BOSTON COLLEGE
Department of Economics
EC 771: Econometrics
Spring 2008
Prof. Baum, Mr. Alva

Problem Set 5: Solutions

Point Distribution:
1), 2): 15 points each
3), 4): 20 points each
5), 6): 15 points each

1) a)

```
. use ~/Documents/Economics/data/grunfeld

. xtreg invest mvalue kstock, fe

Fixed-effects (within) regression               Number of obs      =       200
Group variable (i): company                     Number of groups   =        10

R-sq:  within  = 0.7668                         Obs per group: min =        20
       between = 0.8194                                        avg =      20.0
       overall = 0.8060                                        max =        20

                                                F(2,188)           =    309.01
corr(u_i, Xb)  = -0.1517                        Prob > F           =    0.0000

------------------------------------------------------------------------------
      invest |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      mvalue |   .1101238   .0118567     9.29   0.000     .0867345    .1335131
      kstock |   .3100653   .0173545    17.87   0.000     .2758308    .3442999
       _cons |  -58.74393   12.45369    -4.72   0.000    -83.31086     -34.177
-------------+----------------------------------------------------------------
     sigma_u |  85.732501
     sigma_e |  52.767964
         rho |  .72525012   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(9, 188) =    49.18              Prob > F = 0.0000

. est store xfe
```

```
. xtreg invest mvalue kstock

Random-effects GLS regression                   Number of obs      =       200
Group variable (i): company                     Number of groups   =        10

R-sq:  within  = 0.7668                         Obs per group: min =        20
       between = 0.8196                                        avg =      20.0
       overall = 0.8061                                        max =        20

Random effects u_i ~ Gaussian                   Wald chi2(2)       =    657.67
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0000

------------------------------------------------------------------------------
      invest |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      mvalue |   .1097811   .0104927    10.46   0.000     .0892159    .1303464
      kstock |    .308113   .0171805    17.93   0.000     .2744399    .3417861
       _cons |  -57.83441   28.89893    -2.00   0.045    -114.4753   -1.193537
-------------+----------------------------------------------------------------
     sigma_u |   84.20095
     sigma_e |  52.767964
         rho |  .71800838   (fraction of variance due to u_i)
------------------------------------------------------------------------------

. est store xre

. hausman xfe xre

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |      xfe          xre         Difference          S.E.
-------------+----------------------------------------------------------------
      mvalue |    .1101238     .1097811        .0003427        .0055213
      kstock |    .3100653      .308113        .0019524        .0024516
------------------------------------------------------------------------------
                           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(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        2.33
                Prob>chi2 =      0.3119
```

Rejection of the null hypothesis would indicate that the RE estimator was inconsistent. Since the null hypothesis is not rejected, the RE estimator is appropriate and preferred due to its efficiency.

b)

```
. drop _est*

.
. tsset year company
       panel variable:  year (strongly balanced)
        time variable:  company, 1 to 10

. xi i.company
i.company         _Icompany_1-10      (naturally coded; _Icompany_1 omitted)

. xtreg invest mvalue kstock _Icompany_2 _Icompany_3 _Icompany_4 _Icompany_5  ///
>     _Icompany_6 _Icompany_7 _Icompany_8 _Icompany_9 _Icompany_10, fe

Fixed-effects (within) regression               Number of obs      =       200
Group variable (i): year                        Number of groups   =        20

R-sq:  within  = 0.9482                         Obs per group: min =        10
       between = 0.9379                                        avg =      10.0
       overall = 0.9416                                        max =        10

                                                F(11,169)          =    281.32
corr(u_i, Xb)  = -0.2617                        Prob > F           =    0.0000

------------------------------------------------------------------------------
      invest |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      mvalue |   .1177158   .0137513     8.56   0.000     .0905694    .1448623
      kstock |   .3579163    .022719    15.75   0.000     .3130667    .4027659
 _Icompany_2 |   207.0542   35.17275     5.89   0.000     137.6197    276.4887
 _Icompany_3 |  -135.2308   35.70897    -3.79   0.000    -205.7239   -64.73772
 _Icompany_4 |    95.3538   50.72211     1.88   0.062    -4.776741    195.4844
 _Icompany_5 |  -5.438636   57.83052    -0.09   0.925    -119.6019    108.7246
 _Icompany_6 |   102.8886   54.17388     1.90   0.059    -4.056071    209.8333
 _Icompany_7 |   51.46657   58.17922     0.88   0.378    -63.38505    166.3182
 _Icompany_8 |   67.49048   50.97092     1.32   0.187    -33.13125    168.1122
 _Icompany_9 |   30.21752   55.72307     0.54   0.588    -79.78542    140.2204
_Icompany_10 |   126.8371   58.52545     2.17   0.032     11.30197    242.3722
       _cons |  -134.2277   58.29153    -2.30   0.023     -249.301   -19.15433
-------------+----------------------------------------------------------------
     sigma_u |  23.128818
     sigma_e |  51.724523
```

```
         rho |  .16662954   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(19, 169) =     1.40             Prob > F = 0.1309

. drop _I*

. tsset company year
       panel variable:  company (strongly balanced)
        time variable:  year, 1935 to 1954
```

To avoid having twenty year dummies, I swapped the group and time indices, created dummies for the companies, and ran the fixed effects estimator. The estimates for mvalue and kstock aren't changed very much.

