TESTS OF COEFFICIENTS STATA LAB ANSWERS
Professor Lemke
Fall 2008

```
. * PROBLEM ONE;

. sum;

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        wage |       704     29857.9    35136.31       1086     293787
      lnwage |       704    9.773959    1.079598   6.990256   12.59061
         age |       704    48.89915    13.76339         25         74
        age2 |       704    2580.288    1364.125        625       5476
      school |       704    13.58807    2.939054          9         21
    hispanic |       704    .1775568    .3824109          0          1
       black |       704    .1136364    .3175947          0          1
       white |       704    .7088068    .4546354          0          1
      female |       704    .3508523    .4775761          0          1
        male |       704    .6491477    .4775761          0          1
-------------+--------------------------------------------------------


. * PROBLEM TWO;

. reg wage age age2 school hispanic black female;

      Source |       SS       df       MS              Number of obs =     704
-------------+------------------------------           F(  6,   697) =   25.19
       Model |  1.5465e+11     6  2.5775e+10           Prob > F      =  0.0000
    Residual |  7.1325e+11   697  1.0233e+09           R-squared     =  0.1782
-------------+------------------------------           Adj R-squared =  0.1711
       Total |  8.6790e+11   703  1.2346e+09           Root MSE      =   31989

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   5852.196   682.3649     8.58   0.000     4512.459    7191.933
        age2 |   -64.5738   6.886513    -9.38   0.000     -78.0946   -51.05301
      school |   1748.316   411.1296     4.25   0.000     941.1149    2555.517
    hispanic |  -4565.543   3211.182    -1.42   0.156    -10870.29    1739.206
       black |  -11479.98   3864.626    -2.97   0.003    -19067.69   -3892.281
      female |  -7806.004   2535.325    -3.08   0.002    -12783.79   -2828.214
       _cons |  -108592.8   17109.92    -6.35   0.000    -142185.9    -74999.6
------------------------------------------------------------------------------
```

These results make sense generally: there is an increasing then decreasing return to age, a positive effect from schooling, a negative racial effect but less so for Hispanics, and a substantial gender differential going against women.

```
. sum errors1;

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     errors1 |       704    .0000493    31852.43  -44977.78   259153.5
-------------+--------------------------------------------------------
```

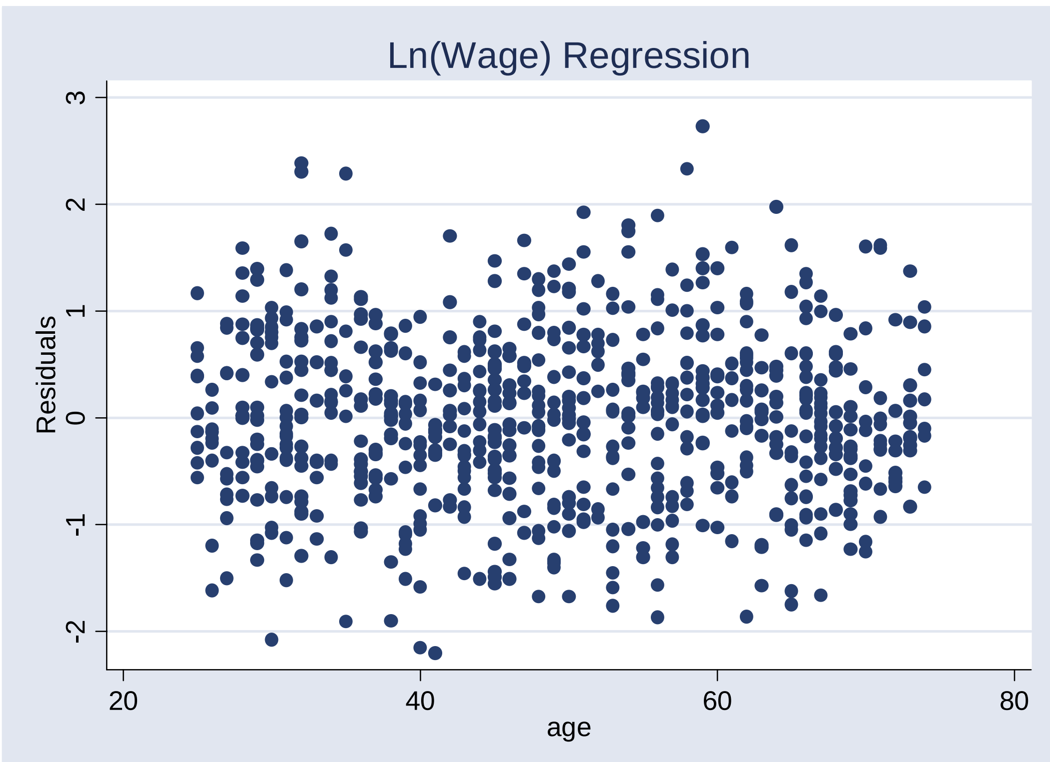
The summary statistics of the residuals are immediately troublesome as the lowest differential is about -$45,000 while the greatest differential is almost six times that at almost $260,000. The worry, of course is possible heteroskedasticity. To look at this further, consider the graph of the residuals against age. One can see that the errors are much more erratic at some ages and the positive errors are much more erratic than the negative errors. This is troublesome.


```
. * PROBLEM THREE;
. reg lnwage age age2 school hispanic black female;

      Source |       SS       df       MS              Number of obs =     704
-------------+------------------------------           F(  6,   697) =   97.67
       Model |  374.244031     6  62.3740052           Prob > F      =  0.0000
    Residual |  445.125611   697  .638630719           R-squared     =  0.4567
-------------+------------------------------           Adj R-squared =  0.4521
       Total |  819.369642   703  1.16553292           Root MSE      =  .79914

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .2909222   .0170466    17.07   0.000     .2574533     .324391
        age2 |   -.003262    .000172   -18.96   0.000    -.0035997   -.0029242
      school |   .0665694   .0102707     6.48   0.000     .0464042    .0867346
    hispanic |  -.1809396   .0802206    -2.26   0.024    -.3384426   -.0234366
       black |  -.2682951   .0965447    -2.78   0.006    -.4578484   -.0787418
      female |  -.1831355   .0633366    -2.89   0.004    -.3074889   -.0587821
       _cons |   3.187207    .427434     7.46   0.000     2.347995     4.02642
------------------------------------------------------------------------------
```

