

```

***** Econ 651: Assignment 3 *****
***** Stata Solutions - Andrew Proctor *****

***** Preliminaries
capture log close _all // Closes any log if open //
set more off
set varabbrev off
*Install ivreg2 if necessary

* Set working dir
cd "C:\Users\AN.4271\Dropbox\HHS 651\Assignments\Assignment 3\"

* Initialize log
log using "Assignment3log", text replace

* Import data
import delimited using "m_d_806.tab", delimiter(tab) clear

***** variables
**** Rename variables
rename yobm yob_moth
rename agem age_mother
rename weeksm wks_wrked_moth
rename weeksd wks_wrked_fath
rename ageqk ageof1stchild
rename hoursd hourswked_fath
rename hoursm hourswked_moth
rename incomelm labinc_moth
rename income2m selfempinc_moth
rename incomeld labinc_fath
rename income2d selfempinc_fath

**** Label variables
label variable wks_wrked_moth "Weeks worked (moth)"
label variable wks_wrked_fath "Weeks worked (fath)"
label variable labinc_moth "Mother's labor income"
label variable selfempinc_moth "Mother's self-employment income"
label variable labinc_fath "Father's labor income"
label variable selfempinc_fath "Father's self-employment income"
label variable age_mother "Mother's age"
label variable hourswked_moth "Mother's hours worked"

**** Recode missing values and reform string Variables
*** Identify string variables
describe
codebook ageq2nd ageq3rd ageq4th ageq5th, tab(100)
foreach variable in ageq2nd ageq3rd ageq4th ageq5th {
    replace `variable' = "." if `variable'=="NA"
    destring `variable', replace
}

replace agemar = . if agemar == 0

***** Generate Timing of Birth and Marriage Variables (to determine
***** whether parents were Married)

*** Recode quarter Married as 0-3 instead of 1-4
replace qtrmar = qtrmar - 1 if ((qtrmar > 0) & !missing(qtrmar))

*** Code year of marriage
gen yr_Married = .
replace yr_Married = yob_moth + agemar if ((qtrbthm<=qtrmar) & ///
!missing(qtrmar) & !missing(qtrbthm))
replace yr_Married = yob_moth + agemar + 1 if ((qtrbthm>qtrmar) ///
& !missing(qtrmar) & !missing(qtrbthm))

*** Code year&quarter of marriage

```

```

gen yr_qtr_Married= .
replace yr_qtr_Married = yr_Married+(qtrmar/4) if ///
  (!missing(qtrmar) & !missing(yr_Married))

*** Code year&quarter of first birth
gen yr_qtr_birth = yobk+((qtrbkid)/4)

/* Generate indicator for if parents were not Married when
   first-birth occurred */
gen unMarriedbirth = 0
replace unMarriedbirth = 1 if ((yr_qtr_Married-yr_qtr_birth >0) ///
  & !missing(yr_qtr_Married) & !missing(yr_qtr_birth))

**** Sex of Child Indicators
gen boy1st = (sexk==0)
label variable boy1st "First child is a boy"

gen boy2nd = (sex2nd==0)
label variable boy2nd "Second child is a boy"

gen boys2 = ((sexk==0) & (sex2nd==0))
label variable boys2 "First two children boys"

gen girls2 = ((sexk==1) & (sex2nd==1))
label variable boys2 "First two children girl"

gen samesex = ((boys2==1) | (girls2==1))
label variable samesex "First two children same sex"

gen morekids = (kidcount>2) if !missing(kidcount)
label variable morekids "Parents had more than 3 children"

**** Race Indicators
gen black_mother = (racem==2)
label variable black_mother "Black / African-American (Mother)"

gen hisp_moth = (racem==12)
label variable hisp_moth "Hispanic / Latina (Mother)"

gen white_moth = (racem==1)
label variable white_moth "White (Mother)"

gen othrace_moth = 1 - black_moth - hisp_moth - white_moth
label variable othrace_moth "Other Race / Ethnicity (Mother)"

**** Mother and Father Labor Supply Variables
scalar inflationfactor = 245.519 / 82.4
  // From CPI-U data on BLS website

gen workedind_moth = (wks_wrked_moth>0) if !missing(wks_wrked_moth)
label variable workedind_moth "Mother worked"

gen workedind_fath = (wks_wrked_fath>0) if !missing(wks_wrked_fath)
label variable workedind_fath "Father worked"

gen totalinc_moth =labinc_moth+max(0,selfempinc_moth)
gen totalinc_fath =labinc_fath+max(0,selfempinc_fath)

replace totalinc_moth = totalinc_moth*inflationfactor
replace totalinc_fath = totalinc_fath*inflationfactor

label variable totalinc_moth "Total Income (mother)"
label variable totalinc_fath "Total Income (father)"

gen totalinc_fam =faminc *inflationfactor
label variable totalinc_fam "Total Income (family)"