c)

```
. reshape wide invest mvalue kstock time, i(year) j(company)
(note: j = 1 2 3 4 5 6 7 8 9 10)

Data                               long   ->   wide
-----------------------------------------------------------------------------
Number of obs.                      200   ->      20
Number of variables                   6   ->      41
j variable (10 values)          company   ->   (dropped)
xij variables:
                                 invest   ->   invest1 invest2 ... invest10
                                 mvalue   ->   mvalue1 mvalue2 ... mvalue10
                                 kstock   ->   kstock1 kstock2 ... kstock10
                                   time   ->   time1 time2 ... time10
-----------------------------------------------------------------------------

. sureg (invest1 mvalue1 kstock1) (invest2 mvalue2 kstock2) (invest3 mvalue3 kstock3) ///
>     (invest4 mvalue4 kstock4) (invest5 mvalue5 kstock5) (invest6 mvalue6 kstock6) ///
>     (invest7 mvalue7 kstock7) (invest8 mvalue8 kstock8) (invest9 mvalue9 kstock9) ///
>     (invest10 mvalue10 kstock10)

Seemingly unrelated regression
----------------------------------------------------------------------
Equation          Obs  Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
invest1            20      2    85.01269    0.9206     284.94   0.0000
invest2            20      2    89.33556    0.4658      32.76   0.0000
invest3            20      2    26.59842    0.6845      51.78   0.0000
invest4            20      2    12.41115    0.9112     208.80   0.0000
invest5            20      2    8.465161    0.6721      38.24   0.0000
invest6            20      2    7.640396    0.9497     447.97   0.0000
invest7            20      2    8.739524    0.7603      80.10   0.0000
```

```
invest8            20      2    9.837619    0.7210       77.06   0.0000
invest9            20      2    8.516405    0.6552       40.80   0.0000
invest10           20      2    1.029905    0.6220       43.88   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
invest1      |
     mvalue1 |   .1138135   .0167455     6.80   0.000     .0809928    .1466342
     kstock1 |   .3861235   .0297382    12.98   0.000     .3278378    .4444092
       _cons |  -135.6061   72.29358    -1.88   0.061    -277.2989    6.086716
-------------+----------------------------------------------------------------
invest2      |
     mvalue2 |   .1627658   .0401126     4.06   0.000     .0841465    .2413851
     kstock2 |   .3406261   .1016225     3.35   0.001     .1414497    .5398026
       _cons |  -10.90599   82.50626    -0.13   0.895    -172.6153    150.8033
-------------+----------------------------------------------------------------
invest3      |
     mvalue3 |   .0349626    .009346     3.74   0.000     .0166448    .0532803
     kstock3 |   .1257302    .020424     6.16   0.000     .0856998    .1657605
       _cons |   -15.8959   20.73065    -0.77   0.443    -56.52724    24.73543
-------------+----------------------------------------------------------------
invest4      |
     mvalue4 |   .0678437   .0159811     4.25   0.000     .0365213    .0991661
     kstock4 |   .3075528   .0253627    12.13   0.000     .2578428    .3572628
       _cons |   1.804334    11.0026     0.16   0.870    -19.76036    23.36903
-------------+----------------------------------------------------------------
invest5      |
     mvalue5 |   .1274473   .0458318     2.78   0.005     .0376187    .2172759
     kstock5 |   .0119871   .0179203     0.67   0.504    -.0231361    .0471104
       _cons |   26.46736   6.065146     4.36   0.000     14.57989    38.35483
-------------+----------------------------------------------------------------
invest6      |
     mvalue6 |   .1333107   .0177916     7.49   0.000     .0984398    .1681817
     kstock6 |   .0540052   .0597388     0.90   0.366    -.0630806    .1710911
       _cons |  -6.193452   3.439334    -1.80   0.072    -12.93442    .5475189
-------------+----------------------------------------------------------------
invest7      |
     mvalue7 |   .1134649   .0457163     2.48   0.013     .0238627    .2030671
     kstock7 |   .1281802   .0146209     8.77   0.000     .0995238    .1568367
       _cons |  -9.770128   8.763982    -1.11   0.265    -26.94722    7.406962
-------------+----------------------------------------------------------------
invest8      |
     mvalue8 |   .0537015   .0082648     6.50   0.000     .0375027    .0699002
     kstock8 |   .0433622   .0345859     1.25   0.210    -.0244249    .1111494
```

```
       _cons |   3.149101   5.056158     0.62   0.533    -6.760786    13.05899
-------------+----------------------------------------------------------------
invest9      |
     mvalue9 |   .0765949   .0210567     3.64   0.000     .0353245    .1178654
     kstock9 |   .0654245   .0219503     2.98   0.003     .0224027    .1084463
       _cons |  -3.156863    7.29719    -0.43   0.665    -17.45909    11.14537
-------------+----------------------------------------------------------------
invest10     |
    mvalue10 |  -.0161291   .0157461    -1.02   0.306    -.0469909    .0147327
    kstock10 |   .3768475   .0573058     6.58   0.000     .2645303    .4891647
       _cons |    1.98935   1.177681     1.69   0.091    -.3188624    4.297562
------------------------------------------------------------------------------

. test [invest1]mvalue1 = [invest2]mvalue2 = [invest3]mvalue3 = [invest4]mvalue4 = ///
>         [invest5]mvalue5 = [invest6]mvalue6 = [invest7]mvalue7 = [invest8]mvalue8 = ///
>         [invest9]mvalue9 = [invest10]mvalue10

 ( 1)  [invest1]mvalue1 - [invest2]mvalue2 = 0
 ( 2)  [invest1]mvalue1 - [invest3]mvalue3 = 0
 ( 3)  [invest1]mvalue1 - [invest4]mvalue4 = 0
 ( 4)  [invest1]mvalue1 - [invest5]mvalue5 = 0
 ( 5)  [invest1]mvalue1 - [invest6]mvalue6 = 0
 ( 6)  [invest1]mvalue1 - [invest7]mvalue7 = 0
 ( 7)  [invest1]mvalue1 - [invest8]mvalue8 = 0
 ( 8)  [invest1]mvalue1 - [invest9]mvalue9 = 0
 ( 9)  [invest1]mvalue1 - [invest10]mvalue10 = 0

           chi2(  9) =   79.36
         Prob > chi2 =    0.0000

. test [invest1]kstock1 = [invest2]kstock2 = [invest3]kstock3 = [invest4]kstock4 = ///
>         [invest5]kstock5 = [invest6]kstock6 = [invest7]kstock7 = [invest8]kstock8 = ///
>         [invest9]kstock9 = [invest10]kstock10

 ( 1)  [invest1]kstock1 - [invest2]kstock2 = 0
 ( 2)  [invest1]kstock1 - [invest3]kstock3 = 0
 ( 3)  [invest1]kstock1 - [invest4]kstock4 = 0
 ( 4)  [invest1]kstock1 - [invest5]kstock5 = 0
 ( 5)  [invest1]kstock1 - [invest6]kstock6 = 0
 ( 6)  [invest1]kstock1 - [invest7]kstock7 = 0
 ( 7)  [invest1]kstock1 - [invest8]kstock8 = 0
 ( 8)  [invest1]kstock1 - [invest9]kstock9 = 0
 ( 9)  [invest1]kstock1 - [invest10]kstock10 = 0

           chi2(  9) =  318.07
         Prob > chi2 =    0.0000
```