These results make sense generally: there is an increasing then decreasing return to age, a positive effect from schooling, a negative racial effect but less so for Hispanics, and a substantial gender differential going against women.

```
. sum errors2;

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     errors2 |       704    3.17e-10    .7957261  -2.203019   2.728446
-------------+--------------------------------------------------------
```

The summary statistics of the residuals are much improved as the minimum error is -2.2 while the maximum error is 2.73. One can further see that the errors are distributed roughly with the same variance across all age groups (as well as for positive and negative errors) by looking at the graph.



```
. * PROBLEM FOUR;

. reg lnwage age age2 school hispanic black female;

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .2909222   .0170466    17.07   0.000     .2574533     .324391
        age2 |   -.003262    .000172   -18.96   0.000    -.0035997   -.0029242
      school |   .0665694   .0102707     6.48   0.000     .0464042    .0867346
    hispanic |  -.1809396   .0802206    -2.26   0.024    -.3384426   -.0234366
       black |  -.2682951   .0965447    -2.78   0.006    -.4578484   -.0787418
      female |  -.1831355   .0633366    -2.89   0.004    -.3074889   -.0587821
       _cons |   3.187207    .427434     7.46   0.000     2.347995     4.02642
------------------------------------------------------------------------------
```

* QUESTION A;

We fail to reject the claim that there is no difference in wages between blacks and Hispanics as the *F*-statistic (0.58) is associated with a high *p*-value of 44.73%. (We would need the *p*-value to be under 10% to reject the claim.)

* QUESTION B;

We reject the claim that there is no effect on wages from race as the *F*-statistic (5.52) is associated with a high *p*-value of 0.42%.

Notice how the test command is different from (A). In (A), the command was simply:

```
test black=hispanic
```

to see if there is a difference between these two races. In (B), however, we must also test to whites. As white is the omitted race, this requires a test command of

```
test black=hispanic=0.
```

* QUESTION C;

We reject the claim that the return to each year of schooling is 9 percent as the *F*-statistic is 5.20 which is associated with a *p*-value of 2.28%.

* QUESTION D;

We fail to reject the claim that the gender differential is negative ten percent as the *F*-statistic (1.72) is associated with a high *p*-value of 18.98%.

