Quantitative Methods in Finance

Tutorial, Part 10:
Instrumental variables estimation.

Example 1: A consulting firm run by Mr. John Chardonnay is investigating the relative efficiency of wine production at 75 wineries. John sets up the production function:

$$Q = \beta_1 + \beta_2 MGT + \beta_3 CAP + \beta_4 LAB + u \ ,$$

where Q is an index of wine output for a winery taking into account both quantity and quality, MGT is a variable that reflects the efficiency of management, CAP is an index of capital input, and LAB is an index of labour input.

Because he cannot get data on management efficiency, John collects observations on the number of years of experience ($XPER$) of each winery manager and uses that variable in place of MGT. The 75 observations are stored in the data file `chard.dta`.

a) Estimate the revised equation using least squares and comment on the results.
b) John is concerned that the proxy variable $XPER$ might be correlated with the disturbance term[1]. He decides to perform the (Wu-)Hausman test using the manager's age (AGE) as an instrument for $XPER$. Regress $XPER$ on AGE, CAP, LAB, and save the residuals. Include these residuals as an extra variable in the equation you estimated in part a), and comment on the outcome of the test.
c) However, the (Wu-)Hausman test is reliable only if the instrument is valid. Check the validity of the two instrument conditions as possible.
d) Use the instrumental variables estimator to estimate the equation:

$$Q = \beta_1 + \beta_2 XPER + \beta_3 CAP + \beta_4 LAB + u \ ,$$

   with AGE as the instrumental variable for $XPER$. Comment on the results and compare them with those obtained in part a). Would you get the same results if you used the instrumental variable instead of the instrumented one?
e) How would you obtain the instrumental variables estimator manually using the 2SLS procedure? Compare the results with those from part d).
f) Calculate manually the regression coefficient in a bivariate instrumental variable regression model $Q = \beta_1 + \beta_2 XPER + u$.
g) Suppose John was not completely happy with using only AGE as an instrument for $XPER$, so he tried to include the square of AGE as an additional instrument. Test the overidentifying restrictions.

---

[1] Keep in mind that experience $XPER$ may be correlated with the disturbance term u, as, e.g., other aspects of management may be missing in the model presently that are relevant for wine output Q and being at the same time correlated with experience $XPER$. In addition, there may be other (non-management) determinants of wine output Q, not covered by the model presently, which are correlated with experience $XPER$.

Computer printout of the results in Stata:

a) OLS estimation

```
. regress q xper cap lab

      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  3,    71) =   30.31
       Model |  690.791296     3  230.263765           Prob > F      =  0.0000
    Residual |   539.35083    71  7.59649056           R-squared     =  0.5616
-------------+------------------------------           Adj R-squared =  0.5430
       Total |  1230.14213    74  16.6235422           Root MSE      =  2.7562

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        xper |   .1468382    .063433     2.31   0.024     .0203563      .27332
         cap |   .4379566   .1175603     3.73   0.000     .2035481    .6723652
         lab |   .2391613   .0998016     2.40   0.019     .0401627      .43816
       _cons |   1.762261   1.055347     1.67   0.099    -.3420408    3.866564
------------------------------------------------------------------------------
```

b) (Wu-)Hausman test

```
. regress xper age cap lab

      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  3,    71) =    4.78
       Model |  335.264398     3  111.754799           Prob > F      =  0.0043
    Residual |   1658.6556    71  23.3613465           R-squared     =  0.1681
-------------+------------------------------           Adj R-squared =  0.1330
       Total |     1993.92    74  26.9448649           Root MSE      =  4.8334

------------------------------------------------------------------------------
        xper |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .1661195   .0530281     3.13   0.003     .0603844    .2718546
         cap |   .4066095   .2070154     1.96   0.053    -.0061675    .8193865
         lab |  -.1150556   .1786841    -0.64   0.522    -.4713415    .2412304
       _cons |   4.715996    2.57399     1.83   0.071    -.4163945    9.848386
------------------------------------------------------------------------------

. predict xper_res, resid

. regress q xper cap lab xper_res

      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  4,    70) =   25.17
       Model |  725.606273     4  181.401568           Prob > F      =  0.0000
    Residual |  504.535852    70  7.20765503           R-squared     =  0.5899
-------------+------------------------------           Adj R-squared =  0.5664
       Total |  1230.14213    74  16.6235422           Root MSE      =  2.6847

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        xper |   .5121021   .1773102     2.89   0.005      .158468    .8657361
         cap |   .3321335   .1242232     2.67   0.009     .0843782    .5798889
         lab |   .2399754   .0972145     2.47   0.016     .0460871    .4338636
    xper_res |  -.4157506   .1891676    -2.20   0.031    -.7930335   -.0384677
       _cons |  -2.486688   2.189597    -1.14   0.260    -6.853701    1.880324
------------------------------------------------------------------------------
```

