PS12 Garcia

Jorge Vladimir Garcia

April 2018

6. Use stargazer to produce a summary table of this data frame. Include it in your Latex writeup, and discuss whether the results make sense. At what rate are the log wages missing? Do you think the *log wage* variable is most likely to be MCAR, MAR, or MNAR.

Table 1:

Statistic	N	Mean	St. Dev.	Min	Max
log wage	1,545	1.652	0.688	−0.956	4.166
hgc	2,229	12.455	2.444	5	18
exper	2,229	6.435	4.867	0.000	25.000
kids	2,229	0.429	0.495	0	1

Mean of kids does not make any sense as having 0.429 is not feasible. Mean of years of schooling *hgc*, on the other hand, could be affected by having a greater number of observations with a lot of education, so using the median would be better. Same with with experience variable.

Also, the *log wage* is more likely to be MAR because it could be the case that people with no college education are more likely to not report their wages.

7. Use stargazer to create one regression table which has the estimates of the first three regression models. Include this table in your.tex writeup

The parameter β_1 as it can be seen in the table below, is statistically significant at the same level across all models. The standard error is smallest for the model with mean imputations, because of the means, and the highest for the Heckman model. The effect of hgc is smaller in the model with imputations as well.

8. Assess the impat of a counterfactual polify in which wives and mothers are no longer allowed to work in union jobs. Do you think the model you estimated is realistic? Why or why not?

They are exactly the same. I don't know if this is due to some errors with my coding, because I am not sure this is what I was supposed to get.

Table 2: Results

	Dependent variable:		
	log wage		
	OLS		*selection*
	(1)	(2)	(3)
hgc	0.059***	0.036***	0.091***
	(0.009)	(0.006)	(0.010)
union1	0.222**	0.068	
	(0.087)	(0.047)	
college1	−0.065	−0.126***	
	(0.106)	(0.048)	
union			0.186**
			(0.084)
college			0.092
			(0.100)
exper	0.050***	0.021***	0.054***
	(0.013)	(0.007)	(0.012)
exper_sqrd	−0.004***	−0.001***	
	(0.001)	(0.0004)	
I(exper_Sqr)			−0.002*
			(0.001)
Constant	0.834***	1.149***	0.446***
	(0.113)	(0.078)	(0.122)
Observations	1,545	2,229	2,229
R^2	0.038	0.020	
Adjusted R^2	0.035	0.018	
ρ			-0.998
Inverse Mills Ratio			-0.695***(0.060)
Residual Std. Error	0.676 (df = 1539)	0.568 (df = 2223)	
F Statistic	12.106*** (df = 5; 1539)	9.207*** (df = 5; 2223)	

Note: *p<0.1; **p<0.05; ***p<0.01