* QUESTION E;

Because the coefficient on age is positive but the coefficient on age-squared is negative, we know that the predicted relationship is a positive one (at a decreasing rate) up to a point. After that point, wages are decreasing at an increasing rate. Specifically, wages increase up to the point where

$$0.291 - 2 * 0.003262 * age = 0$$

which happens at an age of roughly 43.06 years-old.

```
. * PROBLEM FIVE;
```

The model is:

$$lnwage = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 school + \beta_4 hispanic + \beta_5 black + \beta_6 female + \varepsilon.$$

Imposing the restriction that β_3=0.09 produces a restricted model of:

$$lnwage - 0.09school = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_4 hispanic + \beta_5 black + \beta_6 female + \varepsilon.$$

So a new variable needs to be created:

```
gen new_y=lnwage-.09*school;
```

Now estimate the unrestricted model.

```
. reg lnwage age age2 school hispanic black female;

      Source |       SS       df       MS              Number of obs =     704
-------------+------------------------------           F(  6,   697) =   97.67
       Model |  374.244031     6  62.3740052           Prob > F      =  0.0000
    Residual |  445.125611   697  .638630719           R-squared     =  0.4567
-------------+------------------------------           Adj R-squared =  0.4521
       Total |  819.369642   703  1.16553292           Root MSE      =  .79914

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .2909222   .0170466    17.07   0.000     .2574533     .324391
        age2 |   -.003262    .000172   -18.96   0.000    -.0035997   -.0029242
      school |   .0665694   .0102707     6.48   0.000     .0464042    .0867346
    hispanic |  -.1809396   .0802206    -2.26   0.024    -.3384426   -.0234366
       black |  -.2682951   .0965447    -2.78   0.006    -.4578484   -.0787418
      female |  -.1831355   .0633366    -2.89   0.004    -.3074889   -.0587821
       _cons |   3.187207    .427434     7.46   0.000     2.347995     4.02642
------------------------------------------------------------------------------
```

In order to calculate the *F*-statistic and *p*-value in Stata, generate the following values.

```
. gen ssru=_result(4);
. gen numdf=1;
. gen dendf=_result(5);
```

Now estimate the unrestricted model.

```
. reg new_y age age2 hispanic black female;

      Source |       SS       df       MS              Number of obs =     704
-------------+------------------------------           F(  5,   698) =  110.92
       Model |  356.317476     5  71.2634953           Prob > F      =  0.0000
    Residual |  448.449299   698  .642477506           R-squared     =  0.4428
-------------+------------------------------           Adj R-squared =  0.4388
       Total |  804.766776   703  1.14476071           Root MSE      =  .80155
------------------------------------------------------------------------------
```

```
------------------------------------------------------------------------------
       new_y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |    .292338   .0170865    17.11   0.000     .2587908    .3258851
        age2 |  -.0032768   .0001724   -19.00   0.000    -.0036154   -.0029383
    hispanic |  -.1824586   .0804591    -2.27   0.024    -.3404294   -.0244878
       black |  -.2606208   .0967762    -2.69   0.007    -.4506282   -.0706134
      female |  -.1836381   .0635267    -2.89   0.004    -.3083643   -.0589118
       _cons |   2.837554   .4002094     7.09   0.000     2.051796    3.623313
------------------------------------------------------------------------------
```

Finally:

```
. gen ssrr=_result(4);
. gen fstat=((ssrr-ssru)/numdf)/(ssru/dendf);
. gen pval=Ftail(numdf,dendf,fstat);
. list fstat pval if _n==1;

     +---------------------+
     |    fstat       pval |
     |---------------------|
     | 5.204416   .0228304 |
     +---------------------+


. * PROBLEM SIX;
```

First, estimate the unrestricted model, and collect the necessary information.

```
. reg lnwage age age2 school hispanic black female;

      Source |       SS       df       MS              Number of obs =     704
-------------+------------------------------           F(  6,   697) =   97.67
       Model |  374.244031     6  62.3740052           Prob > F      =  0.0000
    Residual |  445.125611   697  .638630719           R-squared     =  0.4567
-------------+------------------------------           Adj R-squared =  0.4521
       Total |  819.369642   703  1.16553292           Root MSE      =  .79914
------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .2909222   .0170466    17.07   0.000     .2574533     .324391
        age2 |   -.003262    .000172   -18.96   0.000    -.0035997   -.0029242
      school |   .0665694   .0102707     6.48   0.000     .0464042    .0867346
    hispanic |  -.1809396   .0802206    -2.26   0.024    -.3384426   -.0234366
       black |  -.2682951   .0965447    -2.78   0.006    -.4578484   -.0787418
      female |  -.1831355   .0633366    -2.89   0.004    -.3074889   -.0587821
       _cons |   3.187207    .427434     7.46   0.000     2.347995     4.02642
------------------------------------------------------------------------------

. replace ssru=_result(4);
. replace numdf=_result(3);
. replace dendf=_result(5);
```