c) Instrument conditions

```
. regress xper age cap lab

      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  3,    71) =    4.78
       Model |  335.264398     3  111.754799           Prob > F      =  0.0043
    Residual |   1658.6556    71  23.3613465           R-squared     =  0.1681
-------------+------------------------------           Adj R-squared =  0.1330
       Total |     1993.92    74  26.9448649           Root MSE      =  4.8334

------------------------------------------------------------------------------
        xper |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .1661195   .0530281     3.13   0.003     .0603844    .2718546
         cap |   .4066095   .2070154     1.96   0.053    -.0061675    .8193865
         lab |  -.1150556   .1786841    -0.64   0.522    -.4713415    .2412304
       _cons |   4.715996    2.57399     1.83   0.071    -.4163945    9.848386
------------------------------------------------------------------------------
```

d) IV estimator

```
. ivregress 2sls q cap lab (xper=age)

Instrumental variables (2SLS) regression               Number of obs =      75
                                                       Wald chi2(3)  =   67.32
                                                       Prob > chi2   =  0.0000
                                                       R-squared     =  0.3568
                                                       Root MSE      =   3.248

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        xper |   .5121021   .2145152     2.39   0.017       .09166    .9325441
         cap |   .3321335    .150289     2.21   0.027     .0375726    .6266945
         lab |   .2399754    .117613     2.04   0.041     .0094581    .4704926
       _cons |  -2.486688   2.649039    -0.94   0.348     -7.67871    2.705334
------------------------------------------------------------------------------
Instrumented:  xper
Instruments:   cap lab age

. estat endog

  Tests of endogeneity
  Ho: variables are exogenous

  Durbin (score) chi2(1)          =  4.84123  (p = 0.0278)
  Wu-Hausman F(1,70)              =  4.83028  (p = 0.0313)

. regress q age cap lab

      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  3,    71) =   32.33
       Model |  710.207975     3  236.735992           Prob > F      =  0.0000
    Residual |   519.93415    71   7.3230162           R-squared     =  0.5773
-------------+------------------------------           Adj R-squared =  0.5595
       Total |  1230.14213    74  16.6235422           Root MSE      =  2.7061

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0850701   .0296895     2.87   0.005      .025871    .1442692
         cap |   .5403591   .1159041     4.66   0.000     .3092529    .7714652
```

```
         lab |   .1810552   .1000419     1.81   0.075    -.0184227     .380533
       _cons |   -.071617   1.441129    -0.05   0.961    -2.945147    2.801913
------------------------------------------------------------------------------
```

e) IV estimator - manual procedure

. regress xper age cap lab

```
      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  3,    71) =    4.78
       Model |  335.264398     3  111.754799           Prob > F      =  0.0043
    Residual |   1658.6556    71  23.3613465           R-squared     =  0.1681
-------------+------------------------------           Adj R-squared =  0.1330
       Total |     1993.92    74  26.9448649           Root MSE      =  4.8334

------------------------------------------------------------------------------
        xper |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .1661195   .0530281     3.13   0.003     .0603844    .2718546
         cap |   .4066095   .2070154     1.96   0.053    -.0061675    .8193865
         lab |  -.1150556   .1786841    -0.64   0.522    -.4713415    .2412304
       _cons |   4.715996    2.57399     1.83   0.071    -.4163945    9.848386
------------------------------------------------------------------------------
```

. predict xper_hat
(option xb assumed; fitted values)

