Econometrics II - Assignment 2

Floris Holstege, Stanislav Avdeev

1

As a base model we use the following model specification:

$$Log(Earnings) = \alpha_0 + \alpha_1 Schooling_{it} + \alpha_2 AGE_{it} + \alpha_3 AGE_{it}^2 + \alpha_4 ETHNICITY_i + \alpha_5 URBAN_{it} + \alpha_6 REGNE_{it} + \alpha_7 REGNCit + \alpha_8 REGW_{it}$$

In order to check the impact of ability, we include a variable $ASVABC$ in the base model. Including an ability allows to account for an omitted variavble bias that most likely occurs in the base model without it.

The results of the base model with and without ability variable are as follows. When we do not account for ability in the model specification, returns to one year of education are higher, i.e. returns to a year of education are 7% and statistically significant at the 1% level. When we inlude the ability variable, returns to one year of education become 4.8% and remain significant at the 1% level. The underlying reason for such a drop is that higher ability students tend to get more education, thus, they tend to get higher earnings.

Table 1: OLS pooled model with and without ability variable

	Dependent variable:	
	EARNINGS	
	(1)	(2)
Schooling	0.070^{***}	0.048^{***}
	(-0.00004)	(-0.00004)
Test_score		0.011^{***}
		(-0.00004)
Ethnicity	-0.192^{***}	-0.096^{***}
	(-0.00004)	(-0.0002)
Constant	-0.079^{***}	-0.386^{***}
	(0.004)	(0.004)
Observations	40,043	40,043
Adjusted R^2	0.292	0.313
Note:	*p<0.1; **p<0.05; ***p<0.01	

From now on we are going to include the ability variable in all model specifications due to two reasons. First, we deal with an omitted variable bias problem (theoretical problem). Second, including this variable improves our model: adjusted R squared is slightlty higher. Moreover, for the sake of saving the space we don't report all coefficients in tables unless they are important for a specific question.

2

As we can see, there is a statistically significant difference between returns to education by ethnicity. Black workers tend to get higher returns.

Table 2: OLS pooled model with heterogeneous effects by ethnicity

	Dependent variable:		
	EARNINGS		
	(1)	(2)	(3)
Schooling	0.046^{***}	0.046^{***}	0.061^{***}
	(-0.00001)	(-0.00000)	(-0.0004)
Ethnicity	-0.295^{***}		
	(-0.001)		
Test_score	0.011^{***}	0.010^{***}	0.014^{***}
	(-0.00002)	(-0.00002)	(-0.0001)
Schooling:Ethnicity	0.016^{***}		
	(0.00004)		
Constant	-0.370^{***}	-0.437^{***}	0.038
	(0.004)	(0.005)	(0.036)
Observations	40,043	35,223	4,820
Adjusted R^2	0.313	0.299	0.314
Note:			*p<0.1; **p<0.05; ***p<0.01

3

Once we use a panel model with random effects, there seems to be no difference returns to education by ethnicity.

4

5

6

To decide between fixed or random effects we run a Hausman test. The p-value is highly significant, thus, it's preferred to use fixed effects.

7

8

9