```

```

gen logincfam =log(max(totalinc_fam ,1))
label variable logincfam "Log of total income (family)"

gen income_nonmoth =totalinc_fam -labinc_moth*inflationfactor
replace income_nonmoth=log(max(1,income_nonmoth ))
label variable income_nonmoth "Total income besides mother"

***** Create variables for sample
**** Ages of Mother and Father and ages when 1st kid was born
*** Generate "year of birth" for father
gen yob_fath = 79 - aged
replace yob_fath= 80 - aged if qtrbthd==0

*** Generate age in quarters for parents
gen ageqm=4*(80-yob_moth)-qtrbthm-1
gen ageqd=4*(80-yob_fath)-qtrbthd

*** Generate age of parents at birth of first child
gen agelstbth_moth =floor((ageqm-ageof1stchild)/4)
label variable agelstbth_moth "Age of mother at birth of first child"
gen agelstbth_fath =floor((ageqd-ageof1stchild)/4)

*** Main Sample
gen Main = 0
replace Main =1 if (inrange(age_mother,21,35) & (kidcount >= 2) & ///
!missing(kidcount) & (ageq2nd > 4) & !missing(ageq2nd) ///
& (agelstbth_moth>=15) & !missing(agelstbth_moth) ///
& asex==0 & aage==0 & aqtrbrth==0 & asex2nd==0 & aage2nd==0 ///
& aqtrbrth==0)
label variable Main "Main sample"

*** Married Sample
gen Married = 0
replace Married = 1 if (!missing(aged) & (timesmar==1) & ///
(marital==0) & (unMarriedbirth==0) & (agelstbth_fath >=15) & ///
(agelstbth_moth>=15) & !missing(agelstbth_fath) & ///
!missing(agelstbth moth) & (Main==1))
label variable Married "Married sample"

***** Descriptive Statistics
/* Reproduce summary statistics for variables in Table 2, cols 1-2 for
Main sample and Married sample */
summarize kidcount morekids boy1st boy2nd boys2 girls2 samesex age_mother ///
agelstbth_moth agelstbth_fath workedind_moth workedind_fath ///
wks_wrked_moth wks_wrked_fath hourswked_moth hourswked_fath ///
totalinc_moth totalinc_fath faminc logincfam income_nonmoth ///
if Main==1

summarize kidcount morekids boy1st boy2nd girls2 samesex age_mother ///
agelstbth_moth agelstbth_fath workedind_moth workedind_fath ///
wks_wrked_moth wks_wrked_fath hourswked_moth hourswked_fath ///
totalinc_moth totalinc_fath faminc logincfam income_nonmoth ///
if Married==1

***** Regressions

/* Note: Easily the most succinct way to run these regression
is using local macros and loops, which I do here. The long
way without these features is included at the very end of the
file. */

*** Macro to store the dependent variables
local dependvars workedind_moth wks_wrked_moth hourswked_moth ///
totalinc_moth logincfam

```

```

*** Macro for the control variables
local controls age mother agelstbth_moth boy1st boy2nd black_mother ///
    hisp_moth othrace_moth

** Perform the same regressions for each the Main and Married sample
foreach sample in Main Married {
    if ("`sample'"=="Married") local dependvars `dependvars' income_nonmoth

    eststo clear
    * Perform the OLS regression for each of the dep. variables.
    foreach outcome in `dependvars' {
        display
        "*****"
        display "                Regression: OLS - " `"`: var label
        `outcome'"" " - `sample' Sample "
        display
        "*****"
        eststo: reg `outcome' morekids `controls' if `sample'==1, robust
    }

    esttab using "OLS`sample'", title("OLS Estimates of Effects of Children on
    Parents' Labor Supply (`sample' sample)") ///
        se label wrap noabbrev rtf compress one replace // Output th table, with
        names and titles using the loop-value of sample
    eststo clear

    *Perform the IV regression for each of the dep. variables.
    foreach outcome in `dependvars' {
        display
        "*****"
        display "                Regression: IV - " `"`: var label
        `outcome'"" " - `sample' Sample "
        display
        "*****"
        eststo: ivreg2 `outcome' `controls' (morekids = samesex) if ///
        `sample'==1, first robust
    }

    esttab using "IV`sample'", title("IV Estimates of Effects of Children on
    Parents' Labor Supply (`sample' Sample)") ///
        se label wrap noabbrev rtf compress one replace // Output IV tables
    eststo clear
}

*****

***** Discussion

/* Question 3(a):
The number of children is certain to suffer from endogeneity to
omitted variables and very likely suffers from simultaneity with
labor supply. For instance, having children may cause someone to
reduce their labor supply, but at the same time, having a good job
or high income might allow someone to have children, or perhaps
more likely, experiencing unemployment / low income may cause someone
to delay plans for children.
*/

/* Question 3(b):
Would-be parents are often perceived to desire children of a certain
sex, or often, to have both a girl and a boy. Hence, if the first
two children are the same sex, in this case parents might be expected