SUR allows each company to have its own coefficient vector, whereas fixed effects only allows the constant term to vary. Furthermore, the error structure is richer with SUR; each company can have a different error variance.

The test for cross-equation equality for each of the slope coefficients is rejected, indicating heterogeneity across companies.

2) a)

```
. use ~/Documents/Economics/data/pntsprd.dta

. reg favwin spread, robust

Linear regression                                      Number of obs =     553
                                                       F(  1,   551) =  101.54
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.1107
                                                       Root MSE      =  .40168

------------------------------------------------------------------------------
             |               Robust
      favwin |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      spread |   .0193655   .0019218    10.08   0.000     .0155905    .0231405
       _cons |   .5769492   .0316568    18.23   0.000     .5147664    .6391321
------------------------------------------------------------------------------

. test _cons == 0.5

 ( 1)  _cons = .5

       F(  1,   551) =    5.91
            Prob > F =    0.0154

. lincom _cons + 10*spread

 ( 1)  10 spread + _cons = 0

------------------------------------------------------------------------------
      favwin |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   .7706044   .0166574    46.26   0.000     .7378847    .8033241
------------------------------------------------------------------------------
```

If spread incorporates all the relevant information, then when the spread is zero, there is no information that would indicate whether one or the other team is more likely to win and so the probability of winning is 0.5 Yes, the coefficient on spread is highly statistically significantly different from 0. The estimated probability of a win given a spread of 10 is 0.77.

b)

```
. probit favwin spread

Iteration 0:   log likelihood = -302.74988
Iteration 1:   log likelihood = -266.49244
Iteration 2:   log likelihood = -263.62542
Iteration 3:   log likelihood = -263.56223
Iteration 4:   log likelihood = -263.56219

Probit regression                                 Number of obs   =        553
                                                  LR chi2(1)      =      78.38
                                                  Prob > chi2     =     0.0000
Log likelihood = -263.56219                       Pseudo R2       =     0.1294

------------------------------------------------------------------------------
      favwin |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      spread |    .092463   .0121811     7.59   0.000     .0685885    .1163374
       _cons |  -.0105926   .1037469    -0.10   0.919    -.2139328    .1927476
------------------------------------------------------------------------------

. test _cons

 ( 1)  _cons = 0

           chi2(  1) =    0.01
         Prob > chi2 =    0.9187

. nlcom normal(_b[_cons] + 10*_b[spread])

       _nl_1:  normal(_b[_cons] + 10*_b[spread])

------------------------------------------------------------------------------
      favwin |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   .8196513   .0194412    42.16   0.000     .7815473    .8577553
------------------------------------------------------------------------------

. mfx, at (spread = 10)

Marginal effects after probit
      y  = Pr(favwin) (predict)
         =  .81965134
```

```
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
  spread |   .0242918      .00271    8.95   0.000   .018971  .029613        10
------------------------------------------------------------------------------
```

When the intercept is zero and spread is zero, the probability is 0.5. So, we see again the reasoning in part a) that the intercept should reflect whether spread incorporates all relevant information. Using the probit model, we cannot reject the hypothesis that spread incorporates all relevant information, which we could using the linear probability model. The estimated probability using the probit model is 0.82 when spread is 10. The marginal effect is 0.024.

c)

```
. probit favwin spread favhome fav25 und25

Iteration 0:   log likelihood = -302.74988
Iteration 1:   log likelihood = -265.47417
Iteration 2:   log likelihood = -262.70317
Iteration 3:   log likelihood = -262.64181
Iteration 4:   log likelihood = -262.64177

Probit regression                                 Number of obs   =        553
                                                  LR chi2(4)      =      80.22
                                                  Prob > chi2     =     0.0000
Log likelihood = -262.64177                       Pseudo R2       =     0.1325

------------------------------------------------------------------------------
      favwin |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      spread |   .0878845   .0129491     6.79   0.000     .0625047    .1132642
     favhome |   .1485753   .1370571     1.08   0.278    -.1200517    .4172024
       fav25 |    .003068     .15869     0.02   0.985    -.3079587    .3140946
       und25 |  -.2198082   .2505842    -0.88   0.380    -.7109443    .2713278
       _cons |  -.0551801    .128763    -0.43   0.668    -.3075509    .1971907
------------------------------------------------------------------------------

. test favhome = fav25 = und25 = 0

 ( 1)  favhome - fav25 = 0
 ( 2)  favhome - und25 = 0
 ( 3)  favhome = 0

           chi2(  3) =    1.84
         Prob > chi2 =    0.6054
```

The test reveals that the coefficients of the additional variables is not statistically significantly different from zero, which means we cannot reject the hypothesis that spread incorporates all relevant information.

