1.200711    .2298634
------------------------------------------------------------



Lambda:

------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+-----------------------------------------------
         h1 |   -1.78e-06    1.56e-10   -11385.17           0
------------------------------------------------------------



Over-identifying restrictions tests: degrees of freedom is 0

------------------------------------
   Variable |  statistics     p-value
------------+-----------------------
    LR test |    1.41e-08           .
    LM test |    1.44e-08           .
    J test  |    1.41e-08           .
------------------------------------


Convergence code for the coefficients: There is no convergence code for optimize
```

Error Cases

Case 1

```
hybreg y ( x = sz1 sz2  sz3 ), trace(0) method(GEL) types(CUE) intercept(0)
tet0(tet0) lars_normalize(0) lars_intercept(0) optfct(optim)

Lasso step:
Selected Instrumental Variables:  sz1
End of Lasso step.



The length of initial value does not match the number of regressors
The length of initial value is 3
The the number of regressors is 1.
You may want to choose one dimensional optimization.

Only when the dimension of regressor is 1, you can choose between the algorithm BFGS
>  or one dimensional optimization. In that case, the former is unreliable. If one d
> imensional optimization is chosen, you should provide the upper and lower bound of
>  the interval in which the algorithm seeks the solution.r(232);
```

Only when the dimension of endogenous variable is 1, you can choose between the algorithm `optim` (BFGS algorithm) or `optimize` (One Dimensional Optimization). In that case, the former is unreliable. If optimize is chosen, you should provide the upper and lower bound of the interval in which the algorithm seeks the solution.

Case 2

```
hybreg y ( x =  sz3 ), trace(0) method(GEL) types(CUE) intercept(0) tet0(tet0)
lars_normalize(0) lars_intercept(0) optfct(optimize)

Lasso step:
Selected Instrumental Variables:
End of Lasso step.

Equation not identified, must have at least as many instruments not in the
regression.
There are 1 endogenous variable, 0 instrumented variables
r(481);
```

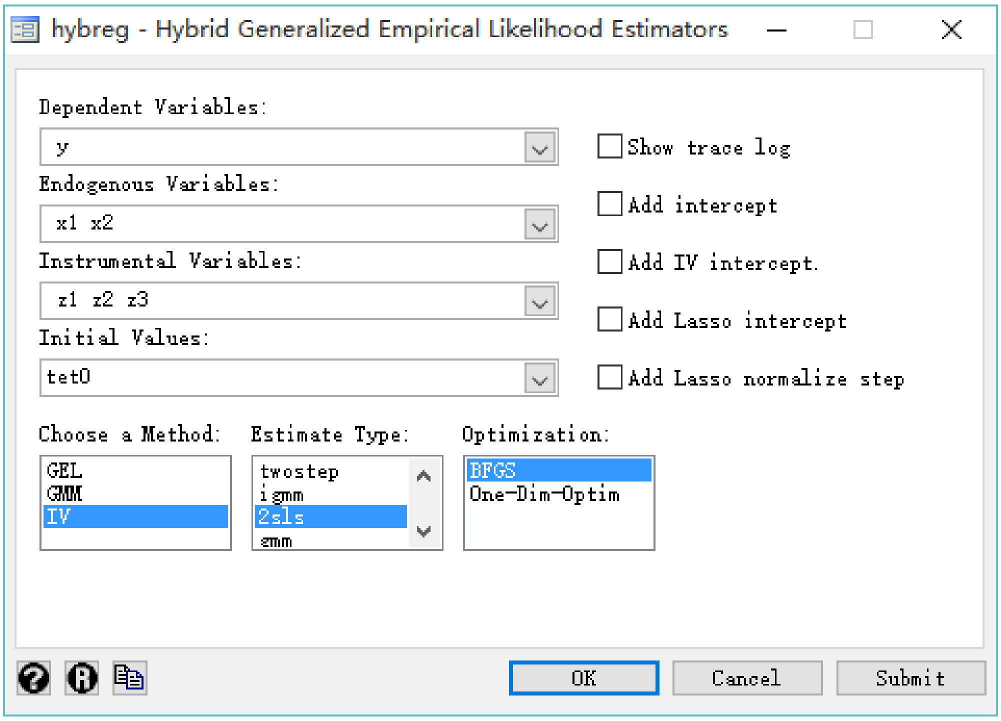
You should provide more instrumental variables.

Multiple Endogenous Variables

There are two endogenous variables. z1 z2 are correlated with x1, and z3 is correlated with x2.

