Statistics/Data Analysis

```
> ----------------------------------------------------------------------------
> --------
      name:  <unnamed>
       log:  C:\Users\ehs588\Downloads\hw5_output.txt
  log type:  text
 opened on:  25 Apr 2018, 15:50:36

1 . /*fuck STATA -- no time to code these things in python tho*/
2 .
3 . /*---------------Question 2---------------*/
4 . use "C:\Users\ehs588\Downloads\wagepan.dta", clear

5 . regress lwage exper expersq educ union black poorhlth

      Source |       SS           df       MS      Number of obs   =     4,360
-------------+----------------------------------   F(6, 4353)      =    157.02
       Model |  220.004631         6  36.6674385   Prob > F        =    0.0000
    Residual |  1016.52501     4,353   .23352286   R-squared       =    0.1779
-------------+----------------------------------   Adj R-squared   =    0.1768
       Total |  1236.52964     4,359  .283672779   Root MSE        =    .48324

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       exper |   .1006233   .0100142    10.05   0.000     .0809904    .1202563
     expersq |   -.003226   .0007083    -4.55   0.000    -.0046147   -.0018373
        educ |   .1018746   .0046038    22.13   0.000     .0928489    .1109004
       union |   .1863272   .0171747    10.85   0.000     .1526561    .2199984
       black |  -.1719465   .0230514    -7.46   0.000     -.217139    -.126754
    poorhlth |  -.0600268   .0567319    -1.06   0.290    -.1712501    .0511966
       _cons |  -.0670423   .0630324    -1.06   0.288     -.190618    .0565334
------------------------------------------------------------------------------

6 . xtreg lwage exper expersq educ union black poorhlth, fe
note: educ omitted because of collinearity
note: black omitted because of collinearity

Fixed-effects (within) regression               Number of obs     =      4,360
Group variable: nr                              Number of groups  =        545

R-sq:                                           Obs per group:
     within  = 0.1768                                         min =          8
     between = 0.0002                                         avg =        8.0
     overall = 0.0563                                         max =          8

                                                F(4,3811)         =     204.57
corr(u_i, Xb)  = -0.1313                        Prob > F          =     0.0000

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       exper |   .1218831   .0081721    14.91   0.000     .1058611    .1379051
     expersq |  -.0044803   .0006013    -7.45   0.000    -.0056593   -.0033014
        educ |          0  (omitted)
       union |   .0831462   .0193041     4.31   0.000     .0452989    .1209935
       black |          0  (omitted)
    poorhlth |  -.0203299   .0472171    -0.43   0.667    -.1129032    .0722434
       _cons |   1.061092   .0266618    39.80   0.000     1.008819    1.113364
-------------+----------------------------------------------------------------
     sigma_u |  .40337877
     sigma_e |  .35152881
         rho |  .56836135   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(544, 3811) = 10.04                     Prob > F = 0.0000
```

```
7 . estimates store fixed

8 . xtreg lwage exper expersq educ union black poorhlth, re

Random-effects GLS regression                   Number of obs     =      4,360
Group variable: nr                              Number of groups  =        545

R-sq:                                           Obs per group:
     within  = 0.1763                                         min =          8
     between = 0.1709                                         avg =        8.0
     overall = 0.1734                                         max =          8

                                                Wald chi2(6)      =     927.41
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       exper |   .1190793   .0080506    14.79   0.000     .1033005    .1348581
     expersq |  -.0043136   .0005885    -7.33   0.000     -.005467   -.0031602
        educ |   .1021501   .0087963    11.61   0.000     .0849096    .1193905
       union |   .1091375     .01785     6.11   0.000     .0741521    .1441228
       black |  -.1622355   .0472358    -3.43   0.001     -.254816   -.0696551
    poorhlth |   -.027118    .046428    -0.58   0.559    -.1181151    .0638792
       _cons |  -.1185205   .1076057    -1.10   0.271    -.3294238    .0923829
-------------+----------------------------------------------------------------
     sigma_u |  .32828128
     sigma_e |  .35152881
         rho |  .46584285   (fraction of variance due to u_i)
------------------------------------------------------------------------------

9 . hausman fixed ., sigmamore

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed          .          Difference          S.E.
-------------+----------------------------------------------------------------
       exper |    .1218831     .1190793        .0028038        .0015039
     expersq |   -.0044803    -.0043136       -.0001667        .0001297
       union |    .0831462     .1091375       -.0259913        .0074597
    poorhlth |   -.0203299     -.027118        .0067881        .0091438
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(4) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       22.85
                Prob>chi2 =      0.0001

10.
11. /*Do it again with dummy variables*/
12. regress lwage exper expersq educ union black poorhlth d81 d82 d83 d84 d85 d86 d87

      Source |       SS           df       MS      Number of obs   =     4,360
-------------+----------------------------------   F(13, 4346)     =     73.63
       Model |  223.191033        13   17.168541   Prob > F        =    0.0000
    Residual |  1013.33861     4,346   .23316581   R-squared       =    0.1805
-------------+----------------------------------   Adj R-squared   =    0.1780
       Total |  1236.52964     4,359  .283672779   Root MSE        =    .48287
```

```
------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       exper |   .0783167    .013666     5.73   0.000     .0515244    .1051089
     expersq |  -.0027401   .0008223    -3.33   0.001    -.0043521    -.001128
        educ |   .0940767   .0051758    18.18   0.000     .0839296    .1042238
       union |   .1888297   .0172128    10.97   0.000     .1550839    .2225756
       black |  -.1675775   .0230745    -7.26   0.000    -.2128154   -.1223396
    poorhlth |   -.056712   .0567191    -1.00   0.317    -.1679103    .0544863
         d81 |   .0605769   .0305124     1.99   0.047     .0007571    .1203967
         d82 |   .0643113   .0333899     1.93   0.054    -.0011499    .1297725
         d83 |    .065683    .036854     1.78   0.075    -.0065697    .1379356
         d84 |   .0935802   .0402989     2.32   0.020     .0145737    .1725867
         d85 |   .1100291   .0435919     2.52   0.012     .0245667    .1954914
         d86 |    .141212   .0466774     3.03   0.002     .0497005    .2327234
         d87 |   .1720964   .0497023     3.46   0.001     .0746544    .2695383
       _cons |    .055925   .0771012     0.73   0.468    -.0952326    .2070827
------------------------------------------------------------------------------
```

