Statistics/Data Analysis

```
      name:  <unnamed>
       log:  C:\Users\eaj628\Documents\hwk1.smcl
  log type:  smcl
 opened on:  11 Sep 2016, 18:15:57

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

4 . cd "\\tsclient\Stat Apps Server\hwk1"
  \\tsclient\Stat Apps Server\hwk1

5 .
6 . * prob 3
7 . use "\\tsclient\Stat Apps Server\Data Sets- STATA\fertil2.dta", clear

8 .
9 . gen heducmissing = heduc==.

10. replace heduc=0 if heducmissing
  (2,405 real changes made)

11.
12. * part 3.a
13. regress children age agesq educ evermarr heduc heducmissing
```

Source	SS	df	MS		
Model	12723.9061	6	2120.65102	Number of obs	= 4,361
Residual	8803.27023	4,354	2.02188108	F(6, 4354)	= 1048.85
				Prob > F	= 0.0000
				R-squared	= 0.5911
				Adj R-squared	= 0.5905
Total	21527.1763	4,360	4.93742577	Root MSE	= 1.4219

children	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.2786184	.0170281	16.36	0.000	.2452347	.3120021
agesq	-.0020174	.0002739	-7.36	0.000	-.0025544	-.0014803
educ	-.0612256	.0065688	-9.32	0.000	-.0741039	-.0483474
evermarr	.1754812	.1356173	1.29	0.196	-.0903978	.4413601
heduc	-.0646732	.007649	-8.46	0.000	-.0796692	-.0496772
heducmissing	-.8460843	.1369495	-6.18	0.000	-1.114575	-.5775937
_cons	-2.809784	.2781238	-10.10	0.000	-3.355048	-2.26452

```
14.
15. * part 3.b
16. regress children age agesq educ heduc heducmissing if evermarr==1
```

Source	SS	df	MS		
Model	4563.36644	5	912.673289	Number of obs	= 2,079
Residual	6293.42144	2,073	3.03590036	F(5, 2073)	= 300.63
				Prob > F	= 0.0000
				R-squared	= 0.4203
				Adj R-squared	= 0.4189
Total	10856.7879	2,078	5.22463324	Root MSE	= 1.7424

children	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.5254392	.0390565	13.45	0.000	.4488451	.6020333
agesq	-.0054568	.000581	-9.39	0.000	-.0065961	-.0043175
educ	-.0759259	.0118672	-6.40	0.000	-.0991988	-.052653
heduc	-.0551317	.0106142	-5.19	0.000	-.0759472	-.0343162
heducmissing	-.8397852	.1683276	-4.99	0.000	-1.169894	-.5096763
_cons	-6.777208	.6338617	-10.69	0.000	-8.02028	-5.534136

```
17. regress children age agesq educ heduc heducmissing if evermarr==0
  note: heduc omitted because of collinearity
  note: heducmissing omitted because of collinearity
```

Source	SS	df	MS		
				Number of obs	= 2,282
				F(3, 2278)	= 972.55
Model	2956.05077	3	985.350258	Prob > F	= 0.0000
Residual	2307.97026	2,278	1.01315639	R-squared	= 0.5616
				Adj R-squared	= 0.5610
Total	5264.02103	2,281	2.30776898	Root MSE	= 1.0066

children	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.2290825	.0176561	12.97	0.000	.1944587	.2637062
agesq	-.0014574	.0003161	-4.61	0.000	-.0020773	-.0008375
educ	-.0549112	.0064503	-8.51	0.000	-.0675604	-.042262
heduc	0	(omitted)				
heducmissing	0	(omitted)				
_cons	-2.87956	.2326078	-12.38	0.000	-3.335706	-2.423415

```
18.
19. * part 3.c
20. regress children age agesq educ electric
```

