Version: October 5, 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:

$$\widehat{wage} = 146.95 + 60.21 educ$$
$$\widehat{educ} = 5.84 + 0.075 IQ$$

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: $\widehat{wage} = -128.89 + 42.06 educ + 5.14 IQ$ What is the interpretation of the coefficient on educ? What is the interpretation of the constant?
- 3. Now, you control for experience and age and estimate the following population regression model:

$$wage_i = \beta_0 + \beta_1 educ_i + \beta_2 IQ_i + \beta_3 exper_i + \beta_4 age_i + \beta_5 age_i^2 + u_i$$

. reg wage educ IQ exper age age_sq, robust Linear regression

Linear regress	Number of	fobs =	935			
				F(5, 929)) =	35.68
				Prob > F	=	0.0000
				R-squared	1 =	0.1704
				Root MSE	=	369.29
		Robust				
wage	Coef.	Std. Err.	t	P> t	[95% Conf.	Interval]
educ	51.65693	7.507317	6.88	0.000	36.92366	66.3902
IQ	5.280274	.8996509	5.87	0.000	3.51469	7.045858
exper	11.14023	3.871415	2.88	0.004	3.542496	18.73796
age	63.78875	98.27005	0.65	0.516	-129.0683	256.6458
age_sq	745563	1.478397	-0.50	0.614	-3.646948	2.155822
_cons	-1688.496	1615.041	-1.05	0.296	-4858.048	1481.057

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

. reg wage educ IQ exper age age_sq feduc meduc, robust										
Linear regress	sion			Number of	obs =	722				
•				F(7, 714)	=	21.49				
				Prob > F	=	0.0000				
				R-squared	=	0.1818				
				Root MSE	=	370.94				
		Robust								
wage	Coef.	Std. Err.	t	P> t	[95% Conf.	Interval]				
educ	40.62398	8.610596	4.72	0.000	23.71886	57.52909				
IQ	4.730038	1.020807	4.63	0.000	2.725896	6.734179				
exper	13.12448	4.84228	2.71	0.007	3.617671	22.63129				
age	76.02757	116.5825	0.65	0.515	-152.858	304.9131				
age_sq	9288294	1.756502	-0.53	0.597	-4.377356	2.519697				
feduc	12.41051	5.080028	2.44	0.015	2.436935	22.38409				
meduc	6.947846	5.25549	1.32	0.187	-3.370215	17.26591				
_cons	-1899.173	1910.946	-0.99	0.321	-5650.918	1852.572				

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