13. xtreg lwage exper expersq educ union black poorhlth d81 d82 d83 d84 d85 d86 d87, fe

```
note: educ omitted because of collinearity
note: black omitted because of collinearity
note: d87 omitted because of collinearity

Fixed-effects (within) regression               Number of obs      =      4,360
Group variable: nr                              Number of groups   =        545

R-sq:                                           Obs per group:
     within  = 0.1792                                         min =          8
     between = 0.0002                                         avg =        8.0
     overall = 0.0558                                         max =          8

                                                F(10,3805)         =      83.08
corr(u_i, Xb)  = -0.1392                        Prob > F           =     0.0000

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       exper |   .1369736    .009645    14.20   0.000     .1180638    .1558835
     expersq |  -.0053384   .0007024    -7.60   0.000    -.0067155   -.0039613
        educ |          0  (omitted)
       union |   .0811079   .0193248     4.20   0.000       .04322    .1189959
       black |          0  (omitted)
    poorhlth |  -.0202912   .0472043    -0.43   0.667    -.1128394    .0722571
         d81 |   .0200537   .0203757     0.98   0.325    -.0198945     .060002
         d82 |  -.0105501   .0202421    -0.52   0.602    -.0502365    .0291362
         d83 |  -.0402297   .0203252    -1.98   0.048    -.0800791   -.0003803
         d84 |   -.036876   .0203275    -1.81   0.070    -.0767298    .0029778
         d85 |  -.0425298   .0202609    -2.10   0.036     -.082253   -.0028066
         d86 |  -.0272064   .0204032    -1.33   0.182    -.0672085    .0127958
         d87 |          0  (omitted)
       _cons |   1.023714   .0299582    34.17   0.000     .9649787     1.08245
-------------+----------------------------------------------------------------
     sigma_u |  .40441167
     sigma_e |  .35128112
         rho |  .56996125   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(544, 3805) = 9.97                     Prob > F = 0.0000
```