3) a)

```
. use ~/Documents/Economics/data/loanapp.dta

. probit approve white

Iteration 0:   log likelihood = -740.34659
Iteration 1:   log likelihood = -701.33221
Iteration 2:   log likelihood = -700.87747
Iteration 3:   log likelihood = -700.87744

Probit regression                                 Number of obs   =       1989
                                                  LR chi2(1)      =      78.94
                                                  Prob > chi2     =     0.0000
Log likelihood = -700.87744                       Pseudo R2       =     0.0533

------------------------------------------------------------------------------
     approve |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       white |   .7839465   .0867118     9.04   0.000     .6139946    .9538984
       _cons |   .5469464    .075435     7.25   0.000     .3990964    .6947963
------------------------------------------------------------------------------

. // probit estimate of probability of approval if white
. nlcom normal(_b[_cons] + _b[white])

       _nl_1:  normal(_b[_cons] + _b[white])

------------------------------------------------------------------------------
     approve |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   .9083879    .007036   129.10   0.000     .8945975    .9221782
------------------------------------------------------------------------------

. // probit estimate of probability of approval if black
. nlcom normal(_b[_cons])

       _nl_1:  normal(_b[_cons])

------------------------------------------------------------------------------
     approve |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   .7077922   .0259133    27.31   0.000      .657003    .7585814
```

```
------------------------------------------------------------------------------

. reg approve white

      Source |       SS       df       MS              Number of obs =    1989
-------------+------------------------------           F(  1,  1987) =  102.23
       Model |  10.4743407     1  10.4743407           Prob > F      =  0.0000
    Residual |   203.59303  1987  .102462521           R-squared     =  0.0489
-------------+------------------------------           Adj R-squared =  0.0485
       Total |  214.067371  1988  .107679764           Root MSE      =   .3201

------------------------------------------------------------------------------
     approve |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       white |   .2005957     .01984    10.11   0.000     .1616864     .239505
       _cons |   .7077922   .0182393    38.81   0.000     .6720221    .7435623
------------------------------------------------------------------------------

. // linear model estimate of probability of approval if white
. lincom _cons + white

 ( 1)  white + _cons = 0

------------------------------------------------------------------------------
     approve |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   .9083879   .0078073   116.35   0.000     .8930766    .9236991
------------------------------------------------------------------------------

. // linear model estimate of probability of approval if black
. lincom _cons

 ( 1)  _cons = 0

------------------------------------------------------------------------------
     approve |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   .7077922   .0182393    38.81   0.000     .6720221    .7435623
------------------------------------------------------------------------------
```

The estimates of the probability of loan approval conditional on race from the probit model and the linear probability model are identical. There appears to be a strong effect of race on this probability.

b)

```
. probit approve white hrat obrat loanprc unem male married dep ///
```

```
>         sch cosign chist pubrec mortlat1 mortlat2 vr

Iteration 0:   log likelihood = -737.97933
Iteration 1:   log likelihood = -604.00737
Iteration 2:   log likelihood = -600.29775
Iteration 3:   log likelihood = -600.27099
Iteration 4:   log likelihood = -600.27099

Probit regression                                 Number of obs   =       1971
                                                  LR chi2(15)     =     275.42
                                                  Prob > chi2     =     0.0000
Log likelihood = -600.27099                       Pseudo R2       =     0.1866

------------------------------------------------------------------------------
     approve |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       white |   .5202525   .0969588     5.37   0.000     .3302168    .7102883
        hrat |   .0078763   .0069616     1.13   0.258    -.0057682    .0215209
       obrat |  -.0276924   .0060493    -4.58   0.000    -.0395488    -.015836
     loanprc |  -1.011969   .2372396    -4.27   0.000     -1.47695   -.5469882
        unem |  -.0366849   .0174807    -2.10   0.036    -.0709464   -.0024234
        male |  -.0370014   .1099273    -0.34   0.736    -.2524549    .1784521
     married |   .2657469   .0942523     2.82   0.005     .0810159    .4504779
         dep |  -.0495756   .0390573    -1.27   0.204    -.1261266    .0269753
         sch |   .0146496   .0958421     0.15   0.879    -.1731974    .2024967
      cosign |   .0860713   .2457509     0.35   0.726    -.3955917    .5677343
       chist |   .5852812   .0959715     6.10   0.000     .3971806    .7733818
      pubrec |  -.7787405     .12632    -6.16   0.000    -1.026323   -.5311578
    mortlat1 |  -.1876236   .2531127    -0.74   0.459    -.6837153     .308468
    mortlat2 |  -.4943562   .3265563    -1.51   0.130    -1.134395    .1456823
          vr |  -.2010621   .0814934    -2.47   0.014    -.3607862    -.041338
       _cons |   2.062327   .3131763     6.59   0.000     1.448512    2.676141
------------------------------------------------------------------------------
```

The coefficient on white is statistically significantly different from 0, which would indicate discrimination against nonwhites.

