

# PS 12

y.mozaffar

April 2020

## 1 Problem 6

[table on next page as Table 1] There are 684 observations, about 30 percent, that are missing the logwage variable. I would guess that this would be MNAR missingness. Assuming this is self-reported data, it's likely that those with higher wages would be more likely to report it.

## 2 Problem 7

[table on next page as Table 2] Both mean and stepwise imputation were far off the mark, implying that the missingness was in fact not random. This is further corroborated by the fact that the Heckit model, accounting for the non-randomness, was within 0.001 of the true value of Beta 1.

## 3 Problem 12

[table on next page as Table 3] The counterfactual is just barely lower than the original probit model. This doesn't quite make sense to me because If two large subsets of the dataset were now unable to join a union, you would think the probability of a given person being in a union would be much smaller

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

Table 2:

	<i>Dependent variable:</i>		
	logwage		
	<i>OLS</i>		<i>selection</i>
	(1)	(2)	(3)
hgc	0.058*** (0.009)	0.058*** (0.009)	0.092*** (0.009)
union1	0.068 (0.073)	0.068 (0.073)	0.171** (0.074)
college1	-0.079 (0.106)	-0.079 (0.106)	0.108 (0.094)
exper	0.016** (0.006)	0.016** (0.006)	0.051*** (0.009)
I(exper^2)			-0.002** (0.001)
Constant	0.891*** (0.112)	0.891*** (0.112)	0.453*** (0.112)
Observations	1,545	1,545	2,229
R <sup>2</sup>	0.032	0.032	
Adjusted R <sup>2</sup>	0.029	0.029	
Log Likelihood			-1,802.229
$\rho$			-0.998*** (0.002)
Residual Std. Error (df = 1540)	0.678	0.678	
F Statistic (df = 4; 1540)	12.600***	12.600***	
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01	

Table 3:

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
predProbit	2,229	0.237	0.407	0	0	0.3	1
counterfact	2,229	0.228	0.401	0	0	0.2	1