. regress q xper_hat cap lab

```
      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  3,    71) =   32.33
       Model |  710.207985     3  236.735995           Prob > F      =  0.0000
    Residual |   519.93414    71  7.32301606           R-squared     =  0.5773
-------------+------------------------------           Adj R-squared =  0.5595
       Total |  1230.14213    74  16.6235422           Root MSE      =  2.7061

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    xper_hat |   .5121021   .1787235     2.87   0.005     .1557375    .8684667
         cap |   .3321335   .1252134     2.65   0.010     .0824651    .5818019
         lab |   .2399753   .0979894     2.45   0.017     .0445901    .4353606
       _cons |  -2.486689    2.20705    -1.13   0.264     -6.88742    1.914043
------------------------------------------------------------------------------
```

. ivregress 2sls q cap lab (xper=age), first

```
First-stage regressions
-----------------------

                                                  Number of obs   =        75
                                                  F(   3,     71) =      4.78
                                                  Prob > F        =    0.0043
                                                  R-squared       =    0.1681
                                                  Adj R-squared   =    0.1330
                                                  Root MSE        =    4.8334

------------------------------------------------------------------------------
        xper |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         cap |   .4066095   .2070154     1.96   0.053    -.0061675    .8193865
         lab |  -.1150556   .1786841    -0.64   0.522    -.4713415    .2412304
         age |   .1661195   .0530281     3.13   0.003     .0603844    .2718546
       _cons |   4.715996    2.57399     1.83   0.071    -.4163945    9.848386
------------------------------------------------------------------------------
```

```
Instrumental variables (2SLS) regression               Number of obs =      75
                                                       Wald chi2(3)  =   67.32
                                                       Prob > chi2   =  0.0000
                                                       R-squared     =  0.3568
                                                       Root MSE      =   3.248

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        xper |   .5121021   .2145152     2.39   0.017       .09166    .9325441
         cap |   .3321335    .150289     2.21   0.027     .0375726    .6266945
         lab |   .2399754    .117613     2.04   0.041     .0094581    .4704926
       _cons |  -2.486688   2.649039    -0.94   0.348     -7.67871    2.705334
------------------------------------------------------------------------------
Instrumented:  xper
Instruments:   cap lab age
```

f) Bivariate instrumental variable regression model

```
. correlate q xper age, cov
(obs=75)

             |        q     xper      age
-------------+---------------------------
           q |  16.6235
        xper |  7.16734  26.9449
         age |  9.88379  18.0746   117.24

. return list

scalars:
                  r(N) =  75
             r(cov_12) =  7.167344648648651
              r(Var_2) =  26.94486486486488
              r(Var_1) =  16.62354223664144

matrices:
                  r(C) :  3 x 3

. matrix C=r(C)
. matrix list C

symmetric C[3,3]
              q       xper        age
   q  16.623542
xper  7.1673446  26.944865
 age  9.8837877  18.074595  117.24036

. scalar cov_q_age=C[1,3]
. display cov_q_age
9.8837877

. scalar cov_xper_age=C[2,3]
. display cov_xper_age
18.074595

. scalar b_IV=cov_q_age/cov_xper_age
. display b_IV
.54683316
```

```
. ivregress 2sls q (xper=age)

Instrumental variables (2SLS) regression               Number of obs =      75
                                                       Wald chi2(1)  =    3.71
                                                       Prob > chi2   =  0.0541
                                                       R-squared     =       .
                                                       Root MSE      =  4.0765

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        xper |   .5468332   .2838817     1.93   0.054    -.0095647    1.103231
       _cons |   2.044378   3.968294     0.52   0.606    -5.733334    9.822091
------------------------------------------------------------------------------
Instrumented:  xper
Instruments:   age

. regress q xper

      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  1,    73) =    9.46
       Model |  141.082221     1  141.082221           Prob > F      =  0.0030
    Residual |   1089.0599    73  14.9186288           R-squared     =  0.1147
-------------+------------------------------           Adj R-squared =  0.1026
       Total |  1230.14213    74  16.6235422           Root MSE      =  3.8625

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        xper |   .2660004   .0864989     3.08   0.003     .0936083    .4383925
       _cons |   5.942337   1.280768     4.64   0.000      3.38977    8.494904
------------------------------------------------------------------------------

. regress q age

      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  1,    73) =    3.85
       Model |  61.6596979     1  61.6596979           Prob > F      =  0.0535
    Residual |  1168.48243    73  16.0066086           R-squared     =  0.0501
-------------+------------------------------           Adj R-squared =  0.0371
       Total |  1230.14213    74  16.6235422           Root MSE      =  4.0008

------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0843036   .0429532     1.96   0.053    -.0013019    .1699092
       _cons |   6.013863   1.901663     3.16   0.002     2.223854    9.803872
------------------------------------------------------------------------------
```