```

to have a weakly greater propensity to have another child. This implies that having two children of the same sex might be a relevant predictor of having a third child.

At the same time, there aren't immediately obvious reasons why having two children of the same sex would be expected to influence the labor supply of parents via other channels, hence it seems quite possible that this also works as a valid instrument.

*/

/* Question 3(d):

As one can see from the first state regressions, the F-test of the excluded instruments indicates that the instrument is relevant since we reject the null of irrelevance in the underidentification test. Furthermore, this F-value is large enough that we know we do not have a problem with weak instruments, which is confirmed by the Stock-Yogo critical values suggesting that the instrument has very good size properties.

*/

/* Question 3(e):

Using the "same sex" as an instrument, we estimate that having three or more children is estimated to reduce the likelihood of a mother working by 11.7 percentage points. It is estimated to reduce mother's weeks worked by 5.559 weeks, and hours worked per week by 4.547 hours. Having three or more children is estimated to reduce mother's income by \$2,701 in 2017 USD, and it is estimated to reduce family income by 2.46 percentage points. All of these estimates are highly statistically significant.

*/

/* Question 3(f):

We find that the reductions in labor supply or income estimated by IV regression is smaller in every case than the effects estimated using ordinary least squares, although they each still remain highly significant.

*/

/* Question 3(g):

The overidentification restrictions (OIR) test is not meaningful because it used to check coherence of instruments in the case where there are more than excluded instruments than endogenous variables. Here, we have the same number of excluded instruments as endogenous variables (just identified). Hence, we cannot conduct the OIR test.

*/

***** For 3 person groups

**** Regressions (4a)

```
local dependvars workedind_moth wks_wrked_moth hourswked_moth ///
    totalinc_moth logincfam income_nonmoth
```

```
local controls age_mother agelstbth_moth boylst black_mother ///
    hisp_moth othrace_moth
```

```
foreach sample in Main Married { // Perform the same regressions for each
the Main and Married sample
```

```
    eststo clear
```

```
        foreach outcome in `dependvars' { // Perform the IV regression for each
outcome in the list of dependent variables
```

```
            display
```

```
            "*****"
            "*****"
```

```

display " Regression: IV - " "`": var label `outcome'' "' - `sample'
Sample - Using Two Sons and Two Daughters Instruments "
display
"*****"
*****"
eststo: ivreg2 `outcome' `controls' (morekids = boys2 girls2) if ///
`sample'==1, first robust
}
}

```

/* Question 4(b):

Having more excluded instruments than endogenous regressors (overidentification) can be advantageous in a couple of ways. First, if the instruments satisfy the necessary assumptions (validity and relevance) and they're not weak, then the overidentified 2SLS has greater asymptotic efficiency than IV regression.

Additionally, 2SLS allows us to perform a means of falsification of our IV strategy. With more than two instruments, we can perform an Overidentifying Restrictions Test, to check if we have evidence against the null hypothesis that are instruments are all valid.

*/

/* Question 4(c):

A potential concern is that the instruments may be weak, in which additional instruments increases the bias of 2SLS.

*/

/* Question 4(d):

Using the F-test for excluded instruments, we once again find the instrument is relevant (we reject the null of underidentification) and we find that weak instruments is not a problem.

*/

/* Question 4(e):

Looking at the results for the OIR test, we reject the null hypothesis that all the instruments are valid for the labor supply variables (mother work, weeks worked, and hours worked), but not for the income variables. This may cast doubt on the validity of these instruments.

*/

/* Fully written out method (looong) for running regressions.

