ECON 5253 Problem Set 7

Ahmed Chaudhry

March 25, 2021

1 Question No. 6

About 25% of the logwage values are missing. I think the logwage variable is most likely to be Missing Not At Random (MNAR). This can be because of two reasons. The first reason could be that some people have not reported their wages at all, and the second reason could be wages are missing for the people who are unemployed.

Table 1 Summary Table: Before Dropping Missing Observations

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
logwage	675	25	1.6	0.4	0.0	1.7	2.3
hgc	17	0	13.1	2.5	0.0	12.0	18.0
tenure	260	1	6.0	5.5	0.0	3.8	25.9
age	13	0	39.2	3.1	34.0	39.0	46.0

Table 2 Summary Table: After Dropping Missing Observations

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
logwage	675	25	1.6	0.4	0.0	1.7	2.3
hgc	17	0	13.1	2.5	0.0	12.0	18.0
tenure	260	1	6.0	5.5	0.0	3.8	25.9
age	13	0	39.2	3.1	34.0	39.0	46.0

2 Question No. 7

In table 3, it can be seen that no model is actually come close to the true $\hat{\beta}_1 = 0.093$. The model 1 which uses listwise deletion and the model 4 which employs regression imputations (bootstrapped) gives us the closest estimates to the true value. While the model 2, which uses mean imputations and model 3 which makes use of regression based predicted imputations are further from the truth - however, model 3 does a better job as compared to model 2.

In summary, it can be concluded that regression based predicted imputation methods have generally done a better job in estimating our model.

Table 3 OLS Regressions: Impact of Years of Schooling on Wages

	Model 1	Model 2	Model 3	Model 4
(Intercept)	0.534	0.708	0.563	0.532
	(0.146)	(0.116)	(0.112)	(0.146)
hgc	0.062	0.050	0.059	0.062
	(0.005)	(0.004)	(0.004)	(0.005)
as.factor(college)not college grad	0.145	0.169	0.178	0.140
	(0.034)	(0.026)	(0.025)	(0.035)
poly(tenure, 2, raw = T)1	0.050	0.038	0.047	0.050
	(0.005)	(0.004)	(0.004)	(0.005)
poly(tenure, 2, raw = T)2	-0.002	-0.001	-0.002	-0.002
	(0.000)	(0.000)	(0.000)	(0.000)
age	0.000	0.000	0.000	0.001
	(0.003)	(0.002)	(0.002)	(0.003)
as.factor(married)single	-0.022	-0.027	-0.028	-0.024
	(0.018)	(0.014)	(0.013)	(0.018)
Num.Obs.	1669	2229	2229	1686
Num.Imp.				10
R2	0.208	0.146	0.223	0.209
R2 Adj.	0.206	0.144	0.221	0.207
AIC	1179.9	1093.8	957.6	
BIC	1223.2	1139.5	1003.2	
Log.Lik.	-581.936	-538.912	-470.776	
F	72.917	63.461	106.192	

3 Question No. 8

I want to investigate if stricter campaign finance laws leads to a decrease in the actual amount of contributions across US states. For this, I have decided to perform a panel-data regression analysis from 2000 to 2015 across the 50 US states. This analysis would be used as a part of my research paper later on.