14. estimates store fixed

15. xtreg lwage exper expersq educ union black poorhlth d81 d82 d83 d84 d85 d86 d87, re

```
Random-effects GLS regression                   Number of obs     =      4,360
Group variable: nr                              Number of groups  =        545

R-sq:                                           Obs per group:
     within  = 0.1787                                         min =          8
     between = 0.1724                                         avg =        8.0
     overall = 0.1753                                         max =          8

                                                Wald chi2(13)     =     940.50
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       exper |   .1126381    .015347     7.34   0.000     .0825586    .1427176
     expersq |  -.0049356   .0006877    -7.18   0.000    -.0062836   -.0035877
        educ |   .0929967   .0106053     8.77   0.000     .0722107    .1137827
       union |   .1079553   .0178752     6.04   0.000     .0729207      .14299
       black |  -.1581144   .0473409    -3.34   0.001    -.2509008    -.065328
    poorhlth |   -.026784   .0464186    -0.58   0.564    -.1177627    .0641947
          d81 |   .0415954   .0247482     1.68   0.093    -.0069101    .0901009
          d82 |   .0314813   .0324952     0.97   0.333    -.0322082    .0951708
          d83 |   .0219667   .0418427     0.52   0.600    -.0600436    .1039769
          d84 |   .0443199   .0516895     0.86   0.391    -.0569897    .1456295
          d85 |   .0576037   .0617221     0.93   0.351    -.0633695    .1785768
          d86 |   .0907605   .0718313     1.26   0.206    -.0500263    .2315472
          d87 |   .1332586   .0820275     1.62   0.104    -.0275123    .2940294
        _cons |   .0096985   .1493042     0.06   0.948    -.2829323    .3023294
-------------+----------------------------------------------------------------
     sigma_u |  .32831442
     sigma_e |  .35128112
         rho |   .4662439   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

16. hausman fixed ., sigmamore

```
Note: the rank of the differenced variance matrix (5) does not equal the number of coe
> fficients being tested (10); be sure this is what you expect, or there may be proble
> ms
        computing the test.  Examine the output of your estimators for anything unexpe
> cted and possibly consider scaling your variables so that the coefficients are on a
> similar
        scale.

                 ---- Coefficients ----
              |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
              |     fixed          .          Difference          S.E.
--------------+----------------------------------------------------------------
       exper |    .1369736     .1126381        .0243355                .
     expersq |   -.0053384    -.0049356       -.0004028        .0001503
       union |    .0811079     .1079553       -.0268474        .0074552
    poorhlth |   -.0202912     -.026784        .0064929        .0091346
          d81 |    .0200537     .0415954       -.0215416                .
          d82 |   -.0105501     .0314813       -.0420314                .
          d83 |   -.0402297     .0219667       -.0621964                .
          d84 |    -.036876     .0443199       -.0811958                .
          d85 |   -.0425298     .0576037       -.1001335                .
          d86 |   -.0272064     .0907605       -.1179668                .
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(5) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       22.16
                Prob>chi2 =      0.0005
                (V_b-V_B is not positive definite)
```

```
17.
18. /*---------------Question 3---------------*/
19.
20. /*part (b)*/
21. use "C:\Users\ehs588\Downloads\MROZ.DTA", clear

22. mprobit inlf kidslt6 huswage kidsge6 educ

  Iteration 0:   log likelihood = -472.19567
  Iteration 1:   log likelihood = -471.73409
  Iteration 2:   log likelihood = -471.73396
  Iteration 3:   log likelihood = -471.73396

  Multinomial probit regression                     Number of obs     =        753
                                                    Wald chi2(4)      =      77.10
  Log likelihood = -471.73396                       Prob > chi2       =     0.0000

  ------------------------------------------------------------------------------
          inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
  -------------+----------------------------------------------------------------
  0            |
       kidslt6 |   .9084417   .1398578     6.50   0.000     .6343255    1.182558
       huswage |   .0634753   .0175097     3.63   0.000      .029157    .0977937
       kidsge6 |  -.0427764   .0508989    -0.84   0.401    -.1425365    .0569837
          educ |  -.2205896   .0328126    -6.72   0.000     -.284901   -.1562781
         _cons |   1.822819   .3915409     4.66   0.000     1.055413    2.590225
  -------------+----------------------------------------------------------------
  1            |  (base outcome)
  ------------------------------------------------------------------------------

