Statistics/Data Analysis

```
      name:  <unnamed>
       log:  C:\Users\eaj628\Documents\hwk3.smcl
  log type:  smcl
 opened on:  12 Oct 2016, 23:27:34

1 . /*
  >         Assignment 3
  >
  >         Evan Johnston
  > */
2 .
3 . set more off

4 . cd "\\tsclient\Stat Apps Server\hwk3"
  \\tsclient\Stat Apps Server\hwk3

5 .
6 . * prob 4
7 . use "\\tsclient\Stat Apps Server\Data Sets- STATA\401ksubs.dta", clear

8 .
9 . * part 4.a
10. regress pira p401k inc incsq age agesq, robust

Linear regression                               Number of obs     =      9,275
                                                F(5, 9269)        =     422.16
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1800
                                                Root MSE          =     .39449

------------------------------------------------------------------------------
             |               Robust
        pira |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       p401k |   .0536598   .0102763     5.22   0.000      .033516    .0738035
         inc |   .0086788    .000484    17.93   0.000     .0077301    .0096276
       incsq |  -.0000228   3.86e-06    -5.91   0.000    -.0000304   -.0000152
         age |  -.0015936   .0032391    -0.49   0.623    -.0079431    .0047558
       agesq |   .0001173   .0000382     3.07   0.002     .0000424    .0001922
       _cons |  -.1977236    .065172    -3.03   0.002     -.325475   -.0699721
------------------------------------------------------------------------------

11. ****************************************
12. * part 4.d
13. * first stage
14. regress p401k e401k, robust

Linear regression                               Number of obs     =      9,275
                                                F(1, 9273)        =    8666.03
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5916
                                                Root MSE          =     .28577

------------------------------------------------------------------------------
             |               Robust
       p401k |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       e401k |   .7044267    .007567    93.09   0.000     .6895937    .7192598
       _cons |  -5.05e-15          .        .       .            .           .
------------------------------------------------------------------------------
```

```
15. predict p401k_hat
  (option xb assumed; fitted values)

16. ***************************************
17. * part 4.e
18. * second stage
19. regress pira p401k_hat inc incsq age agesq, robust
```

Linear regression

Number of obs	9,275
F(5, 9269)	411.13
Prob > F	0.0000
R-squared	0.1774
Root MSE	.3951

pira	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
p401k_hat	.0202433	.0129526	1.56	0.118	-.0051467	.0456332
inc	.0090213	.0004882	18.48	0.000	.0080642	.0099783
incsq	-.0000241	3.89e-06	-6.20	0.000	-.0000317	-.0000165
age	-.0012444	.0032578	-0.38	0.702	-.0076303	.0051416
agesq	.0001131	.0000384	2.94	0.003	.0000378	.0001885
_cons	-.2060891	.0655143	-3.15	0.002	-.3345117	-.0776666

```
20. * iv command version
21. ivregress 2sls pira (p401k=e401k) inc incsq age agesq, robust
```

Instrumental variables (2SLS) regression

Number of obs	9,275
Wald chi2(5)	2063.76
Prob > chi2	0.0000
R-squared	0.1789
Root MSE	.39461

pira	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
p401k	.0207012	.0132278	1.56	0.118	-.0052249	.0466273
inc	.0089982	.0004911	18.32	0.000	.0080356	.0099608
incsq	-.0000241	3.88e-06	-6.22	0.000	-.0000317	-.0000165
age	-.0011466	.003248	-0.35	0.724	-.0075126	.0052193
agesq	.0001121	.0000383	2.92	0.003	.000037	.0001872
_cons	-.2073136	.0653589	-3.17	0.002	-.3354147	-.0792125

Instrumented: p401k
Instruments: inc incsq age agesq
e401k

```
22.
23. ********************************************************************************
24. * prob 5
25. use "\\tsclient\Stat Apps Server\Data Sets- STATA\voucher.dta", clear

26.
27. * part a
28. tab choiceyrs selectyrs
```

years attended choice school	years selected to attend choice school 0	1	2	3	4	Total
0	467	8	2	5	8	490
1	1	107	75	20	8	211
2	0	1	71	41	9	122
3	0	0	0	84	27	111
4	0	0	0	0	56	56
Total	468	116	148	150	108	990

```
29. ****************************************
30. * part b
31. regress choiceyrs selectyrs, robust
```

Linear regression

Number of obs	990
F(1, 988)	1665.27
Prob > F	0.0000
R-squared	0.7898
Root MSE	.576

choiceyrs	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
selectyrs	.7668317	.0187913	40.81	0.000	.7299562	.8037073
_cons	.0199189	.0105037	1.90	0.058	-.0006931	.040531

```
32. ****************************************
33. * part c
34. regress mnce choiceyrs, robust
```

Linear regression

Number of obs	990
F(1, 988)	13.58
Prob > F	0.0002
R-squared	0.0122
Root MSE	20.754

mnce	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
choiceyrs	-1.837014	.4985704	-3.68	0.000	-2.815393	-.858636
_cons	46.2344	.8973782	51.52	0.000	44.47342	47.99539

```
35. regress mnce choiceyrs black hispanic female, robust
```