g) Extra instruments and the Sargan test

```
. gen age2=age^2

. ivregress 2sls q cap lab (xper=age age2)

Instrumental variables (2SLS) regression               Number of obs =      75
                                                       Wald chi2(3)  =   70.90
                                                       Prob > chi2   =  0.0000
                                                       R-squared     =  0.3956
                                                       Root MSE      =  3.1484
```

```
------------------------------------------------------------------------------
           q |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        xper |    .475636   .2062569     2.31   0.021     .0713798    .8798921
         cap |   .3426984   .1454797     2.36   0.018     .0575634    .6278333
         lab |   .2398941   .1140064     2.10   0.035     .0164457    .4633425
       _cons |  -2.062495   2.549404    -0.81   0.419    -7.059235    2.934246
------------------------------------------------------------------------------
Instrumented:  xper
Instruments:   cap lab age age2

. predict uhat, resid

. regress uhat cap lab age age2

      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  4,    70) =    0.46
       Model |  18.9418782     4  4.73546954           Prob > F      =  0.7666
    Residual |  724.507607    70  10.3501087           R-squared     =  0.0255
-------------+------------------------------           Adj R-squared = -0.0302
       Total |  743.449485    74  10.0466147           Root MSE      =  3.2172

------------------------------------------------------------------------------
        uhat |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         cap |  -.0181275   .1387993    -0.13   0.896    -.2949539    .2586989
         lab |   .0128411   .1196042     0.11   0.915     -.225702    .2513842
         age |  -.4006049   .3051015    -1.31   0.193    -1.009111    .2079009
        age2 |   .0047165   .0035148     1.34   0.184    -.0022936    .0117266
       _cons |   7.972852   6.364428     1.25   0.214    -4.720599     20.6663
------------------------------------------------------------------------------

. generate Sargan=e(N)*e(r2)
. display Sargan, invchi2tail(1, 0.05), chi2tail(1, Sargan)
1.9108775 3.8414588 .16686582

. quietly ivregress 2sls q cap lab (xper=age age2)
. estat overid

  Tests of overidentifying restrictions:

  Sargan (score) chi2(1) =  1.91088  (p = 0.1669)
  Basmann chi2(1)        =  1.83011  (p = 0.1761)
```

■

Example 2: The data file `fertility.dta` contains information on married women aged 21–35 with two or more children. We are interested in the relationship between:

– fertility (*morekids*), defined by a dummy variable that takes value 1 if the mother had more than two kids and 0 otherwise, and
– labour supply (*weeksm1*), defined as the number of mother's weeks worked in 1979.

a) Produce summary statistics and explain the relevant relationships among the variables.
b) Regress *weeksm1* on *morekids* and explain the results. Is this regression appropriate for estimating the causal effect[2] of fertility (*morekids*) on labour supply (*weeksm1*)?

[2] Keep in mind that both fertility (*morekids*) and labour supply (*weeksm1*) are choice variables. Mother with a positive labour supply disturbance, i.e. mother that worked more than average, may also be less likely to have an

c) Can we use the variable *samesex* as an instrument for the IV regression of *weeksm1* on *morekids*? Is it a proper instrument?
d) Estimate the regression of *weeksm1* on *morekids* using *samesex* as an instrument. Compare the results with the OLS estimates from part b).
e) Estimate the same regression manually using 2SLS procedure and compare the results.
f) Include also the variables *agem1*, *black*, *hispan*, *othrace*. Do the results change?

