Homework #5 Answers

1a.) What is the R_2 for this regression? Using an F-test, conclude at the 95% level whether the variables of the model (Educ, Exper, and Tenure) have any power in explaining wage

```
. regress wage educ exper tenure

      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  3,   931) =   53.00
       Model |  22278193.8     3  7426064.59           Prob > F      =  0.0000
    Residual |   130437974   931  140105.236           R-squared     =  0.1459
-------------+------------------------------           Adj R-squared =  0.1431
       Total |   152716168   934  163507.675           Root MSE      =  374.31

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   74.41486   6.286993    11.84   0.000     62.07654    86.75318
       exper |   14.89164    3.25292     4.58   0.000     8.507732    21.27554
      tenure |   8.256811   2.497628     3.31   0.001     3.355178    13.15844
       _cons |  -276.2405   106.7018    -2.59   0.010    -485.6444   -66.83653
```

The R^2 is 0.1459. This means the that model matches 14.5% of the variation in the dependent variable, wage.

To test whether this regression tells us anything, we use the "full exclusion" F test. For this the null and alternative hypotheses are the following:

H0: $B_{Educ}=B_{Exper}=B_{Tenure}=0$
HA: H0 Not true.

To test this, construct the F Statistic, which is 53. For a F-Test with 3 restrictions and 931 degrees of freedom, the critical value is 2.6. Clearly, we can reject the null hypothesis in favor of the null. In words, we lose a lot by the restrictions in H0, and thus we reject them in favor of the model without them.

1b.) Does education significantly affect wages? That is, can you conclude that β_{Educ} is significantly different from zero? Test this hypothesis at the 95% level.

Looking at the table, we see that the t-statistic is 11.84. For a two sided test, the null and alternative hypotheses are the following:

H0: $\beta Educ=0$
HA: $\beta Educ$ not equal to 0

We reject the null in favor of the alternative if:

|tstat|>tcrit

The tcrit for a two-sided test at 95% is 1.96. Clearly, we can reject the null in favor of the alternative. That is, education has a statistically significant effect of wages.

1c.) Please construct a 99% confidence interval for β_{Exper} . Please interpret your results.

On an exam, I’m unlikely to just give you the confidence interval. You will have to calculate it using the standard error and the critical value. And, the table above only gives you the confidence interval for 95%. For this homework, you can either (1) calculate it directly or (2) run another regression specifying that the level of confidence is 99%. Doing so yields:

```
. regress wage educ exper tenure, level(99)

      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  3,   931) =   53.00
       Model |  22278193.8     3  7426064.59           Prob > F      =  0.0000
    Residual |   130437974   931  140105.236           R-squared     =  0.1459
-------------+------------------------------           Adj R-squared =  0.1431
       Total |   152716168   934  163507.675           Root MSE      =  374.31

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [99% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   74.41486   6.286993    11.84   0.000     58.18738    90.64235
       exper |   14.89164    3.25292     4.58   0.000     6.495459    23.28782
      tenure |   8.256811   2.497628     3.31   0.001     1.810133    14.70349
       _cons |  -276.2405   106.7018    -2.59   0.010    -551.6507   -.8302472
------------------------------------------------------------------------------
```

Interpreting, the marginal effect of one additional year of education is between $58 and $90 per month.

1d.) Suppose that I reject the hypothesis that $\beta_{Tenure} = 0$ in favor of a two-sided alternative. What is the probability that I’m wrong?

I am asking for the p-value. Looking at the regression results, which was the intention of this question, we see that the p-value is 0.001, which means that the probability of wrongly rejecting the hypothesis that tenure makes a difference is 0.1%.

