# PS7 Garcia

Vladimir Garcia

March 2018

Table 1:

| Statistic | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| logwage | 1,628 | 1.627 | 0.385 | 0.005 | 2.261 |
| hgc | 2,176 | 13.122 | 2.495 | 4 | 18 |
| tenure | 2,176 | 6.115 | 5.495 | 0.083 | 25.917 |
| age | 2,176 | 39.130 | 3.056 | 34 | 46 |

At what rate are log wages missing? Do you think the logwage variable is most likely to be MCAR, MAR, or MNAR?

The wages are missing at a rate of 0.25, for which I believe the data for this variable is MNAR since it seems to be pretty substantial.

Tell me about the progress you've made on your project. What data are you using? What kinds of modeling approaches do you think you're going to take?

So far, I have been checking how to perform the data scraping part of my project and hope to making more progress over the spring break. I am still unsure about the modeling approaches.

Table 2: results

| | *Dependent variable:* |
|---|---|
| | logwage |
| hgc | 0.064*** |
| | (0.006) |
| collegenot college grad | 0.154*** |
| | (0.036) |
| tenure | 0.024*** |
| | (0.002) |
| age | -0.0004 |
| | (0.003) |
| marriedsingle | -0.025 |
| | (0.018) |
| Constant | 0.586*** |
| | (0.149) |
| Observations | 1,628 |
| $R^2$ | 0.200 |
| Adjusted $R^2$ | 0.197 |
| Residual Std. Error | 0.345 (df = 1622) |
| F Statistic | 81.015*** (df = 5; 1622) |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 |