c)

```
. mfx, varlist(white)

Marginal effects after probit
      y  = Pr(approve) (predict)
         =  .91065604
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
```

```
---------+--------------------------------------------------------------------
   white*|    .105747      .02386    4.43   0.000   .058988  .152506   .846271
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1

. logit approve white hrat obrat loanprc unem male married dep ///
>         sch cosign chist pubrec mortlat1 mortlat2 vr

Iteration 0:   log likelihood = -737.97933
Iteration 1:   log likelihood = -634.97536
Iteration 2:   log likelihood = -601.41194
Iteration 3:   log likelihood = -600.49724
Iteration 4:   log likelihood = -600.49616

Logistic regression                               Number of obs   =       1971
                                                  LR chi2(15)     =     274.97
                                                  Prob > chi2     =     0.0000
Log likelihood = -600.49616                       Pseudo R2       =     0.1863

------------------------------------------------------------------------------
     approve |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       white |   .9377643   .1729041     5.42   0.000     .5988784     1.27665
        hrat |   .0132631   .0128802     1.03   0.303    -.0119816    .0385078
       obrat |  -.0530338   .0112803    -4.70   0.000    -.0751427   -.0309249
     loanprc |  -1.904951   .4604412    -4.14   0.000      -2.8074   -1.002503
        unem |  -.0665789   .0328086    -2.03   0.042    -.1308825   -.0022753
        male |  -.0663851   .2064288    -0.32   0.748    -.4709781    .3382078
     married |   .5032817    .177998     2.83   0.005     .1544121    .8521513
         dep |  -.0907336   .0733341    -1.24   0.216    -.2344657    .0529986
         sch |   .0412287   .1784035     0.23   0.817    -.3084356    .3908931
      cosign |    .132059   .4460933     0.30   0.767    -.7422677    1.006386
       chist |   1.066577   .1712117     6.23   0.000      .731008    1.402146
      pubrec |  -1.340665   .2173657    -6.17   0.000    -1.766694   -.9146362
    mortlat1 |  -.3098821   .4635193    -0.67   0.504    -1.218363     .598599
    mortlat2 |  -.8946755   .5685807    -1.57   0.116    -2.009073    .2197222
          vr |  -.3498279   .1537248    -2.28   0.023    -.6511231   -.0485328
       _cons |    3.80171   .5947054     6.39   0.000     2.636109    4.967311
------------------------------------------------------------------------------

. mfx, varlist(white)

Marginal effects after logit
      y  = Pr(approve) (predict)
         =   .9141792
------------------------------------------------------------------------------
```

```
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
   white*|   .0967431      .02249    4.30   0.000   .052659  .140828   .846271
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

The coefficient is markedly different, but a meaningful comparison of coefficients across logit and probit models cannot be made. However, the marginal effects can be compared. The probit model yields a larger marginal effect than does the logit model, but they are close in value.

d)

```
. qui probit approve white hrat obrat loanprc unem male married ///
>         dep sch cosign chist pubrec mortlat1 mortlat2 vr

. margeff

Average partial effects after probit
      y  = Pr(approve)

------------------------------------------------------------------------------
    variable |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       white |   .1042245   .0149126     6.99   0.000     .0749962    .1334527
        hrat |   .0013078   .0011553     1.13   0.258    -.0009566    .0035723
       obrat |  -.0045983   .0010014    -4.59   0.000    -.0065611   -.0026355
     loanprc |  -.1680353   .0392957    -4.28   0.000    -.2450534   -.0910172
        unem |  -.0060914   .0028994    -2.10   0.036    -.0117741   -.0004088
        male |  -.0060727    .018379    -0.33   0.741    -.0420948    .0299494
     married |   .0459954   .0141381     3.25   0.001     .0182853    .0737055
         dep |  -.0082333   .0064869    -1.27   0.204    -.0209474    .0044808
         sch |   .0024421   .0158594     0.15   0.878    -.0286417    .0335259
      cosign |   .0137141   .0374253     0.37   0.714    -.0596381    .0870662
       chist |   .1196853   .0145369     8.23   0.000     .0911936     .148177
      pubrec |  -.1804182   .0379961    -4.75   0.000    -.2548892   -.1059473
    mortlat1 |  -.0340738    .050086    -0.68   0.496    -.1322406     .064093
    mortlat2 |  -.1030246    .082264    -1.25   0.210    -.2642591    .0582099
          vr |  -.0339854   .0151069    -2.25   0.024    -.0635943   -.0043765
------------------------------------------------------------------------------
```

The estimate does not differ very much, but in principle could. Whereas the **mfx** command calculates the marginal effect at the "average" individual, the **margeff** command determines the marginal effect for each individual, and then averages across each of these.

4) a)

```
. use ~/Documents/Economics/data/fringe.dta

. gen zeropen = 0

. replace zeropen = 1 if pension == 0
(172 real changes made)

. reg zeropen

      Source |       SS       df       MS              Number of obs =     616
-------------+------------------------------           F(  0,   615) =    0.00
       Model |           0     0           .           Prob > F      =       .
    Residual |  123.974026   615  .201583782           R-squared     =  0.0000
-------------+------------------------------           Adj R-squared =  0.0000
       Total |  123.974026   615  .201583782           Root MSE      =  .44898

------------------------------------------------------------------------------
     zeropen |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   .2792208     .01809    15.44   0.000     .2436952    .3147464
------------------------------------------------------------------------------

. summ pension if !zeropen

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     pension |       444    905.0439    550.3696       7.28    2880.27
```

The pension is 0 for about 28% of the workers in the sample. The range of non-zero pensions is from 7.28 to 2880.27. The Tobit model is appropriate because we have censoring, since some workers might work at jobs that don't offer pension benefits rather than have no pension amount at jobs that do offer these benefits.

b)

```
. tobit pension exper age tenure educ depends married white male, ll(0)

Tobit regression                                  Number of obs   =        616
                                                  LR chi2(8)      =     184.70
                                                  Prob > chi2     =     0.0000
Log likelihood = -3672.9635                       Pseudo R2       =     0.0245

------------------------------------------------------------------------------
     pension |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       exper |   5.203458    6.00928     0.87   0.387    -6.598007    17.00492
         age |  -4.638944   5.710741    -0.81   0.417    -15.85412    6.576228
```

```
      tenure |   36.02385   4.564348     7.89   0.000     27.06005    44.98765
        educ |   93.21262   10.89133     8.56   0.000     71.82343    114.6018
     depends |   35.28461   21.91691     1.61   0.108    -7.757432    78.32666
     married |   53.68858   71.73266     0.75   0.454    -87.18528    194.5624
       white |   144.0855   102.0753     1.41   0.159    -56.37738    344.5485
        male |   308.1505    69.8903     4.41   0.000     170.8948    445.4062
       _cons |  -1252.429   219.0692    -5.72   0.000    -1682.653   -822.2048
-------------+----------------------------------------------------------------
      /sigma |   677.7383   24.13815                      630.3341    725.1426
------------------------------------------------------------------------------
  Obs. summary:        172  left-censored observations at pension<=0
                       444     uncensored observations
                         0 right-censored observations
```

Males do have statistically significantly higher benefits that females, but whites do not.

c)

