ECON4150 - Introductory Econometrics Seminar 8

Stock and Watson EE12.1

April 28, 2015

Empirical exercise E12.1: Data

- Fertility.dta contains data on 254,654 women between the age of 21 and 35 indoor workers taken from the 1980 Census.
- The data set contains information on the number of children the women had, their gender, the weeks worked by the mother and other characteristics.
- We are going to investigate the effect of fertility on female labor supply.
- How much a woman's labor supply fall when she has an additional children? behavior.

Empirical exercise E12.1: Data

Variable	**description**
morekids	=1 if mom had more than 2 children, =0 otherwise
boy1st	=1 if 1st child was a boy, =0 otherwise
boy2nd	=1 if 2nd child was a boy, =0 otherwise
samesex	=1 if 1st two children same sex, =0 otherwise
agem1	age of mom at census
black	=1 if mom is black, =0 otherwise
hispan	=1 if mom is Hispanic, =0 otherwise
othrace	=1 if mom is not black, Hispanic or white, =0 otherwise
weeksm1	mom's weeks worked in 1979

Empirical exercise E12.1: Data

```
clear all
set more off
cap log close
log using AE12_1.log , replace

cd M:\pc\Desktop\courses\introductory_econometrics\seminar_8

use "fertility.dta",clear
summ
```

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

a)

```
reg weeksm1 morekids, robust

Linear regression                               Number of obs =  254654
                                                F(  1,254652) = 3820.91
                                                Prob > F      =  0.0000
                                                R-squared     =  0.0143
                                                Root MSE      =   21.71

------------------------------------------------------------------------------
             |               Robust
     weeksm1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    morekids |  -5.386996   .0871491   -61.81   0.000    -5.557806   -5.216186
       _cons |   21.06843   .0560681   375.76   0.000     20.95854    21.17832
------------------------------------------------------------------------------
```

Women with more than 2 children work on average 5.39 fewer weeks per year than women with 2 or fewer children.

b)

- Possible omitted variable bias
- More educated women may both work more and be less likely to have an additional child than less educated women
- This would imply that a woman who works more than average may also be a woman who is less likely to have an extra child.
- In turn, this would implies that *Morekids* is positively correlated with the regression error so that the OLS estimator $\beta_{morekids}$ is positively biased

c)

```
reg morekids samesex, robust

Linear regression                               Number of obs =  254654
                                                F(  1,254652) = 1238.17
                                                Prob > F      =  0.0000
                                                R-squared     =  0.0048
                                                Root MSE      =  .48435

------------------------------------------------------------------------------
             |               Robust
    morekids |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     samesex |   .0675253    .001919    35.19   0.000     .0637641    .0712865
       _cons |   .3464248    .001341   258.34   0.000     .3437965    .3490531
------------------------------------------------------------------------------
```

couples with the first two kids of the same sex are 6.6that couples with the first two kids of different sex.

t-statistic$= 35.2 > 1.96$ so that the effect is statistically significant at a 5% significance level.

d) and (e)

- **Instrument relevance** $corr(morekids_i, samesex_i) \neq 0$. From the previous regression F= 1238.17, so that the instrument is relevant
- **Instrument Exogeneity** $samesex_i$ is uncorrelated with the error term $corr(samesex_i, u_i) = 0$ and has no direct effect on labor supply $weeksm1_i$. Plausible to think that this condition holds.

f)

```
 ivregress 2sls weeksm1 (morekids = samesex), robust

Instrumental variables (2SLS) regression          Number of obs =  254654
                                                  Wald chi2(1)  =   24.53
                                                  Prob > chi2   =  0.0000
                                                  R-squared     =  0.0139
                                                  Root MSE      =  21.715

------------------------------------------------------------------------------
             |               Robust
     weeksm1 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    morekids |  -6.313685   1.274681    -4.95   0.000    -8.812013   -3.815357
       _cons |   21.42109   .4872487    43.96   0.000      20.4661    22.37608
------------------------------------------------------------------------------
Instrumented:  morekids
Instruments:   samesex
```

The coefficient is -6.31, suggesting that that women with more than 2 children work 6.31 fewer weeks per year than women with 2 or fewer children.

g)

```
 ivregress 2sls weeksm1 agem1 black hispan othrace (morekids = samesex), robust

Instrumental variables (2SLS) regression              Number of obs =  254654
                                                      Wald chi2(5)  = 6954.98
                                                      Prob > chi2   =  0.0000
                                                      R-squared     =  0.0437
                                                      Root MSE      =  21.384

------------------------------------------------------------------------------
             |               Robust
     weeksm1 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    morekids |  -5.821051   1.246386    -4.67   0.000    -8.263923   -3.378179
       agem1 |   .8315975   .0226406    36.73   0.000     .7872228    .8759722
       black |   11.62327   .2317953    50.14   0.000     11.16896    12.07758
      hispan |   .4041802   .2607962     1.55   0.121     -.106971    .9153314
     othrace |   2.130962   .2109857    10.10   0.000     1.717438    2.544486
       _cons |  -4.791894   .3897868   -12.29   0.000    -5.555862   -4.027925
------------------------------------------------------------------------------
Instrumented:  morekids
Instruments:   agem1 black hispan othrace samesex

log close
```

There are no important changes in the result. This fact suggests that all other regressors are uncorrelated to *samesex*, so that there is no omitted variable bias in IV regression in question (f)