Now estimate the restricted regression, which, for the *F*-test includes no regressors.

```
. reg lnwage;

      Source |       SS       df       MS              Number of obs =     704
-------------+------------------------------           F(  0,   703) =    0.00
       Model |           0     0           .           Prob > F      =       .
    Residual |  819.369642   703  1.16553292           R-squared     =  0.0000
-------------+------------------------------           Adj R-squared =  0.0000
       Total |  819.369642   703  1.16553292           Root MSE      =  1.0796

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   9.773959   .0406889   240.21   0.000     9.694073    9.853845
------------------------------------------------------------------------------

. replace ssrr=_result(4);
. replace fstat=((ssrr-ssru)/numdf)/(ssru/dendf);
. replace pval=Ftail(numdf,dendf,fstat);
. list fstat pval if _n==1;

     +------------------+
     |    fstat    pval |
     |------------------|
     | 97.66834       0 |
     +------------------+
```

Thus, the value of the *F*-stat is 97.67 with a *p*-value of 0, which is what the Stata results from the unrestricted regression also report.

```
. * PROBLEM SEVEN;
```

First, estimate the unrestricted model and collect the results.

```
. reg lnwage age age2 school female;

      Source |       SS       df       MS              Number of obs =     704
-------------+------------------------------           F(  4,   699) =  141.91
       Model |  367.193223     4  91.7983058           Prob > F      =  0.0000
    Residual |  452.176419   699  .646890442           R-squared     =  0.4481
-------------+------------------------------           Adj R-squared =  0.4450
       Total |  819.369642   703  1.16553292           Root MSE      =  .80429

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .2895059   .0171497    16.88   0.000     .2558348    .3231771
        age2 |  -.0032428    .000173   -18.74   0.000    -.0035826   -.0029031
      school |   .0672973   .0103296     6.52   0.000     .0470166     .087578
      female |  -.1832476   .0635309    -2.88   0.004    -.3079818   -.0585134
       _cons |   3.134705   .4298208     7.29   0.000     2.290811      3.9786
------------------------------------------------------------------------------

. replace ssru=_result(4);
. replace numdf=(3-1)*(_result(3)+1);
. replace dendf=_result(1)-3*(_result(3)+1);
```

Now estimate the three restricted regressions, where the restrictions are by race. For each collect the sum of squared residuals. (In fact, the coefficient estimates of the unrestricted models, as always, are completely ignored in the test, so I will not report them here.)

```
. reg lnwage age age2 school female if white==1;

      Source |       SS       df       MS              Number of obs =     499
-------------+------------------------------           F(  4,   494) =  122.02
       Model |  296.719281     4  74.1798202           Prob > F      =  0.0000
    Residual |  300.319325   494  .607933855           R-squared     =  0.4970
-------------+------------------------------           Adj R-squared =  0.4929
       Total |  597.038605   498   1.1988727           Root MSE      =   .7797
------------------------------------------------------------------------------

. gen ssrw=_result(4);

. reg lnwage age age2 school female if hispanic==1;

      Source |       SS       df       MS              Number of obs =     125
-------------+------------------------------           F(  4,   120) =   20.95
       Model |  66.8914979     4  16.7228745           Prob > F      =  0.0000
    Residual |  95.7830093   120  .798191744           R-squared     =  0.4112
-------------+------------------------------           Adj R-squared =  0.3916
       Total |  162.674507   124  1.31189119           Root MSE      =  .89342
------------------------------------------------------------------------------

. gen ssrh=_result(4);

. reg lnwage age age2 school female if black==1;

      Source |       SS       df       MS              Number of obs =      80
-------------+------------------------------           F(  4,    75) =    8.58
       Model |   17.895377     4  4.47384426           Prob > F      =  0.0000
    Residual |  39.1230099    75  .521640131           R-squared     =  0.3139
-------------+------------------------------           Adj R-squared =  0.2773
       Total |  57.0183869    79  .721751733           Root MSE      =  .72225
------------------------------------------------------------------------------

. gen ssrb=_result(4);
```

Now it is straightforward to calculate the test statistic and *p*-value:

```
. replace fstat=((ssru-ssrw-ssrh-ssrb)/numdf)/((ssrw+ssrh+ssrb)/dendf);
. replace pval=Ftail(numdf,dendf,fstat);
. list fstat pval if _n==1;

     +---------------------+
     |    fstat       pval |
     +---------------------+
     | 2.683503   .0031481 |
     +---------------------+
```

Notice that we reject the claim that all of the coefficients are identical across races as the *p*-value is extremely low at 0.31%.