Linear regression

Number of obs	990
F(4, 985)	20.28
Prob > F	0.0000
R-squared	0.0868
Root MSE	19.986

mnce	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
choiceyrs	-.5652475	.4940605	-1.14	0.253	-1.53478	.4042845
black	-16.01743	1.926572	-8.31	0.000	-19.79808	-12.23677
hispanic	-13.40287	2.41094	-5.56	0.000	-18.13404	-8.671704
female	1.352745	1.279764	1.06	0.291	-1.158633	3.864123
_cons	57.12192	1.879984	30.38	0.000	53.43268	60.81115

```
36. ****************************************
37. * part e
38. ivregress 2sls mnce (choiceyrs=selectyrs) black hispanic female, robust
```

Instrumental variables (2SLS) regression

Number of obs	990
Wald chi2(**4**)	80.28
Prob > chi2	0.0000
R-squared	0.0864
Root MSE	19.939

mnce	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
choiceyrs	-.2413189	.5917809	-0.41	0.683	-1.401188	.9185504
black	-16.31692	1.947611	-8.38	0.000	-20.13417	-12.49967
hispanic	-13.7754	2.41278	-5.71	0.000	-18.50436	-9.04644
female	1.319709	1.277918	1.03	0.302	-1.184964	3.824383
_cons	57.06804	1.87722	30.40	0.000	53.38875	60.74732

Instrumented: choiceyrs
Instruments: black hispanic
female selectyrs

```
39. ****************************************
40. * part f
41. regress mnce choiceyrs black hispanic female mnce90, robust
```

Linear regression

Number of obs	=	328
F(5, 322)	=	49.77
Prob > F	=	0.0000
R-squared	=	0.4237
Root MSE	=	16.029

mnce	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
choiceyrs	.4105823	.7544828	0.54	0.587	-1.073756	1.894921
black	-8.305183	2.552212	-3.25	0.001	-13.3263	-3.284067
hispanic	-4.10498	3.465353	-1.18	0.237	-10.92257	2.712612
female	-.882847	1.784385	-0.49	0.621	-4.393373	2.627679
mnce90	.6203655	.0473119	13.11	0.000	.5272861	.7134449
_cons	22.1529	3.599737	6.15	0.000	15.07092	29.23487

```
42. ivregress 2sls mnce (choiceyrs=selectyrs) black hispanic female mnce90, robust
```

Instrumental variables (2SLS) regression

Number of obs	=	328
Wald chi2(5)	=	258.25
Prob > chi2	=	0.0000
R-squared	=	0.4173
Root MSE	=	15.969

mnce	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
choiceyrs	1.799385	.9378768	1.92	0.055	-.0388202	3.637589
black	-9.067109	2.556081	-3.55	0.000	-14.07694	-4.057283
hispanic	-5.00373	3.43925	-1.45	0.146	-11.74453	1.737076
female	-1.020484	1.773235	-0.58	0.565	-4.495961	2.454992
mnce90	.6288128	.0468642	13.42	0.000	.5369606	.7206649
_cons	21.53886	3.585963	6.01	0.000	14.5105	28.56722

