Version: October 5, 2020 Fall 2020
EC200 Econometrics and Applications

In-Class Exercise - Multiple Linear Regression

Consider a dataset on earnings in the United States. We are interested in the returns to education - how much an extra year of schooling "buys" you in terms of weekly wages (...as of 1980). You're also worried about whether one's education suffers from omitted variable bias.

1. You estimate two equations:

$$\hat{education} = 146.95 + 60.21 educ$$
$$\hat{educ} = 5.84 + 0.075 IQ$$

   Based on these results, is 60.21 an overestimate or underestimate of the returns to education? How do you know?

2. You estimate another equation:

$$\hat{education} = -128.89 + 42.06 educ + 5.14 IQ$$

   What is the interpretation of the coefficient on *educ*? What is the interpretation of the constant?

3. Now, you control experience and age and estimate the following population regression model:

$$wage_i = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 age + \beta_4 age^2 + u$$

```
. reg wage educ IQ, robust
Linear regression                               Number of obs     =        935
                                                F(2, 932)         =      64.47
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1339
                                                Root MSE          =     376.73
```

wage	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
educ	42.05762	6.810074	6.18	0.000	28.69276	55.42247
IQ	5.137958	.9266458	5.54	0.000	3.319404	6.956512
_cons	-128.8899	93.09396	-1.38	0.167	-311.5879	53.80818

   A one-year increase in age is associated with what change in wages? (mind the squared term)

4. Finally, because you are worried about omitted variable bias, you include father's and mother's education.

   (a) Why might parent's education might directly affect wage?

   (b) Which other independent variables do you think parent's education might affect? Explain.

```
. reg wage educ exper age age_sq feduc meduc, robust
Linear regression                               Number of obs     =        722
                                                F(6, 715)         =      20.81
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1620
                                                Root MSE          =     375.15
------------------------------------------------------------------------------
             |               Robust
        wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   56.30609   8.000723     7.04   0.000     40.59837    72.01381
       exper |   14.26107   4.866297     2.93   0.003     4.707136    23.81501
         age |   28.32564   115.6423     0.24   0.807    -198.7134    255.3647
      age_sq |  -.2297112   1.741619    -0.13   0.895     -3.64901    3.189588
       feduc |   14.10016   5.116401     2.76   0.006     4.055198    24.14513
       meduc |   9.678735   5.345395     1.81   0.071    -.8158121    20.17328
       _cons |  -884.1403   1887.477    -0.47   0.640      -4589.8    2821.519
------------------------------------------------------------------------------
```

   (c) How did controlling for parent's education affect the returns to education? The returns to IQ?