# PS7

Joel McGuire

March 2018

## 1 Project Progess

I'm using Twitter data to compare the future orientation of tweets between states and see if that has predictive power for macroeconomic indicators like savings rate. I'm not sure about modeling yet. I have all the code written, I just need to collect about a hundred million tweets which is difficult because you have to capture about 10x as much information than you need and 100 million tweets would already be much more than 20 gigabytes.

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 |

Table 2:

| | *Dependent variable:* | | |
|---|---|---|---|
| | logwage | | |
| | (1) | (2) | (3) |
| hgc | 0.062*** | 0.049*** | 0.052*** |
| | (0.005) | (0.004) | (0.005) |
| collegenot college grad | 0.146*** | 0.160*** | 0.167*** |
| | (0.035) | (0.026) | (0.027) |
| tenure | 0.023*** | 0.015*** | 0.016*** |
| | (0.002) | (0.001) | (0.001) |
| age | −0.001 | −0.001 | −0.003 |
| | (0.003) | (0.002) | (0.002) |
| marriedsingle | −0.024 | −0.029** | −0.022 |
| | (0.018) | (0.014) | (0.014) |
| Constant | 0.639*** | 0.833*** | 0.839*** |
| | (0.146) | (0.115) | (0.118) |
| Observations | 1,669 | 2,229 | 2,229 |
| $R^2$ | 0.195 | 0.132 | 0.146 |
| Adjusted $R^2$ | 0.192 | 0.130 | 0.144 |
| Residual Std. Error | 0.346 (df = 1663) | 0.311 (df = 2223) | 0.320 (df = 2223) |
| *Note:* | | | *p<0.1; **p<0.05; ***p<0.01 |