

# PS7 Yarberry

Megan N. Yarberry

March 10, 2020

## 1 Question 6

Table 1:

Statistic Max	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)
logwage 2.261	1,669	1.625	0.386	0.005	1.362	1.936
hgc 18	2,229	13.101	2.524	0	12	15
tenure 25.917	2,229	5.971	5.507	0.000	1.583	9.333
age 46	2,229	39.152	3.062	34	36	42

At what rate are log wages missing?

- Over 25 percent of logwage data is missing

Do you think the logwage variable is most likely to be MCAR, MAR, or MNAR?

- MAR able to determine the data using a regression set and the current available data at hand

## 2 Question 7

The true value of  $\beta_1 = 0.093$ . Based on the following two equations the true value will go up the true value goes up after adding the random missing variables.

Table 2: Regression Results

	<i>Dependent variable:</i>		
		logwage	
	(1)	(2)	(3)
hgc	0.062*** (0.005)	0.049*** (0.004)	0.062*** (0.004)
collegenot college grad	0.146*** (0.035)	0.160*** (0.026)	0.146*** (0.025)
tenure	0.023*** (0.002)	0.015*** (0.001)	0.023*** (0.001)
age	-0.001 (0.003)	-0.001 (0.002)	-0.001 (0.002)
marriedsingle	-0.024 (0.018)	-0.029** (0.014)	-0.024* (0.013)
Constant	0.639*** (0.146)	0.833*** (0.115)	0.639*** (0.111)
Observations	1,669	2,229	2,229
R <sup>2</sup>	0.195	0.132	0.268
Adjusted R <sup>2</sup>	0.192	0.130	0.266
Residual Std. Error	0.346 (df = 1663)	0.311 (df = 2223)	0.300 (df = 2223)
F Statistic	80.508*** (df = 5; 1663)	67.496*** (df = 5; 2223)	162.884*** (df = 5; 2223)

*Note:*

\*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

### 3 Question 8

For my final project I have made some slow progress in scrapping the data for my project. I have a clear direct since I want to use something similar to what my thesis is going to be over for my master's program. Now it's just in the process of scrapping that extra data to prove my hypothesis. With everything going on right now with my internship, teaching and my own classes I haven't been able to get very far but look forward to doing some deep diving over spring break when I have a little more time.