**** Main Sample

*** OLS Estimates

eststo clear

```

eststo: reg workedind_moth morekids age_mother agelstbth_moth ///
      boy1st boy2nd black_mother hisp_moth othrace_moth //
      if Main==1, robust

```

```

eststo: reg wks_wrked_moth morekids age_mother agelstbth_moth ///
      boy1st boy2nd black_mother hisp_moth othrace_moth ///
      if Main==1, robust

```

```

eststo: reg hourswked_moth morekids age_mother agelstbth_moth ///
      boy1st boy2nd black_mother hisp_moth othrace_moth ///
      if Main==1, robust

```

```

eststo: reg totalinc_moth morekids age_mother agelstbth_moth ///
      boy1st boy2nd black_mother hisp_moth othrace_moth ///
      if Main==1, robust
eststo: reg logincfam morekids age_mother agelstbth_moth ///
      boy1st boy2nd black_mother hisp_moth othrace_moth ///
      if Main==1, robust
esttab using "OLSMain", title("OLS Estimates of Effects of Children on
Parents' Labor Supply (Main sample)") ///
      label wrap noabbrev rtf compress one replace

*** IV Regression Esimates
eststo clear
eststo: ivreg2 workedind_moth age_mother agelstbth_moth boy1st ///
      boy2nd black_mother hisp_moth othrace_moth ///
      (morekids = samesex) if Main==1, first robust

eststo: ivreg2 wks_wrked_moth age_mother agelstbth_moth boy1st ///
      boy2nd black_mother hisp_moth othrace_moth ///
      (morekids = samesex) if Main==1, first robust
eststo: ivreg2 hourswked_moth age_mother agelstbth_moth boy1st ///
      boy2nd black_mother hisp_moth othrace_moth ///
      (morekids = samesex) if Main==1, first robust
eststo: ivreg2 totalinc_moth age_mother agelstbth_moth boy1st ///
      boy2nd black_mother hisp_moth othrace_moth ///
      (morekids = samesex) if Main==1, first robust
eststo: ivreg2 logincfam age_mother age_mother agelstbth_moth boy1st ///
      boy2nd black_mother hisp_moth othrace_moth ///
      (morekids = samesex) if Main==1, first robust
esttab using "IVMain", title("IV Estimates of Effects of Children on Parents
Labor Supply (Main sample)") ///
      label wrap noabbrev rtf compress one replace

***** Married Sample
*** OLS Estimates
eststo clear
eststo: reg workedind_moth morekids age_mother agelstbth_moth ///
      boy1st boy2nd black mother hisp_moth othrace_moth ///
      if Married==1, robust
eststo: reg wks_wrked_moth morekids age_mother agelstbth_moth ///
      boy1st boy2nd black_mother hisp_moth othrace_moth ///
      if Married==1, robust
eststo: reg hourswked_moth morekids age_mother agelstbth_moth ///
      boy1st boy2nd black mother hisp_moth othrace_moth ///
      if Married==1, robust
eststo: reg totalinc_moth morekids age_mother agelstbth_moth ///
      boy1st boy2nd black_mother hisp_moth othrace_moth ///
      if Married==1, robust
eststo: reg logincfam morekids age_mother agelstbth_moth ///
      boy1st boy2nd black_mother hisp_moth othrace_moth ///
      if Married==1, robust
esttab using "OLSMarried", ///
      title("OLS Estimates of Effects of Children on Parents Labor Supply (Married
sample)") ///
      label wrap noabbrev rtf compress one replace

*** IV Regression Esimates
eststo clear
eststo: ivreg2 workedind_moth age_mother agelstbth_moth boy1st ///
      boy2nd black_mother hisp_moth othrace_moth ///
      (morekids = samesex ) if (Married==1), robust
eststo: ivreg2 wks_wrked_moth age_mother agelstbth_moth boy1st ///
      boy2nd black_mother hisp_moth othrace_moth ///
      (morekids = samesex ) if (Married==1), robust
eststo: ivreg2 hourswked_moth age_mother agelstbth_moth boy1st ///
      boy2nd black_mother hisp_moth othrace_moth ///
      (morekids = samesex ) if (Married==1), robust
eststo: ivreg2 totalinc_moth age_mother agelstbth_moth boy1st ///

```

```
    boy2nd black_mother hisp_moth othrace_moth ///
    (morekids = samesex ) if (Married==1), robust
eststo: ivreg2 logincfam age mother agelstbth moth boylst ///
    boy2nd black_mother hisp_moth othrace_moth ///
    (morekids = samesex ) if (Married==1), robust
esttab using "IVMarried", ///
    title("IV Estimates of Effects of Children on Parents Labor Supply (Married
    sample)") ///
    label wrap noabbrev rtf compress one replace
*/
```