```
sysuse hybreg_multi

generate var7 = 0.5 in 1
replace var7 = 0.5 in 2
rename var7 tet0

db hybreg
```

1. IV



```
hybreg y ( x1 x2 = z1 z2 z3 ), trace(0) method(IV) types(2sls) intercept(0) tet0(t
> et0) lars_normalize(0) lars_intercept(0) optfct(optim)



Lasso step:
Selected Instrumental Variables:  z1 z2 z3
End of Lasso step.

note: z3 dropped due to collinearity

Instrumental variables (2SLS) regression               Number of obs =      100
                                                       Wald chi2(2)  =        .
                                                       Prob > chi2   =        .
                                                       R-squared     =        .
                                                       Root MSE      =    10.63

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -2.474199   15.54964    -0.16   0.874    -32.95093    28.00253
          x2 |    7.23444   26.98233     0.27   0.789    -45.64996    60.11884
------------------------------------------------------------------------------
Instrumented:  x1 x2
Instruments:   z1 z2
```

2. GMM

GMM: onestep

```
hybreg y ( x1 x2 = z1 z2 z3 ), trace(0) method(GMM) types(onestep) intercept(0) te
> t0(tet0) lars_normalize(0) lars_intercept(0) optfct(optim)

Lasso step:
Selected Instrumental Variables:  z1 z2 z3
End of Lasso step.


Step 1
Iteration 0:   GMM criterion Q(b) =  21.065446
Iteration 1:   GMM criterion Q(b) =  .05915936
Iteration 2:   GMM criterion Q(b) =  .05915936

GMM estimation

Number of parameters =   2
Number of moments    =   3
Initial weight matrix: Unadjusted                    Number of obs   =     100

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       /b_x1 |   1.077991   .2798355     3.85   0.000     .5295239    1.626459
       /b_x2 |   1.066366   .3252707     3.28   0.001      .428847    1.703885
------------------------------------------------------------------------------
Instruments for equation 1: z1 z2 z3
```

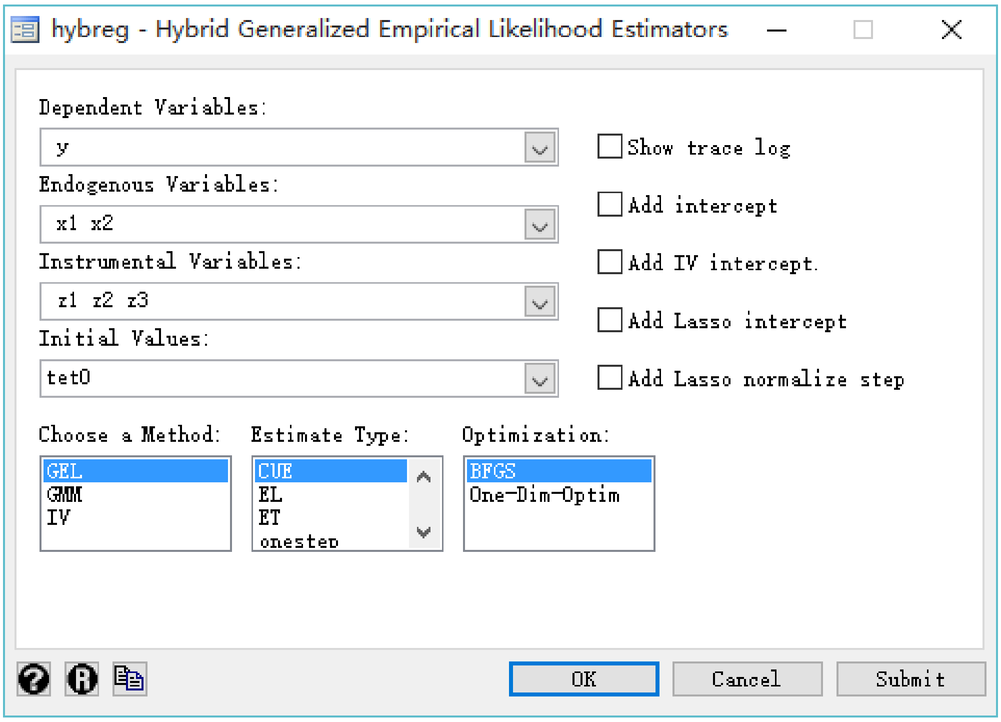
GMM: twostep