Computer printout of the results in Stata:

```
a) Relationships among the variables

. summarize

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
    morekids |     254654    .3805634    .4855263          0          1
      boy1st |     254654    .5143607    .4997947          0          1
      boy2nd |     254654    .5125504    .4998434          0          1
     samesex |     254654    .5055683      .49997          0          1
       agem1 |     254654    30.39327    3.386447         21         35
-------------+---------------------------------------------------------
       black |     254654    .0516623    .2213447          0          1
      hispan |     254654    .0742066    .2621073          0          1
     othrace |     254654    .0563431    .2305836          0          1
     weeksm1 |     254654    19.01833    21.86728          0         52

. summarize weeksm1, detail

                 Mom's weeks worked in 1979
-------------------------------------------------------------
      Percentiles      Smallest
 1%            0              0
 5%            0              0
10%            0              0       Obs             254,654
25%            0              0       Sum of wgt.     254,654

50%            5                      Mean           19.01833
                        Largest       Std. dev.      21.86728
75%           44             52
90%           52             52       Variance       478.1778
95%           52             52       Skewness       .5360685
99%           52             52       Kurtosis       1.524324

. inspect weeksm1

weeksm1:  Mom's weeks worked in 1979            Number of observations
------------------------------------   ---------------------------------------
                                             Total      Integers   Nonintegers
|  #                           Negative          -             -             -
|  #                           Zero        120,141       120,141             -
|  #                           Positive    134,513       134,513             -
|  #                                   -----------   -----------   -----------
|  #              #            Total       254,654       254,654             -
|  #   .   .   .  #            Missing           -
+----------------------                -----------
0                     52                   254,654
  (53 unique values)
```

---

additional child. This would imply that *morekids* is correlated with the regression disturbance term. Conversely, a mother with more than two kids is likely to work less weeks, implying simultaneity.

b) OLS estimation

. regress weeksm1 morekids

```
      Source |       SS           df       MS      Number of obs   =   254,654
-------------+----------------------------------   F(1, 254652)    =   3696.02
       Model |  1742078.14         1  1742078.14   Prob > F        =    0.0000
    Residual |   120027337   254,652  471.338679   R-squared       =    0.0143
-------------+----------------------------------   Adj R-squared   =    0.0143
       Total |   121769415   254,653  478.177816   Root MSE        =     21.71

------------------------------------------------------------------------------
     weeksm1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    morekids |  -5.386996   .0886093   -60.79   0.000    -5.560667   -5.213324
       _cons |   21.06843   .0546629   385.42   0.000     20.96129    21.17557
------------------------------------------------------------------------------
```

c) Choice of an instrument

. regress morekids samesex

```
      Source |       SS           df       MS      Number of obs   =   254,654
-------------+----------------------------------   F(1, 254652)    =   1237.22
       Model |  290.247937         1  290.247937   Prob > F        =    0.0000
    Residual |  59740.5888   254,652  .234596975   R-squared       =    0.0048
-------------+----------------------------------   Adj R-squared   =    0.0048
       Total |  60030.8368   254,653  .235735832   Root MSE        =    .48435

------------------------------------------------------------------------------
    morekids |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     samesex |   .0675253   .0019197    35.17   0.000     .0637626    .0712879
       _cons |   .3464248    .001365   253.79   0.000     .3437494    .3491002
------------------------------------------------------------------------------
```

. tabulate morekids samesex, chi2

```
 =1 if mom |
  had more |  =1 if 1st two kids
    than 2 |       same sex
      kids |         0          1 |     Total
-----------+----------------------+----------
         0 |    82,291     75,451 |   157,742 
         1 |    43,618     53,294 |    96,912 
-----------+----------------------+----------
     Total |   125,909    128,745 |   254,654 

          Pearson chi2(1) =  1.2e+03   Pr = 0.000
```

d) IV estimation

. ivregress 2sls weeksm1 (morekids=samesex)

```
Instrumental variables (2SLS) regression          Number of obs =  254654
                                                  Wald chi2(1)  =   24.54
                                                  Prob > chi2   =  0.0000
                                                  R-squared     =  0.0139
                                                  Root MSE      =  21.715
```

```
------------------------------------------------------------------------------
     weeksm1 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    morekids |  -6.313685   1.274599    -4.95   0.000    -8.811853   -3.815517
       _cons |   21.42109   .4869706    43.99   0.000     20.46665    22.37554
------------------------------------------------------------------------------
Instrumented:  morekids
Instruments:   samesex
```