23. tobit hours kidslt6 huswage kidsge6 educ, ll

  Tobit regression                                  Number of obs     =        753
                                                    LR chi2(4)        =      90.20
                                                    Prob > chi2       =     0.0000
  Log likelihood =   -3909.79                       Pseudo R2         =     0.0114

  ------------------------------------------------------------------------------
         hours |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
  -------------+----------------------------------------------------------------
       kidslt6 |  -826.1761   115.7003    -7.14   0.000    -1053.312   -599.0406
       huswage |  -57.01664   13.77307    -4.14   0.000    -84.05506   -29.97823
       kidsge6 |  -34.27016    39.1025    -0.88   0.381    -111.0337    42.49338
          educ |   137.6396   24.04347     5.72   0.000     90.43904    184.8402
         _cons |  -722.9432   297.4544    -2.43   0.015    -1306.887   -138.9996
  -------------+----------------------------------------------------------------
        /sigma |   1291.782   48.61892                      1196.336    1387.227
  ------------------------------------------------------------------------------
            325  left-censored observations at hours <= 0
            428     uncensored observations
              0 right-censored observations

24. truncreg hours kidslt6 huswage kidsge6 educ, ll(0)
  (note: 325 obs. truncated)

  Fitting full model:

  Iteration 0:   log likelihood = -3418.4751
  Iteration 1:   log likelihood = -3407.7415
  Iteration 2:   log likelihood = -3407.6796
  Iteration 3:   log likelihood = -3407.6796

  Truncated regression
  Limit:   lower =          0                       Number of obs     =        428
           upper =       +inf                       Wald chi2(4)      =      28.50
  Log likelihood = -3407.6796                       Prob > chi2       =     0.0000
```

```
------------------------------------------------------------------------------
       hours |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     kidslt6 |  -505.8588   161.2838    -3.14   0.002    -821.9693   -189.7482
     huswage |  -35.24644   15.96651    -2.21   0.027    -66.54022   -3.952647
     kidsge6 |  -146.6171     42.453    -3.45   0.001    -229.8235   -63.41074
        educ |  -12.56781   24.22717    -0.52   0.604    -60.05218    34.91656
       _cons |   1776.841   304.6413     5.83   0.000     1179.755    2373.927
-------------+----------------------------------------------------------------
      /sigma |   899.3909    48.3947    18.58   0.000     804.5391    994.2428
------------------------------------------------------------------------------
```

```
25.
26. /*part (c)*/
27. kdensity hours

28. graph export "C:\Users\ehs588\Downloads\q3_graph.png", replace
  (file C:\Users\ehs588\Downloads\q3_graph.png written in PNG format)

29.
30. /*---------------Question 4---------------*/
31. use "C:\Users\ehs588\Downloads\cps09mar.dta", clear

32.
33. /*part (b)*/
34. kdensity wage, normal

35. graph export "C:\Users\ehs588\Downloads\q4_graph1.png", replace
  (file C:\Users\ehs588\Downloads\q4_graph1.png written in PNG format)

36.
37. /*part (c)*/
38. kdensity lwage, normal

39. graph export "C:\Users\ehs588\Downloads\q4_graph2.png", replace
  (file C:\Users\ehs588\Downloads\q4_graph2.png written in PNG format)

40.
41. /*part d*/
42. lpoly wage exp if female==1 & white==1

43. graph export "C:\Users\ehs588\Downloads\q4_graph3.png", replace
  (file C:\Users\ehs588\Downloads\q4_graph3.png written in PNG format)

44. lpoly wage exp if female==0 & white==1

45. graph export "C:\Users\ehs588\Downloads\q4_graph4.png", replace
  (file C:\Users\ehs588\Downloads\q4_graph4.png written in PNG format)

46. lpoly lwage exp if female==1 & white==1

47. graph export "C:\Users\ehs588\Downloads\q4_graph5.png", replace
  (file C:\Users\ehs588\Downloads\q4_graph5.png written in PNG format)

48. lpoly lwage exp if female==0 & white==1

49. graph export "C:\Users\ehs588\Downloads\q4_graph6.png", replace
  (file C:\Users\ehs588\Downloads\q4_graph6.png written in PNG format)

50.
```