Instrumented: choiceyrs
Instruments: black hispanic
female mnce90
selectyrs

```
. ****************************************
. * part h
. ivregress 2sls mnce (choiceyrs1 choiceyrs2 choiceyrs3 choiceyrs4 = ///
>         selectyrs1 selectyrs2 selectyrs3 selectyrs4) black hispanic female, robust

Instrumental variables (2SLS) regression          Number of obs   =        990
                                                  Wald chi2(7)    =      83.85
                                                  Prob > chi2     =     0.0000
                                                  R-squared       =     0.0850
                                                  Root MSE        =     19.955

------------------------------------------------------------------------------
             |               Robust
        mnce |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  choiceyrs1 |    .3899757   2.461086     0.16   0.874    -4.433664    5.213616
  choiceyrs2 |    .7736516   3.919113     0.20   0.844    -6.907669    8.454972
  choiceyrs3 |   -4.284797   3.559596    -1.20   0.229    -11.26148    2.691882
  choiceyrs4 |    2.407061   4.071456     0.59   0.554    -5.572846    10.38697
       black |   -16.29717   2.016116    -8.08   0.000    -20.24869   -12.34566
    hispanic |   -13.36599   2.568347    -5.20   0.000    -18.39985   -8.332119
      female |     1.36639   1.279023     1.07   0.285    -1.140448    3.873228
       _cons |    56.88582   1.895133    30.02   0.000     53.17143    60.60021
------------------------------------------------------------------------------
Instrumented:  choiceyrs1
               choiceyrs2
               choiceyrs3
               choiceyrs4
Instruments:   black hispanic
               female selectyrs1
               selectyrs2
               selectyrs3
               selectyrs4
. ****************************************
. * part i
. ivregress 2sls mnce (choiceyrs = selectyrs1 selectyrs2 selectyrs3 selectyrs4) ///
>         black hispanic female, robust

Instrumental variables (2SLS) regression          Number of obs   =        990
                                                  Wald chi2(4)    =      80.28
                                                  Prob > chi2     =     0.0000
                                                  R-squared       =     0.0865
                                                  Root MSE        =     19.939

------------------------------------------------------------------------------
             |               Robust
        mnce |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   choiceyrs |   -.252745   .5906766    -0.43   0.669     -1.41045    .9049599
       black |   -16.30635   1.946441    -8.38   0.000    -20.12131   -12.4914
    hispanic |   -13.76226     2.4129    -5.70   0.000    -18.49146   -9.033064
      female |    1.320875   1.278125     1.03   0.301    -1.184203    3.825953
       _cons |    57.06994   1.876938    30.41   0.000     53.39121    60.74867
------------------------------------------------------------------------------
Instrumented:  choiceyrs
Instruments:   black hispanic
               female selectyrs1
               selectyrs2
               selectyrs3
               selectyrs4
```

```
49.
50. * regression from d/e
51. ivregress 2sls mnce (choiceyrs=selectyrs) black hispanic female, robust

Instrumental variables (2SLS) regression          Number of obs   =        990
                                                  Wald chi2(4)    =      80.28
                                                  Prob > chi2     =     0.0000
                                                  R-squared       =     0.0864
                                                  Root MSE        =     19.939

------------------------------------------------------------------------------
             |               Robust
        mnce |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   choiceyrs |   -.2413189   .5917809    -0.41   0.683    -1.401188    .9185504
       black |   -16.31692   1.947611    -8.38   0.000    -20.13417   -12.49967
    hispanic |   -13.7754    2.41278     -5.71   0.000    -18.50436    -9.04644
      female |    1.319709   1.277918     1.03   0.302    -1.184964    3.824383
       _cons |    57.06804    1.87722    30.40   0.000     53.38875    60.74732
------------------------------------------------------------------------------
Instrumented:  choiceyrs
Instruments:   black hispanic
               female selectyrs

52. ****************************************
53. ivregress gmm mnce (choiceyrs = selectyrs1 selectyrs2 selectyrs3 selectyrs4) ///
>         black hispanic female, robust

Instrumental variables (GMM) regression           Number of obs   =        990
                                                  Wald chi2(4)    =      82.53
                                                  Prob > chi2     =     0.0000
                                                  R-squared       =     0.0864
GMM weight matrix: Robust                         Root MSE        =     19.939

------------------------------------------------------------------------------
             |               Robust
        mnce |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   choiceyrs |   -.2437544   .5907315    -0.41   0.680    -1.401567    .9140581
       black |   -16.46039   1.938796    -8.49   0.000    -20.26036   -12.66042
    hispanic |   -13.83923   2.409442    -5.74   0.000    -18.56164   -9.116806
      female |     1.34696   1.275053     1.06   0.291    -1.152097    3.846017
       _cons |    57.17299   1.873959    30.51   0.000     53.50009    60.84588
------------------------------------------------------------------------------
Instrumented:  choiceyrs
Instruments:   black hispanic
               female selectyrs1
               selectyrs2
               selectyrs3
               selectyrs4

54. estat overid

    Test of overidentifying restriction:

    Hansen's J chi2(3) = 1.71092 (p = 0.6345)

55. ****************************************
56. * part j
57. regress mnce selectyrs black hispanic female, robust

Linear regression                                 Number of obs     =        990
                                                  F(4, 985)         =      19.97
                                                  Prob > F          =     0.0000
                                                  R-squared         =     0.0859
                                                  Root MSE          =     19.996
```

mnce	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
selectyrs	-.1842306	.4532439	-0.41	0.684	-1.073665	.705204
black	-16.32402	1.947134	-8.38	0.000	-20.14503	-12.50301
hispanic	-13.78579	2.41386	-5.71	0.000	-18.52269	-9.048886
female	1.311901	1.27987	1.03	0.306	-1.199684	3.823485
_cons	57.07264	1.882335	30.32	0.000	53.37879	60.76648

```
58. * regression from c
59. regress mnce choiceyrs black hispanic female, robust
```

Linear regression

Number of obs	=	990
F(4, 985)	=	20.28
Prob > F	=	0.0000
R-squared	=	0.0868
Root MSE	=	19.986

mnce	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
choiceyrs	-.5652475	.4940605	-1.14	0.253	-1.53478	.4042845
black	-16.01743	1.926572	-8.31	0.000	-19.79808	-12.23677
hispanic	-13.40287	2.41094	-5.56	0.000	-18.13404	-8.671704
female	1.352745	1.279764	1.06	0.291	-1.158633	3.864123
_cons	57.12192	1.879984	30.38	0.000	53.43268	60.81115

```
60.
61. log close
      name:  <unnamed>
       log:  C:\Users\eaj628\Documents\hwk3.smcl
  log type:  smcl
 closed on:  12 Oct 2016, 23:27:36
```