```
hybreg y ( x1 x2 = z1 z2 z3 ), trace(0) method(GMM) types(twostep) intercept(0) te
> t0(tet0) lars_normalize(0) lars_intercept(0) optfct(optim)


Lasso step:
Selected Instrumental Variables:  z1 z2 z3
End of Lasso step.


Step 1
Iteration 0:   GMM criterion Q(b) =  6.0499342
Iteration 1:   GMM criterion Q(b) =  .05915936
Iteration 2:   GMM criterion Q(b) =  .05915936

Step 2
Iteration 0:   GMM criterion Q(b) =  .00775748
Iteration 1:   GMM criterion Q(b) =  .00774086
Iteration 2:   GMM criterion Q(b) =  .00774086

GMM estimation

Number of parameters =   2
Number of moments    =   3
Initial weight matrix: Unadjusted                     Number of obs   =      100
GMM weight matrix:     Robust

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       /b_x1 |   1.066874   .2792201     3.82   0.000     .5196123    1.614135
       /b_x2 |    1.07838   .3246282     3.32   0.001     .4421205     1.71464
------------------------------------------------------------------------------
Instruments for equation 1: z1 z2 z3

.
```

GMM: igmm

```
hybreg y ( x1 x2 = z1 z2 z3 ), trace(0) method(GMM) types(igmm) intercept(0) tet0(
> tet0) lars_normalize(0) lars_intercept(0) optfct(optim)

Lasso step:
Selected Instrumental Variables:  z1 z2 z3
End of Lasso step.


Step 1
Iteration 0:   GMM criterion Q(b) =  21.065446
Iteration 1:   GMM criterion Q(b) =  .05915936
Iteration 2:   GMM criterion Q(b) =  .05915936

Step 2
Iteration 0:   GMM criterion Q(b) =  .00775748
Iteration 1:   GMM criterion Q(b) =  .00774086
Iteration 2:   GMM criterion Q(b) =  .00774086

Step 3
Iteration 0:   GMM criterion Q(b) =  .00774782
Iteration 1:   GMM criterion Q(b) =   .0077478

Step 4
Iteration 0:   GMM criterion Q(b) =  .00774844
Iteration 1:   GMM criterion Q(b) =  .00774844

Step 5
Iteration 0:   GMM criterion Q(b) =  .00774847
Iteration 1:   GMM criterion Q(b) =  .00774847
iterative GMM weight matrix converged
iterative GMM parameter vector converged

GMM estimation

Number of parameters =   2
Number of moments    =   3
Initial weight matrix: Unadjusted                     Number of obs   =     100
GMM weight matrix:     Robust

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       /b_x1 |   1.066662   .2791977     3.82   0.000     .5194451     1.61388
       /b_x2 |   1.078813   .3246118     3.32   0.001     .4425856    1.715041
------------------------------------------------------------------------------
Instruments for equation 1: z1 z2 z3
```

3. GEL



GEL: CUE

Stata result:

```
hybreg y ( x1 x2 = z1 z2 z3 ), trace(0) method(GEL) types(CUE) intercept(0) tet0(t
> et0) lars_normalize(0) lars_intercept(0) optfct(optim)
Dependent Variables: y
Endogenous Variables:  x1 x2
Instrumental Variables:  z1 z2 z3

Lasso step:
Selected Instrumental Variables:  z1 z2 z3
End of Lasso step.