```
. mfx, predict(e(0,.)) at(white = 0 male = 0 age = 35 married = 0 depends = 0 ///
>         educ = 16 exper = 10)

warning: no value assigned in at() for variables tenure;
   means used for tenure

Marginal effects after tobit
      y  = E(pension|pension>0) (predict, e(0,.))
         =  718.46441
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
   exper |   2.675663     3.04738    0.88   0.380   -3.2971  8.64842        10
     age |  -2.385385     2.90285   -0.82   0.411  -8.07486  3.30409        35
  tenure |   18.52377     3.12086    5.94   0.000    12.407  24.6405   7.75162
    educ |   47.93073     8.51085    5.63   0.000   31.2498  64.6117        16
 depends |   18.14365       11.07    1.64   0.101  -3.55406  39.8414         0
 married*|   28.20525      38.106    0.74   0.459  -46.4804  102.891         0
   white*|   78.44177      53.152    1.48   0.140  -25.7335  182.617         0
    male*|   178.6284      46.694    3.83   0.000   87.1108  270.146         0
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1

. mfx, predict(e(0,.)) at(white = 0 male = 1 age = 35 married = 0 depends = 0 ///
>         educ = 16 exper = 10)

warning: no value assigned in at() for variables tenure;
   means used for tenure
```

```
Marginal effects after tobit
      y  = E(pension|pension>0) (predict, e(0,.))
         =  897.09282
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
   exper |   3.363472     3.77305    0.89   0.373  -4.03157  10.7585        10
     age |  -2.998575     3.60197   -0.83   0.405  -10.0583  4.06116        35
  tenure |   23.28552      3.8088    6.11   0.000   15.8204  30.7506   7.75162
    educ |   60.25186      10.241    5.88   0.000   40.1794  80.3244        16
 depends |   22.80768      13.864    1.65   0.100  -4.36577  49.9811         0
 married*|   35.33099      47.248    0.75   0.455   -57.274  127.936         0
   white*|   97.63182      66.421    1.47   0.142  -32.5517  227.815         0
    male*|   178.6284      46.694    3.83   0.000   87.1108  270.146         1
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1

. mfx, predict(e(0,.)) at(white = 1 male = 0 age = 35 married = 0 depends = 0 ///
>          educ = 16 exper = 10)

warning: no value assigned in at() for variables tenure;
   means used for tenure

Marginal effects after tobit
      y  = E(pension|pension>0) (predict, e(0,.))
         =  796.90618
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
   exper |   2.992289     3.41108    0.88   0.380  -3.69331  9.67788        10
     age |   -2.66766     3.24648   -0.82   0.411  -9.03065  3.69533        35
  tenure |   20.71579     3.08734    6.71   0.000   14.6647  26.7669   7.75162
    educ |   53.60263     8.10036    6.62   0.000   37.7262   69.479        16
 depends |   20.29069       12.56    1.62   0.106  -4.32687  44.9082         0
 married*|   31.49666      42.191    0.75   0.455  -51.1964   114.19         0
   white*|   78.44177      53.152    1.48   0.140  -25.7335  182.617         1
    male*|   197.8185      48.388    4.09   0.000    102.98  292.657         0
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

The marginal effect is calculated first for a nonwhite female. One could just sum up the marginal effect of being white and that of being male at this point, and one would obtain approximately 257. However, this would ignore the cross-derivative effect, and so a better estimate would be obtained by changing either one of male or white and then calculating the marginal effect at this point. There are two paths that lead from nonwhite female to white male. The first estimate (first the effect from female to male and then nonwhite to white) yields $178.63 + 97.63 =$

276.26, whereas the second estimate (first the effect from nonwhite to white and then from female to male) yields 78.44 + 197.82 = 276.26.

If one interprets the question as asking for the the unconditional expectation of the difference in pensions, then the answer one obtains can be different.

d)

```
. tobit pension exper age tenure educ depends married white male union, ll(0)

Tobit regression                                  Number of obs   =        616
                                                  LR chi2(9)      =     233.52
                                                  Prob > chi2     =     0.0000
Log likelihood = -3648.5515                       Pseudo R2       =     0.0310

------------------------------------------------------------------------------
     pension |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       exper |   4.393523   5.830946     0.75   0.451    -7.057754     15.8448
         age |  -1.653532   5.555708    -0.30   0.766    -12.56427    9.257211
      tenure |   28.77837   4.504963     6.39   0.000     19.93116    37.62557
        educ |   106.8277   10.77274     9.92   0.000     85.67134    127.9841
     depends |   41.46623   21.21414     1.95   0.051    -.1957922    83.12824
     married |   19.74554   69.50047     0.28   0.776     -116.745    156.2361
       white |   159.2972   98.96747     1.61   0.108    -35.06298    353.6575
        male |   257.2457   68.02051     3.78   0.000     123.6615    390.8298
       union |    439.046   62.48832     7.03   0.000     316.3265    561.7656
       _cons |  -1571.506   218.5445    -7.19   0.000    -2000.701   -1142.311
-------------+----------------------------------------------------------------
      /sigma |   652.8974   23.16287                      607.4083    698.3865
------------------------------------------------------------------------------
  Obs. summary:        172  left-censored observations at pension<=0
                       444     uncensored observations
                         0 right-censored observations
```

Union membership increases the pension significantly, both statistically and economically. The effect of being white is still not statistically significantly different from 0.

