Problem Set 7

Justin Canova

March 27, 2025

1 Question 6

The logwage is missing at 25 percentage. I would say that the missing data is missing at random due to the missingness being restricted to a single variable.

	Unique	Missing Pct.	Mean	SD	Min	Median	Max
logwage	670	25	1.6	0.4	0.0	1.7	2.3
hgc	16	0	13.1	2.5	0.0	12.0	18.0
tenure	259	0	6.0	5.5	0.0	3.8	25.9
age	13	0	39.2	3.1	34.0	39.0	46.0
		N	%				
college	college grad	530	23.8				
	not college grad	1699	76.2				
married	married	1431	64.2				
	single	798	35.8				

2 Question 7

As we see from the below table the $\beta 1$ only varies by a max range of .006 though all estimates are off by about .03. I may not have done the estimates correctly, but I they are all very close and I do not see much of a difference between the imputed models in the given sample.

	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.685	0.702	0.618	0.700	0.656
	(0.125)	(0.125)	(0.126)	(0.127)	(0.128)
hgc	0.060	0.061	0.065	0.061	0.059
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
tenure	0.023	0.024	0.022	0.024	0.023
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
age	0.000	−0.002	0.000	−0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
marriedsingle	−0.032	−0.023	−0.020	−0.019	−0.024
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
collegenot college grad	0.105	0.132	0.110	0.125	0.108
	(0.028)	(0.028)	(0.028)	(0.029)	(0.029)
Num.Obs.	2246	2246	2246	2246	2246
R2	0.234	0.232	0.233	0.223	0.213
R2 Adj.	0.232	0.230	0.231	0.221	0.211
AIC	1520.1	1520.9	1562.9	1606.2	1633.5
BIC	1560.1	1561.0	1602.9	1646.2	1673.5
Log.Lik.	−753.037	−753.469	−774.427	−796.080	−809.727
F	136.666	135.038	136.213	128.617	121.000
RMSE	0.34	0.34	0.34	0.34	0.35

3 Question 8

My project is slow, I have be prioritizing the completion of my second year paper. My project is a web scraping longitudinal analysis of meeting minutes for a makerspace in Dallas. I am scraping meeting attendance and committees and will do a poisson regression to help understand how organizations organically grow with membership.