e) IV estimation - manual estimation

```
. regress morekids samesex

      Source |       SS           df       MS      Number of obs   =   254,654
-------------+----------------------------------   F(1, 254652)    =   1237.22
       Model |  290.247937         1  290.247937   Prob > F        =    0.0000
    Residual |  59740.5888   254,652  .234596975   R-squared       =    0.0048
-------------+----------------------------------   Adj R-squared   =    0.0048
       Total |  60030.8368   254,653  .235735832   Root MSE        =    .48435

------------------------------------------------------------------------------
    morekids |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     samesex |   .0675253   .0019197    35.17   0.000     .0637626    .0712879
       _cons |   .3464248    .001365   253.79   0.000     .3437494    .3491002
------------------------------------------------------------------------------

. predict morekids_hat
(option xb assumed; fitted values)

. regress weeksm1 morekids_hat

      Source |       SS           df       MS      Number of obs   =   254,654
-------------+----------------------------------   F(1, 254652)    =     24.20
       Model |  11570.0433         1  11570.0433   Prob > F        =    0.0000
    Residual |   121757845   254,652  478.134259   R-squared       =    0.0001
-------------+----------------------------------   Adj R-squared   =    0.0001
       Total |   121769415   254,653  478.177816   Root MSE        =    21.866

------------------------------------------------------------------------------
     weeksm1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
morekids_hat |  -6.313684   1.283484    -4.92   0.000    -8.829278   -3.798091
       _cons |   21.42109   .4903651    43.68   0.000     20.45999    22.38219
------------------------------------------------------------------------------

. ivregress 2sls weeksm1 (morekids=samesex), first

First-stage regressions
-----------------------

                                                  Number of obs   =     254654
                                                  F(   1, 254652) =    1237.22
                                                  Prob > F        =     0.0000
                                                  R-squared       =     0.0048
                                                  Adj R-squared   =     0.0048
                                                  Root MSE        =     0.4844

------------------------------------------------------------------------------
    morekids |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     samesex |   .0675253   .0019197    35.17   0.000     .0637626    .0712879
       _cons |   .3464248    .001365   253.79   0.000     .3437494    .3491002
------------------------------------------------------------------------------
```

```
Instrumental variables (2SLS) regression               Number of obs =  254654
                                                       Wald chi2(1)  =   24.54
                                                       Prob > chi2   =  0.0000
                                                       R-squared     =  0.0139
                                                       Root MSE      =  21.715

------------------------------------------------------------------------------
     weeksm1 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    morekids |  -6.313685   1.274599    -4.95   0.000    -8.811853   -3.815517
       _cons |   21.42109   .4869706    43.99   0.000     20.46665    22.37554
------------------------------------------------------------------------------
Instrumented:  morekids
Instruments:   samesex
```

f) Additional exogenous explanatory variables

. ivregress 2sls weeksm1 (morekids=samesex) agem1 black hispan othrace

```
Instrumental variables (2SLS) regression               Number of obs =  254654
                                                       Wald chi2(5)  = 6677.36
                                                       Prob > chi2   =  0.0000
                                                       R-squared     =  0.0437
                                                       Root MSE      =  21.384

------------------------------------------------------------------------------
     weeksm1 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    morekids |  -5.821051   1.246295    -4.67   0.000    -8.263744   -3.378358
       agem1 |   .8315975   .0228862    36.34   0.000     .7867414    .8764536
       black |   11.62327   .2289286    50.77   0.000     11.17458    12.07197
      hispan |   .4041802   .2598548     1.56   0.120    -.1051259    .9134863
     othrace |   2.130962   .2058553    10.35   0.000     1.727493    2.534431
       _cons |  -4.791894   .4065695   -11.79   0.000    -5.588755   -3.995032
------------------------------------------------------------------------------
Instrumented:  morekids
Instruments:   agem1 black hispan othrace samesex
```

■