5) a)

```
. use ~/Documents/Economics/data/mroz.dta

. reg lwage educ exper expersq if inlf == 1

      Source |       SS       df       MS              Number of obs =     428
-------------+------------------------------           F(  3,   424) =   26.29
```

```
       Model |  35.0223023     3  11.6741008           Prob > F      =  0.0000
    Residual |  188.305149   424  .444115917           R-squared     =  0.1568
-------------+------------------------------           Adj R-squared =  0.1509
       Total |  223.327451   427  .523015108           Root MSE      =  .66642

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .1074896   .0141465     7.60   0.000     .0796837    .1352956
       exper |   .0415665   .0131752     3.15   0.002     .0156697    .0674633
     expersq |  -.0008112   .0003932    -2.06   0.040    -.0015841   -.0000382
       _cons |  -.5220407   .1986321    -2.63   0.009    -.9124668   -.1316145
------------------------------------------------------------------------------
```

The return to education is statistically significantly different from 0, estimated at 0.107

b)

```
. heckman lwage educ exper expersq, twostep select(nwifeinc age kidslt6 ///
>         kidsge6 educ exper expersq)

Heckman selection model -- two-step estimates   Number of obs      =       753
(regression model with sample selection)        Censored obs       =       325
                                                Uncensored obs     =       428

                                                Wald chi2(6)       =    180.10
                                                Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lwage        |
        educ |   .1090655    .015523     7.03   0.000     .0786411      .13949
       exper |   .0438873   .0162611     2.70   0.007     .0120163    .0757584
     expersq |  -.0008591   .0004389    -1.96   0.050    -.0017194    1.15e-06
       _cons |  -.5781033   .3050062    -1.90   0.058    -1.175904    .0196979
-------------+----------------------------------------------------------------
select       |
    nwifeinc |  -.0120237   .0048398    -2.48   0.013    -.0215096   -.0025378
         age |  -.0528527   .0084772    -6.23   0.000    -.0694678   -.0362376
     kidslt6 |  -.8683285   .1185223    -7.33   0.000    -1.100628    -.636029
     kidsge6 |    .036005   .0434768     0.83   0.408     -.049208    .1212179
        educ |   .1309047   .0252542     5.18   0.000     .0814074     .180402
       exper |   .1233476   .0187164     6.59   0.000     .0866641    .1600311
     expersq |  -.0018871      .0006    -3.15   0.002     -.003063   -.0007111
       _cons |   .2700768    .508593     0.53   0.595    -.7267472    1.266901
```

```
-------------+----------------------------------------------------------------
mills        |
      lambda |   .0322619   .1336246     0.24   0.809    -.2296376    .2941613
-------------+----------------------------------------------------------------
         rho |    0.04861
       sigma |  .66362876
      lambda |  .03226186   .1336246
------------------------------------------------------------------------------
```

The estimate for the return to education is quite close to that found by the OLS regression. The coefficient on the inverse Mills ratio, denote λ is not statistically significantly different from 0, indicating that selection bias is not a problem.

c)

```
. heckman lwage educ exper expersq, select(nwifeinc age kidslt6 kidsge6 ///
>         educ exper expersq)

Iteration 0:   log likelihood = -832.89777
Iteration 1:   log likelihood =  -832.8851
Iteration 2:   log likelihood = -832.88509

Heckman selection model                         Number of obs      =       753
(regression model with sample selection)        Censored obs       =       325
                                                Uncensored obs     =       428

                                                Wald chi2(3)       =     59.67
Log likelihood = -832.8851                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lwage        |
        educ |   .1083502   .0148607     7.29   0.000     .0792238    .1374767
       exper |   .0428369   .0148785     2.88   0.004     .0136755    .0719983
     expersq |  -.0008374   .0004175    -2.01   0.045    -.0016556   -.0000192
       _cons |  -.5526974   .2603784    -2.12   0.034     -1.06303   -.0423652
-------------+----------------------------------------------------------------
select       |
    nwifeinc |  -.0121321   .0048767    -2.49   0.013    -.0216903    -.002574
         age |  -.0528287   .0084792    -6.23   0.000    -.0694476   -.0362098
     kidslt6 |  -.8673988   .1186509    -7.31   0.000     -1.09995   -.6348472
     kidsge6 |   .0358723   .0434753     0.83   0.409    -.0493377    .1210824
        educ |   .1313415   .0253823     5.17   0.000     .0815931    .1810899
       exper |   .1232818   .0187242     6.58   0.000     .0865831    .1599806
     expersq |  -.0018863   .0006004    -3.14   0.002     -.003063   -.0007095
```

```
       _cons |   .2664491   .5089578     0.52   0.601    -.7310898    1.263988
-------------+----------------------------------------------------------------
     /athrho |    .026614    .147182     0.18   0.857    -.2618573    .3150854
    /lnsigma |  -.4103809   .0342291   -11.99   0.000    -.4774687   -.3432931
-------------+----------------------------------------------------------------
         rho |   .0266078   .1470778                     -.2560319    .3050564
       sigma |   .6633975   .0227075                      .6203517    .7094303
      lambda |   .0176515   .0976057                     -.1736521    .2089552
------------------------------------------------------------------------------
LR test of indep. eqns. (rho = 0):   chi2(1) =     0.03   Prob > chi2 = 0.8577
------------------------------------------------------------------------------
```

The estimate for return to education is pretty much the same, and the likelihood ratio test fails to reject the null hypothesis that ρ is statistically significantly different from 0, which implies that ignoring selection is acceptable.