Type of GEL: CUE

Coefficients:

-------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+------------------------------------------------
         x1 |    1.067199    .2788049    3.827761    .0001293
         x2 |    1.072571    .3136692    3.419432    .0006275
-------------------------------------------------------------



Lambda:

-------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+------------------------------------------------
         h1 |    .0397516    .0504093    .7885762    .4303597
         h2 |   -.0404333    .0520053   -.7774844     .436873
         h3 |    .0008132    .0024646     .329964    .7414272
-------------------------------------------------------------



Over-identifying restrictions tests: degrees of freedom is 1

-------------------------------------
   Variable |  statistics     p-value
------------+------------------------
    LR test |    .7740856    .3789561
    LM test |     .608162    .4354807
    J test  |     .985876    .3207524
-------------------------------------


Convergence code for the coefficients: Yes
```

GEL: ET

Stata result:

```
. hybreg y ( x1 x2 = z1 z2 z3 ), trace(0) method(GEL) types(ET) intercept(0) tet0(te
> t0) lars_normalize(0) lars_intercept(0) optfct(optim)
Dependent Variables: y
Endogenous Variables:  x1 x2
Instrumental Variables:  z1 z2 z3

Lasso step:
Selected Instrumental Variables:  z1 z2 z3
End of Lasso step.

Type of GEL: ET

Coefficients:

-------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+------------------------------------------------
         x1 |    1.065136    .2807192    3.794313    .0001481
         x2 |     1.07529    .3149465      3.4142    .0006397
-------------------------------------------------------------



Lambda:

-------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+------------------------------------------------
         h1 |    .0441171    .0493637     .893715    .3714744
         h2 |   -.0449138    .0509266   -.8819327    .3778132
         h3 |    .0008183    .0024135    .3390621    .7345629
-------------------------------------------------------------



Over-identifying restrictions tests: degrees of freedom is 1

-------------------------------------
   Variable |  statistics     p-value
------------+------------------------
    LR test |    .8306702    .3620788
    LM test |    .7817546    .3766051
     J test |    .9443374    .3311649
-------------------------------------



Convergence code for the coefficients: Yes
```

GEL: EL

This runs a little bit slower.

Stata result:

```
. hybreg y ( x1 x2 = z1 z2 z3 ), trace(0) method(GEL) types(EL) intercept(0) tet0(te
> t0) lars_normalize(0) lars_intercept(0) optfct(optim)
Dependent Variables: y
Endogenous Variables:  x1 x2
Instrumental Variables:  z1 z2 z3

Lasso step:
Selected Instrumental Variables:  z1 z2 z3
End of Lasso step.

numerical derivatives are approximate
nearby values are missing
.....
numerical derivatives are approximate
nearby values are missing

Type of GEL: EL

Coefficients:

------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+-----------------------------------------------
         x1 |    1.064391    .2828743     3.76277     .000168
         x2 |    1.075661    .3163445    3.400284    .0006732
------------------------------------------------------------



Lambda:

------------------------------------------------------------
   Variable |    Estimate     t value  Std. Error    Pr(>|t|)
------------+-----------------------------------------------
         h1 |    .0472081    .0484491    .9743853    .3298653
         h2 |   -.0480656     .049983   -.9616395    .3362307
         h3 |    .0007425    .0023687    .3134372    .7539485
------------------------------------------------------------



Over-identifying restrictions tests: degrees of freedom is 1

------------------------------------
   Variable |  statistics     p-value
------------+-----------------------
    LR test |    .8707327    .3507525
    LM test |    .9292014    .3350705
    J test  |    .9094023    .3402731
------------------------------------


Convergence code for the coefficients: Yes
```

Error Cases

```
hybreg y ( x1 x2 = z1 ), trace(0) method(GEL) types(ET) intercept(0) tet0(tet0) la
> rs_normalize(0) lars_intercept(0) optfct(optim)
Dependent Variables: y
Endogenous Variables:  x1 x2
Instrumental Variables:  z1

Lasso step:
equation not identified, must have at least as many instruments not in the regressio
> n as there are instrumented variables.
```

Ramdon Data Generation

hybreg_data.dta

There is only one endogenous variable, and z1 is correlated with x.

```
x = z1 + nu
eps = sqrt(z1^2+z2^2+z3^2)*eps
y = x + eps
```

hybreg_multi.dta

There are two endogenous variables. z1 z2 are correlated with x1, and z3 is correlated with x2.

```
z  = c(z1,z2,z3)
beta1 = c(1,1,0)
beta2 = c(0,0,1)
x1 = z %*% beta1 + nu
x2 = z %*% beta2 + nu
eps = sqrt(z1^2+z2^2+z3^2)*eps
y = x1 + x2 + eps
```