The alternate way is to use the t-statistic (which is 3.31) and calculate the p-value using the z-table. However, for this question, the z-table does not go up to 3.31. For the exam, I will surely give you a question for which you must calculate the p-value using the z-table, so make sure you know how to do it.

```
. gen educ2=educ^2
. regress wage educ educ2 exper tenure

      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  4,   930) =   40.19
       Model |  22506916.3     4  5626729.06           Prob > F      =  0.0000
    Residual |   130209252   930  140009.948           R-squared     =  0.1474
-------------+------------------------------           Adj R-squared =  0.1437
       Total |   152716168   934  163507.675           Root MSE      =  374.18
------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   169.1691   74.40102     2.27   0.023     23.15579    315.1825
       educ2 |  -3.347401   2.618983    -1.28   0.202    -8.487203    1.792401
       exper |   15.50532   3.287069     4.72   0.000     9.054384    21.95625
      tenure |   8.293652   2.496945     3.32   0.001     3.393352    13.19395
       _cons |   -936.446   527.4386    -1.78   0.076    -1971.554    98.66164
------------------------------------------------------------------------------
```

1e.) **Do not answer this.**

1f.) First, generate the squared terms in education. Then run the regression.

```
. gen educ2=educ^2
. regress lwage educ educ2 exper tenure

      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  4,   930) =   43.85
       Model |  26.2869712     4  6.57174281           Prob > F      =  0.0000
    Residual |  139.369323   930  .149859487           R-squared     =  0.1587
-------------+------------------------------           Adj R-squared =  0.1551
       Total |  165.656294   934  .177362199           Root MSE      =  .38712

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .2272601   .0769736     2.95   0.003     .0761981    .3783221
       educ2 |  -.0053837   .0027095    -1.99   0.047    -.0107013   -.0000662
       exper |   .0163155   .0034007     4.80   0.000     .0096415    .0229895
      tenure |   .0134341   .0025833     5.20   0.000     .0083643    .0185038
       _cons |   4.434866   .5456756     8.13   0.000     3.363968    5.505764
------------------------------------------------------------------------------
```

Diminishing returns is the case where the returns to education are positive to start, but diminishing as education increases. This is the case if the coefficient on educ2 is significantly negative. This is clearly the case in the above table. To see this:

H0: βEduc2=0
HA: βEduc2 not equal to 0

The Tstat is -1.99 which is less than -1.96. Thus, we can reject the null that the coefficient on education squared is negative.

1g.) A similar approach to above.

```
. gen exper2=exper^2
. regress lwage educ exper exper2 tenure

      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  4,   930) =   42.70
       Model |  25.7026698     4  6.42566746           Prob > F      =  0.0000
    Residual |  139.953624   930  .150487768           R-squared     =  0.1552
-------------+------------------------------           Adj R-squared =  0.1515
       Total |  165.656294   934  .177362199           Root MSE      =  .38793
------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0746848   .0065659    11.37   0.000      .061799    .0875706
       exper |   .0181845   .0133623     1.36   0.174    -.0080393    .0444082
      exper2 |  -.0001241   .0005619    -0.22   0.825    -.0012268    .0009786
      tenure |   .0134295   .0026003     5.16   0.000     .0083263    .0185326
       _cons |   5.484652   .1232959    44.48   0.000     5.242682    5.726623
```

However, we see that the coefficient on exper2 is negative, but not significantly different from zero. That is, the tstat is -0.22, which is way too small. Thus, there are not significant diminishing returns to experience.

1h.)

Unrestricted model

```
. regress lwage educ educ2 exper exper2 tenure

      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  5,   929) =   35.04
       Model |  26.2869744     5  5.25739488           Prob > F      =  0.0000
    Residual |   139.36932   929  .150020796           R-squared     =  0.1587
-------------+------------------------------           Adj R-squared =  0.1542
       Total |  165.656294   934  .177362199           Root MSE      =  .38733

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .2272176   .0775668     2.93   0.003     .0749911    .3794441
       educ2 |  -.0053824   .0027273    -1.97   0.049    -.0107347     -.00003
       exper |   .0163749    .013373     1.22   0.221    -.0098699    .0426197
      exper2 |  -2.59e-06   .0005644    -0.00   0.996    -.0011102     .001105
      tenure |   .0134352   .0025963     5.17   0.000     .0083399    .0185304
       _cons |   4.434884   .5459833     8.12   0.000     3.363381    5.506388
```