6) a)

```
. use ~/Documents/Economics/data/cps91.dta

. summ inlf

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        inlf |      5634    .5832446    .4930654          0          1

. reg lwage educ exper expersq black hispanic

      Source |       SS       df       MS              Number of obs =    3286
-------------+------------------------------           F(  5,  3280) =  169.08
       Model |  185.581829     5  37.1163657           Prob > F      =  0.0000
    Residual |  720.007572  3280  .219514504           R-squared     =  0.2049
-------------+------------------------------           Adj R-squared =  0.2037
       Total |  905.589401  3285  .275674095           Root MSE      =  .46852

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0991502   .0035898    27.62   0.000     .0921118    .1061887
       exper |   .0198554   .0032856     6.04   0.000     .0134133    .0262974
     expersq |  -.0003489    .000077    -4.53   0.000    -.0004999   -.0001979
       black |  -.0295532   .0343431    -0.86   0.390    -.0968892    .0377828
    hispanic |   .0136158   .0363565     0.37   0.708    -.0576679    .0848996
       _cons |    .648842   .0599659    10.82   0.000     .5312675    .7664164
------------------------------------------------------------------------------
```

About 58% of working women report being in the labor force. Neither race nor ethnicity seem to cause wage differentials.

b)

```
. probit inlf educ exper expersq black hispanic nwifeinc kidlt6

Iteration 0:   log likelihood =  -3826.743
Iteration 1:   log likelihood =  -3539.706
Iteration 2:   log likelihood = -3537.2549
Iteration 3:   log likelihood = -3537.2544

Probit regression                                 Number of obs   =       5634
                                                  LR chi2(7)      =     578.98
                                                  Prob > chi2     =     0.0000
Log likelihood = -3537.2544                       Pseudo R2       =     0.0756

------------------------------------------------------------------------------
        inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0964837   .0077854    12.39   0.000     .0812246    .1117428
       exper |   .0077141   .0072385     1.07   0.287    -.0064732    .0219014
     expersq |  -.0006143   .0001577    -3.90   0.000    -.0009234   -.0003052
       black |   .0167548   .0755896     0.22   0.825    -.1313981    .1649077
    hispanic |  -.1219554   .0704695    -1.73   0.084     -.260073    .0161623
    nwifeinc |  -.0091239   .0006775   -13.47   0.000    -.0104518    -.007796
      kidlt6 |   -.500167   .0452776   -11.05   0.000    -.5889096   -.4114245
       _cons |  -.4393231   .1338545    -3.28   0.001    -.7016732    -.176973
------------------------------------------------------------------------------
```

I would expect that having young children in the house would reduce the probability of working, as would having greater household income from other sources. This expectation is borne out by the probit estimates, both of which are statistically significant.

To test for selection, we need to find factors that help explain the decision to work. However, these factors cannot also explain the amount that is earned if in fact a woman decides to work i.e. they have to be exogenous to the process that determines wages in order to identify the selection effect.

c)

```
. // calculate the nonselection hazard (i.e. the inverse mills ratio)
. predict latent, xb

. gen imr = normalden(latent)/normal(latent)

. reg lwage educ exper expersq black hispanic imr
```

```
      Source |       SS       df       MS              Number of obs =    3286
-------------+------------------------------           F(  6,  3279) =  141.52
       Model |  186.268673     6  31.0447788           Prob > F      =  0.0000
    Residual |  719.320728  3279  .219371982           R-squared     =  0.2057
-------------+------------------------------           Adj R-squared =  0.2042
       Total |  905.589401  3285  .275674095           Root MSE      =  .46837

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .1032796   .0042807    24.13   0.000     .0948865    .1116726
       exper |   .0204788   .0033034     6.20   0.000     .0140019    .0269557
     expersq |  -.0003781   .0000787    -4.80   0.000    -.0005325   -.0002237
       black |  -.0251464   .0344221    -0.73   0.465    -.0926374    .0423447
    hispanic |   .0056534   .0366222     0.15   0.877    -.0661514    .0774581
         imr |   .0918995   .0519368     1.77   0.077    -.0099322    .1937313
       _cons |    .538856   .0863552     6.24   0.000     .3695403    .7081716
------------------------------------------------------------------------------

. // test to see if we get the same estimation results has the heckman twostep
. heckman lwage educ exper expersq black hispanic, twostep select(educ exper ///
>         expersq black hispanic nwifeinc kidlt6)

Heckman selection model -- two-step estimates   Number of obs      =      5634
(regression model with sample selection)        Censored obs       =      2348
                                                Uncensored obs     =      3286

                                                Wald chi2(10)      =   1015.38
                                                Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lwage        |
        educ |   .1032796   .0042981    24.03   0.000     .0948553    .1117038
       exper |   .0204788   .0033184     6.17   0.000     .0139749    .0269827
     expersq |  -.0003781   .0000791    -4.78   0.000     -.000533   -.0002231
       black |  -.0251464   .0345951    -0.73   0.467    -.0929515    .0426588
    hispanic |   .0056534   .0367559     0.15   0.878    -.0663869    .0776936
       _cons |    .538856   .0867082     6.21   0.000      .368911     .708801
-------------+----------------------------------------------------------------
select       |
        educ |   .0964837   .0077854    12.39   0.000     .0812246    .1117428
       exper |   .0077141   .0072385     1.07   0.287    -.0064732    .0219014
     expersq |  -.0006143   .0001577    -3.90   0.000    -.0009234   -.0003052
       black |   .0167548   .0755896     0.22   0.825    -.1313981    .1649077
```

```
    hispanic |  -.1219554   .0704695    -1.73   0.084     -.260073    .0161623
    nwifeinc |  -.0091239   .0006775   -13.47   0.000    -.0104518   -.007796
      kidlt6 |   -.500167   .0452776   -11.05   0.000    -.5889096   -.4114245
       _cons |  -.4393231   .1338545    -3.28   0.001    -.7016732   -.176973
-------------+----------------------------------------------------------------
mills        |
      lambda |   .0918995   .0521032     1.76   0.078    -.0102208    .1940199
-------------+----------------------------------------------------------------
         rho |    0.19437
       sigma |  .47280408
      lambda |  .09189953   .0521032
------------------------------------------------------------------------------
```

The two-sided p-value is 0.077. The addition of the inverse Mills ratio (non-selection hazard) does not change the results from the OLS regression very much. Thus, while there is an indication of selection, since at the 10% level the coefficient on IMR is significant, the effect of this selection on the coefficients is minimal. Our conclusions about discrimination (or lack thereof) are unchanged.

Notice that we essentially ran a two-step Heckman model.