Source	SS	df	MS		
				Number of obs	= 4,358
				F(4, 4353)	= 1451.87
Model	12294.619	4	3073.65474	Prob > F	= 0.0000
Residual	9215.41316	4,353	2.11702577	R-squared	= 0.5716
				Adj R-squared	= 0.5712
Total	21510.0321	4,357	4.93689055	Root MSE	= 1.455

children	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.3370397	.0165166	20.41	0.000	.3046588	.3694206
agesq	-.0026696	.0002718	-9.82	0.000	-.0032025	-.0021367
educ	-.0788944	.0062513	-12.62	0.000	-.0911502	-.0666387
electric	-.3777901	.0673176	-5.61	0.000	-.5097669	-.2458133
_cons	-4.247568	.2406003	-17.65	0.000	-4.719267	-3.775869

```
21.
22. * part 3.d
23. gen interact1 = electric*age*educ
  (3 missing values generated)

24. regress children age agesq educ electric interact1
```

Source	SS	df	MS		
				Number of obs	= 4,358
				F(5, 4352)	= 1169.11
Model	12330.1911	5	2466.03822	Prob > F	= 0.0000
Residual	9179.84104	4,352	2.10933847	R-squared	= 0.5732
				Adj R-squared	= 0.5727
Total	21510.0321	4,357	4.93689055	Root MSE	= 1.4524

children	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.3420078	.0165309	20.69	0.000	.3095988	.3744168
agesq	-.0026829	.0002713	-9.89	0.000	-.0032148	-.002151
educ	-.0662242	.006961	-9.51	0.000	-.0798714	-.052577
electric	.0243015	.1187531	0.20	0.838	-.208515	.257118
interact1	-.0017923	.0004364	-4.11	0.000	-.0026479	-.0009366
_cons	-4.440265	.2447042	-18.15	0.000	-4.92001	-3.96052

```
25.
26. gen elec_partial = _b[electric]+_b[interact1]*age*educ

27. histogram elec_partial
  (bin=36, start=-1.4686629, width=.04147123)

28. graph export elec_partial_hist.png, replace
  (file elec_partial_hist.png written in PNG format)

29. sum elec_partial
```

Variable	Obs	Mean	Std. Dev.	Min	Max
elec_partial	**4,361**	**-.2444093**	**.2073983**	**-1.468663**	**.0243015**

```
30.
31. * prob 4
32. use "\\tsclient\Stat Apps Server\Data Sets- STATA\card.dta", clear

33. drop if _n>3000
  (10 observations deleted)

34.
35. * part 4.b
36. regress lwage educ exper expersq, robust
```

Linear regression

Number of obs	**3,000**
F(3, 2996)	**231.11**
Prob > F	**0.0000**
R-squared	**0.1957**
Root MSE	**.3987**

lwage	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
educ	**.0932231**	**.0036867**	**25.29**	**0.000**	**.0859944**	**.1004518**
exper	**.0897303**	**.0070862**	**12.66**	**0.000**	**.0758361**	**.1036246**
expersq	**-.0024818**	**.000341**	**-7.28**	**0.000**	**-.0031504**	**-.0018132**
_cons	**4.467962**	**.0704221**	**63.45**	**0.000**	**4.329881**	**4.606042**

```
37.
38. lincom 0.0897303+2*(-0.0024818)*exper

  ( 1)  - .0049636*exper = -.0897303
```

lwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(1)	**.0892849**	**.0000352**	**2538.46**	**0.000**	**.0892159**	**.0893539**

```
39. display 0.0897303+2*(-0.0024818)*10-1.645*0.0000352
  .0400364

40. display 0.0897303+2*(-0.0024818)*10+1.645*0.0000352
  .0401522

41.
```

```
42. * part 4.c
43. regress lwage educ exper expersq if (_n<1001), robust

Linear regression                               Number of obs     =      1,000
                                                F(3, 996)         =      73.27
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1928
                                                Root MSE          =     .38612

------------------------------------------------------------------------------
             |               Robust
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0841381   .0062012    13.57   0.000     .0719692    .096307
       exper |   .0950498    .012834     7.41   0.000      .069865   .1202347
     expersq |  -.002494    .0006395    -3.90   0.000    -.0037489  -.0012391
       _cons |   4.628301   .1212139    38.18   0.000     4.390437   4.866165
------------------------------------------------------------------------------

44. regress lwage educ exper expersq if (1000<_n & _n<2001), robust

Linear regression                               Number of obs     =      1,000
                                                F(3, 996)         =      57.79
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1449
                                                Root MSE          =     .39067

------------------------------------------------------------------------------
             |               Robust
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0813257   .0065912    12.34   0.000     .0683915   .0942599
       exper |   .0777722   .0114661     6.78   0.000     .0552718   .1002726
     expersq |  -.0020686   .0005382    -3.84   0.000    -.0031248  -.0010124
       _cons |   4.677698   .1210746    38.63   0.000     4.440107   4.915288
------------------------------------------------------------------------------

45. regress lwage educ exper expersq if (_n>2000), robust

Linear regression                               Number of obs     =      1,000
                                                F(3, 996)         =      89.21
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2283
                                                Root MSE          =      .4036

------------------------------------------------------------------------------
             |               Robust
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0997974   .0063513    15.71   0.000     .0873339   .1122609
       exper |   .0816429   .0121005     6.75   0.000     .0578975   .1053884
     expersq |  -.0022062    .000564    -3.91   0.000    -.0033131  -.0010994
       _cons |   4.363985   .1232061    35.42   0.000     4.122212   4.605758
------------------------------------------------------------------------------

46.
47. * part 4.c.iii
48. display 1/3*(0.0062012+0.0065912+0.0063513)
  .00638123
```

```
49. display ( (0.0062012-.00638123)^2+(0.0065912-.00638123)^2+(0.0063513-.00638123)^2 )^
  > (1/2)
  .0002782

50.
51. * part 4.d.ii
52. bootstrap sig=e(rmse) rsq=e(r2) _b, reps(1000): regress lwage educ exper expersq
  (running regress on estimation sample)

  Bootstrap replications (1000)
  ----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
  .................................................    50
  .................................................   100
  .................................................   150
  .................................................   200
  .................................................   250
  .................................................   300
  .................................................   350
  .................................................   400
  .................................................   450
  .................................................   500
  .................................................   550
  .................................................   600
  .................................................   650
  .................................................   700
  .................................................   750
  .................................................   800
  .................................................   850
  .................................................   900
  .................................................   950
  ................................................... 1000

  Linear regression                               Number of obs     =      3,000
                                                  Replications      =      1,000

        command:  regress lwage educ exper expersq
      [_eq2]sig:  e(rmse)
      [_eq2]rsq:  e(r2)

------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_eq1         |
       educ  |   .0932231    .003819    24.41   0.000     .0857379    .1007083
      exper  |   .0897303   .0071192    12.60   0.000      .075777    .1036837
    expersq  |  -.0024818   .0003407    -7.28   0.000    -.0031497    -.001814
       _cons |   4.467962   .0732877    60.96   0.000      4.32432    4.611603
-------------+----------------------------------------------------------------
_eq2         |
         sig |    .398702   .0055474    71.87   0.000     .3878293    .4095748
         rsq |   .1957357   .0137448    14.24   0.000     .1687965     .222675
------------------------------------------------------------------------------

53. matrix list e(ci_normal)

  e(ci_normal)[2,6]
              _eq1:       _eq1:       _eq1:       _eq1:       _eq2:       _eq2:
              educ       exper     expersq       _cons         sig         rsq
  ll    .08573793   .07577697  -.00314966   4.3243202   .38782927   .16879645
  ul    .10070832   .10368373  -.00181397   4.6116028   .40957477   .22267497
```

```
54. matrix list e(ci_percentile)

  e(ci_percentile)[2,6]
              _eq1:       _eq1:       _eq1:       _eq1:       _eq2:       _eq2:
              educ       exper     expersq       _cons         sig         rsq
  ll    .08563615   .07575226  -.00316833   4.3266301   .38762943   .16943438
  ul    .10057391   .10373978  -.00181378   4.6073582   .40899366   .22215603

55.
56. log close
        name:  <unnamed>
         log:  C:\Users\eaj628\Documents\hwk1.smcl
    log type:  smcl
   closed on:  11 Sep 2016, 18:16:35
```