. * PROBLEM EIGHT;

In order to estimate a model that allows for a different slope coefficient on schooling for each race, we need to define three (actually two, but soon it will be shown why you might want to define three) new variables.

```
. gen schoolb=school*black;
. gen schoolh=school*hispanic;
. gen schoolw=school*white;
```

Now estimate the model with school and two of the three new variables. In this case, the coefficient on school is the return to education for whites. The coefficient on school plus the coefficient on schoolb is the return to education for blacks while the coefficient on schoolb is simply the difference between the return to education between blacks and whites. (One interprets the coefficient on schoolh similarly.)

```
. reg lnwage age age2 school schoolb schoolh hispanic black female;

      Source |       SS       df       MS              Number of obs =     704
-------------+------------------------------           F(  8,   695) =   73.05
       Model |  374.256578     8  46.7820722           Prob > F      =  0.0000
    Residual |  445.113064   695  .640450452           R-squared     =  0.4568
-------------+------------------------------           Adj R-squared =  0.4505
       Total |  819.369642   703  1.16553292           Root MSE      =  .80028

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .2908242   .0170982    17.01   0.000     .2572539    .3243945
        age2 |   -.003261   .0001726   -18.89   0.000    -.0035998   -.0029221
      school |    .067356   .0118637     5.68   0.000     .0440631    .0906489
     schoolb |  -.0042654   .0348725    -0.12   0.903    -.0727335    .0642027
     schoolh |  -.0024487   .0290629    -0.08   0.933    -.0595103    .0546128
    hispanic |  -.1474884    .405027    -0.36   0.716    -.9427117    .6477349
       black |  -.2113987    .474454    -0.45   0.656    -1.142934    .7201364
      female |  -.1829657   .0634396    -2.88   0.004    -.3075219   -.0584095
       _cons |   3.178707   .4339842     7.32   0.000      2.32663    4.030784
------------------------------------------------------------------------------
```

Now, to test the claim that there is no difference in the return to schooling for all three races, the test command is:

```
. test schoolh=schoolb=0;
```

The results fail to reject the claim as the *F*-statistic has a *p*-value of 99.03%.

Next, to test the claim that there is no difference in the return to schooling between blacks and Hispanics, the test command is:

```
. test schoolh=schoolb;
```

Again, the results fail to reject the claim as the *F*-stat has a *p*-value of 96.57%.

Alternatively, we could include schoolw in the model and omit school. Under this specification (which is the same in terms of the data), the coefficients on the schooling variables are the returns to schooling for each race. Thus, to test differences across the races, you must include all of the variables.

That is, first, one estimates the model:

```
. reg lnwage age age2 schoolw schoolb schoolh hispanic black female;
```

Then, the test command is:

```
. test schoolw=schoolh=schoolb;
```

And this will provide the exact same *F*-statistic and *p*-value as before. Testing part (B) would require the command:

```
. test schoolh=schoolb;
```

And this too provides the exact same *F*-statistic and *p*-value as before.

```
. * PROBLEM NINE;
```

First, in order to estimate the fully interacted model, we need to create more interaction terms. We already have the school interaction variables, so:

```
. gen agew=age*white;
. gen ageb=age*black;
. gen ageh=age*hispanic;
. gen age2w=age2*white;
. gen age2b=age2*black;
. gen age2h=age2*hispanic;
. gen femalew=female*white;
. gen femaleb=female*black;
. gen femaleh=female*hispanic;
```

