PS 12

Audrey Hopewell

April 30, 2020

1 Question 6

Table 1:

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
logwage	1,545	1.652	0.688	−0.956	1.201	2.120	4.166
hgc	2,229	12.455	2.444	5	11	14	18
exper	2,229	6.435	4.867	0.000	2.452	9.778	25.000
kids	2,229	0.429	0.495	0	0	1	1

over 30% of logwages are missing. They are most likely missing not at random because women with lower wages (related to fewer years of school or experience) are probably less likely to report their earnings.

2 Question 7

Table 2: Results

		Dependent variable:	
		logwage	
	OLS		selection
	(1)	(2)	(3)
hgc	0.059***	0.036***	0.091***
	(0.009)	(0.006)	(0.010)
union1	0.222**	0.068	0.186**
	(0.087)	(0.047)	(0.084)
college1	−0.065	−0.126***	0.092
	(0.106)	(0.048)	(0.100)
exper	0.050***	0.021***	0.054***
	(0.013)	(0.007)	(0.012)
I(exper^2)	−0.004***	−0.001***	−0.002*
	(0.001)	(0.0004)	(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

Model 3 (using Heckman selection) is the only one to correctly estimate β_1.
Model 1 is more accurate than Model 2, probably because mean imputation
undermines the correlation between variables in the data.

3 Question 9

The counterfactual policy would result in a 1% decrease in preference for union jobs, according to the model. This is probably not a very realistic model at all because we cannot account for all the factors involved in choosing a union vs. non-union job.