Restricted Model

```
. regress lwage educ exper tenure

      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  3,   931) =   56.97
       Model |  25.6953278     3  8.56510927           Prob > F      =  0.0000
    Residual |  139.960966   931  .150334013           R-squared     =  0.1551
-------------+------------------------------           Adj R-squared =  0.1524
       Total |  165.656294   934  .177362199           Root MSE      =  .38773

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0748638   .0065124    11.50   0.000      .062083    .0876446
       exper |   .0153285   .0033696     4.55   0.000     .0087156    .0219413
      tenure |   .0133748   .0025872     5.17   0.000     .0082974    .0184522
       _cons |   5.496696   .1105282    49.73   0.000     5.279782    5.713609
```

H0: βEduc2=βExper2=0
HA: H0 not equal to 0

SSR_R=139.96

SSR_U=139.37

Q=2

DOF_UR=929

Fstat=(139.96-139.37)/2/(139.37/929)= **1.966**

Fcrit=3.00

Thus, we cannot reject the null hypothesis. We can safely remove the squared terms from the regression.

Problem 2a

H0: Theta=βExper-βTenure=0
HA: Theta not equal to 0

First, write the equation

$$wage = \beta_0 + \beta_{educ} educ + \beta_{exper} exper + \beta_{tenure} tenure + u$$

Then, substitute Theta=βExper-βTenure and simplify.

$$wage = \beta_0 + \beta_{educ} educ + (\theta + \beta_{tenure}) exper + \beta_{tenure} tenure + u$$
$$wage = \beta_0 + \beta_{educ} educ + \theta exper + \beta_{tenure} (exper + tenure) + u$$

Thus, run a regression with exper+tenure rather than tenure.

```
. regress wage educ exper tenure

      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  3,   931) =   53.00
       Model |  22278193.8     3  7426064.59           Prob > F      =  0.0000
    Residual |   130437974   931  140105.236           R-squared     =  0.1459
-------------+------------------------------           Adj R-squared =  0.1431
       Total |   152716168   934  163507.675           Root MSE      =  374.31

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   74.41486   6.286993    11.84   0.000     62.07654    86.75318
       exper |   14.89164    3.25292     4.58   0.000     8.507732    21.27554
      tenure |   8.256811   2.497628     3.31   0.001     3.355178    13.15844
       _cons |  -276.2405   106.7018    -2.59   0.010    -485.6444   -66.83653
```

Looking at the coefficient on exper, we see that it is insignificantly different from zero. Thus, there we cannot reject the null hypothesis that βExper=βTenure.

Problem 2b

$$wage = \beta_0 + \beta_{educ} educ + \beta_{exper} exper + \beta_{tenure} tenure + u$$

First, write the prediction for a person with a predicted wage of a person with two years of education, 3 years of experience and 0 years of tenure

$$\theta = \beta_0 + \beta_{educ} 2 + \beta_{exper} 3$$

Substitute in for β_0:

$$wage = \theta - \beta_{educ} 2 - \beta_{\exp er} 3 + \beta_{educ} educ + \beta_{\exp er} \exp er + \beta_{tenure} tenure + u$$
$$wage = \theta + \beta_{educ} (educ - 2) + \beta_{\exp er} (\exp er - 3) + \beta_{tenure} tenure + u$$

Thus, generate new variables and run an adjusted regression.

```
. gen educ_2=educ-2
. gen exper_3=exper-3

. regress wage educ_2 exper_3 tenure

      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  3,   931) =   53.00
       Model |  22278193.8     3  7426064.59           Prob > F      =  0.0000
    Residual |   130437974   931  140105.236           R-squared     =  0.1459
-------------+------------------------------           Adj R-squared =  0.1431
       Total |   152716168   934  163507.675           Root MSE      =  374.31

------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      educ_2 |   74.41486   6.286993    11.84   0.000     62.07654    86.75318
     exper_3 |   14.89164    3.25292     4.58   0.000     8.507732    21.27554
      tenure |   8.256811   2.497628     3.31   0.001     3.355178    13.15844
       _cons |  -82.73581    88.4294    -0.94   0.350    -256.2799    90.80825
------------------------------------------------------------------------------
```