Now, estimate the fully interacted model. In this case, we are including all of the variables from before and all of the black and Hispanic interaction terms.

```
. reg lnwage age ageb ageh age2 age2b age2h school schoolb schoolh female femaleb
             femaleh hispanic black ;

      Source |       SS       df       MS              Number of obs =     704
-------------+------------------------------           F( 14,   689) =   43.44
       Model |  384.144298    14  27.4388785           Prob > F      =  0.0000
    Residual |  435.225344   689  .631676841           R-squared     =  0.4688
-------------+------------------------------           Adj R-squared =  0.4580
       Total |  819.369642   703  1.16553292           Root MSE      =  .79478
------------------------------------------------------------------------------
```

lnwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	.3116017	.0201491	15.46	0.000	.2720407	.3511627
ageb	-.1228372	.0558818	-2.20	0.028	-.2325563	-.0131181
ageh	-.0149718	.0457634	-0.33	0.744	-.1048243	.0748807
age2	-.0034561	.0002023	-17.08	0.000	-.0038533	-.0030589
age2b	.0012847	.0005662	2.27	0.024	.0001731	.0023963
age2h	.0000829	.000467	0.18	0.859	-.0008339	.0009998
school	.0673256	.0117837	5.71	0.000	.0441894	.0904618
schoolb	-.0215787	.035305	-0.61	0.541	-.0908971	.0477397
schoolh	-.0058852	.0289235	-0.20	0.839	-.0626741	.0509037
female	-.3053795	.0745975	-4.09	0.000	-.4518452	-.1589139
femaleb	.4100695	.2026353	2.02	0.043	.0122128	.8079261
femaleh	.3601824	.1759047	2.05	0.041	.0148089	.705556
hispanic	.2992804	1.131305	0.26	0.791	-1.921938	2.520499
black	2.568591	1.476077	1.74	0.082	-.329558	5.46674
_cons	2.710002	.5014309	5.40	0.000	1.725487	3.694518

Now, because of this model specification, to test if the return for blacks is different from the return to whites, the test is simply:

```
. test schoolb;
```

And we fail to reject the claim that the return is the same for whites and blacks as the *F*-statistic is associated with a *p*-value of 54.13%.

To test if the return for Hispanics is different from the return to whites:

```
. test schoolh;
```

And we fail to reject the claim that the return is the same for Hispanics and whites as the *F*-statistic is associated with a *p*-value of 83.88%.

To test if the return for Hispanics is different from the return to blacks:

```
. test schoolb=schoolh;
```

And we fail to reject the claim that the return is the same for blacks and Hispanics as the *F*-statistic is associated with a *p*-value of 71.20%.

Finally, to test if the return is the same for all three races:

```
. test schoolb=schoolh=0;
```

And we fail to reject the claim that the return is the same for all three races as the *F*-statistic is associated with a *p*-value of 82.36%.

Alternatively, one could include all of the interaction terms so that each coefficient is the expected return for each race. That is:

```
. reg lnwage agew ageb ageh age2w age2b age2h schoolw schoolb schoolh femalew femaleb
             femaleh hispanic black ;

      Source |       SS       df       MS              Number of obs =     704
-------------+------------------------------           F( 14,   689) =   43.44
       Model |  384.144298    14  27.4388785           Prob > F      =  0.0000
    Residual |  435.225344   689  .631676841           R-squared     =  0.4688
-------------+------------------------------           Adj R-squared =  0.4580
       Total |  819.369642   703  1.16553292           Root MSE      =  .79478

------------------------------------------------------------------------------
      lnwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        agew |   .3116017   .0201491    15.46   0.000     .2720407    .3511627
        ageb |   .1887645   .0521229     3.62   0.000     .0864258    .2911032
        ageh |   .2966299    .041089     7.22   0.000     .2159552    .3773046
       age2w |  -.0034561   .0002023   -17.08   0.000    -.0038533   -.0030589
       age2b |  -.0021714   .0005288    -4.11   0.000    -.0032096   -.0011332
       age2h |  -.0033732   .0004209    -8.01   0.000    -.0041996   -.0025468
     schoolw |   .0673256   .0117837     5.71   0.000     .0441894    .0904618
     schoolb |   .0457469   .0332805     1.37   0.170    -.0195964    .1110903
     schoolh |   .0614404   .0264143     2.33   0.020     .0095782    .1133026
     femalew |  -.3053795   .0745975    -4.09   0.000    -.4518452   -.1589139
     femaleb |   .1046899   .1884045     0.56   0.579    -.2652259    .4746058
     femaleh |   .0548029   .1593037     0.34   0.731    -.2579761    .3675819
    hispanic |   .2992804   1.131305     0.26   0.791    -1.921938    2.520499
       black |   2.568591   1.476077     1.74   0.082     -.329558     5.46674
       _cons |   2.710002   .5014309     5.40   0.000     1.725487    3.694518
------------------------------------------------------------------------------
```

