Assignment 1

Students

January 6, 2023

1 Question 1

1.1 i)

Table 1: OLS regression for log-earnings on schooling, age, and age squared.

	Dependent variable:
	logwage
schooling	0.216^{***}
	(0.032)
age	-0.342
	(0.521)
I(age^2)	-0.011
	(0.008)
Constant	26.409^{***}
	(8.057)
Observations	416
R^2	0.815
Adjusted R^2	0.813
Residual Std. Error	1.499 (df = 412)
F Statistic	604.261^{***} (df = 3; 412)
Note:	*p<0.1; **p<0.05; ***p<0.01

From Table 1 can be observed that only the intercept and `schooling` are significant. Both are significant at the 1%-significance level. For a given worker, an additional year of schooling is associated with a $(e^{0.216} - 1) \cdot 100 \approx 24.11\%$ increase in wage. The intercept and explanatory variables explain 81.5% of the variation in `logwage`.

1.2 ii)

`logwage` is only observed when a worker earns a wage.

Let Z_i', γ, and V_i denote some exclusion restriction, the corresponding coefficient, and error term for individual i, respectively. Then, we have the selection equation

$$I_i^* = Z_i^{'}\gamma + V_i,$$

where I_{i} takes value 1 if the wage for individual i $I_i^* > 0$, and value 0 otherwise. In addition, let Y_i^* denote the latent variable `logwage`, $X_i^{'}$ the regressor(s), β the coefficient (vector) and U_i the error term. The second regression equation is

$$Y_i^* = X_i^{'}\beta + U_i.$$

However, instead of the true `logwage` Y_i^* we observe

1.3 iii)

1.4 iv)

1.5 v)

2 Question 2

2.1 i)

2.2 ii)

2.3 iii)