PS7$_T$*hatcher*

rachel.e.thatcher-1

March 2018

Question 6

Table 1:

Statistic	N	Mean	St. Dev.	Min	Max
logwage	1,669	1.625	0.386	0.005	2.261
hgc	2,229	13.101	2.524	0	18
tenure	2,229	5.971	5.507	0.000	25.917
age	2,229	39.152	3.062	34	46

There are 1,669 logwages recorded and 2,229 in total, so they are missing at a rate of .749. I think it could be either MAR or MNAR, but I am not sure. I looked up a bunch of definitions but I'm still having trouble really differentiating them. Maybe leaning towards MAR?

Question 7
See table at end of document

Looking at the three models, B^1 appears to be .0062 for the first two, which is close to .093 but not quite there. I think they are fairly accurate, but I feel like I would need to see examples of how close it should be to gauge more - if that makes sense. Like .062 is fairly close to .093 in the scheme of all numbers, but if you are only looking at numbers between .06 and .095, then they are fairly far apart. Looking at regression 3 it is event farther away at .051.

The last two methods - or the second to last one since we can't do the last - has a value of .06, which is close to the rest of our methods.

Question 8
As we talked about after class the other day I am going to be working with twitter data as well as the data from the site about economists. I haven't really started, but I am hoping to get a fair amount done this upcoming week before spring break.

Table 2:

	Dependent variable:		
	logwage		
	(1)	(2)	(3)
hgc	0.062***	0.062***	0.051***
	(0.005)	(0.005)	(0.005)
collegenot college grad	0.146***	0.146***	0.174***
	(0.035)	(0.035)	(0.027)
tenure	0.023***	0.023***	0.015***
	(0.002)	(0.002)	(0.001)
age	−0.001	−0.001	−0.001
	(0.003)	(0.003)	(0.002)
marriedsingle	−0.024	−0.024	−0.031**
	(0.018)	(0.018)	(0.014)
Constant	0.639***	0.639***	0.794***
	(0.146)	(0.146)	(0.119)
Observations	1,669	1,669	2,229
R^2	0.195	0.195	0.127
Adjusted R^2	0.192	0.192	0.125
Residual Std. Error	0.346 (df = 1663)	0.346 (df = 1663)	0.321 (df = 2223)
F Statistic	80.508*** (df = 5; 1663)	80.508*** (df = 5; 1663)	64.879*** (df = 5; 2223)
Note:			*p<0.1; **p<0.05; ***p<0.01