At this point, every test from above can be carried out, and one will obtain the exact same F-statistics and p-values by entering:

```
. test schoolw=schoolb;
. test schoolw=schoolh;
. test schoolb=schoolh;
. test schoolw=schoolb=schoolh;
```

```
#delimit;
set more 1;
log using project4.log, replace;

use project4.dta;

* PROBLEM ONE;
gen hispanic=(race==1);
gen black=(race==2);
gen white=(race==3);
gen female=(sex=="F");
gen male=(sex=="M");
gen age2=age*age;
gen lnwage=ln(wage);
keep  wage lnwage age age2 school hispanic black white female male;
order wage lnwage age age2 school hispanic black white female male;
sum;

* PROBLEM TWO;
reg wage age age2 school hispanic black female;
predict errors1, resid;
sum errors1;
scatter errors1 age, ti("Wage Regression") saving(e1_age, replace);

* PROBLEM THREE;
reg lnwage age age2 school hispanic black female;
predict errors2, resid;
sum errors2;
scatter errors2 age, ti("Ln(Wage) Regression") saving(e2_age, replace);

* PROBLEM FOUR;
reg lnwage age age2 school hispanic black female;

      * QUESTION A;
        test hispanic=black;

      * QUESTION B;
        test hispanic=black=0;

      * QUESTION C;
        test school=.09;

      * QUESTION D;
        test female=-0.10;

      * QUESTION E;
      * nothing for stata to do;
```

```
* PROBLEM FIVE;
gen new_y=lnwage-.09*school;
reg lnwage age age2 school hispanic black female;
gen ssru=_result(4);
gen numdf=1;
gen dendf=_result(5);
reg new_y age age2 hispanic black female;
gen ssrr=_result(4);
gen fstat=((ssrr-ssru)/numdf)/(ssru/dendf);
gen pval=Ftail(numdf,dendf,fstat);
list fstat pval if _n==1;

* PROBLEM SIX;
reg lnwage age age2 school hispanic black female;
replace ssru=_result(4);
replace numdf=_result(3);
replace dendf=_result(5);
reg lnwage;
replace ssrr=_result(4);
replace fstat=((ssrr-ssru)/numdf)/(ssru/dendf);
replace pval=Ftail(numdf,dendf,fstat);
list fstat pval if _n==1;

* PROBLEM SEVEN;
reg lnwage age age2 school female;
replace ssru=_result(4);
replace numdf=(3-1)*(_result(3)+1);
replace dendf=_result(1)-3*(_result(3)+1);
reg lnwage age age2 school female if white==1;
gen ssrw=_result(4);
reg lnwage age age2 school female if hispanic==1;
gen ssrh=_result(4);
reg lnwage age age2 school female if black==1;
gen ssrb=_result(4);
replace fstat=(
(ssru-ssrw-ssrh-ssrb)/numdf)/((ssrw+ssrh+ssrb)/dendf);
replace pval=Ftail(numdf,dendf,fstat);
list fstat pval if _n==1;

* PROBLEM EIGHT;
gen schoolb=school*black;
gen schoolh=school*hispanic;
gen schoolw=school*white;
reg lnwage age age2 school schoolb schoolh hispanic black female;
test schoolh=schoolb=0;
test schoolh=schoolb;
      * Alternatively;
        reg lnwage age age2 schoolw schoolb schoolh hispanic black female;
        test schoolw=schoolh=schoolb;
        test schoolh=schoolb;
```

```
* PROBLEM NINE;
gen agew=age*white;
gen ageb=age*black;
gen ageh=age*hispanic;
gen age2w=age2*white;
gen age2b=age2*black;
gen age2h=age2*hispanic;
gen femalew=female*white;
gen femaleb=female*black;
gen femaleh=female*hispanic;
reg lnwage age ageb ageh age2 age2b age2h school schoolb schoolh
           female femaleb femaleh hispanic black ;
test schoolb;
test schoolh;
test schoolb=schoolh;
test schoolb=schoolh=0;
      * Alternatively;
        reg lnwage agew ageb ageh age2w age2b age2h schoolw schoolb schoolh
                   femalew femaleb femaleh hispanic black ;
        test schoolw=schoolb;
        test schoolw=schoolh;
        test schoolb=schoolh;
        test schoolw=schoolb=schoolh;

save coeftests2, replace;
clear;

log close;
```