# PS7

Travis Richardson

March 2019

## 1 Wages

Table 1:

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

Out of the 2,229 variables, there are only 1,669 logwage variables. So .25 of them are missing, about 1 out of every 4 is missing. These missing values would most likely be MNAR. The wage seems like a very important piece of information given the name of the data set is wages. This missing data does not seem like it can be ignored.

## 2 Listwise Table

## 3 Regression Imputation Table

## 4 Project

So far I have obtained the data I am going to use. I decided to use EPL data for the past 12 seasons. The data includes the places each team finished along with many variables about the seasons stats. I am planning an comparing the teams that have been involved in all 12 season to teams that have not been in every season to understand the differences in the two sets of teams.

Table 2:

| | *Dependent variable:* |
|---|---|
| | logwage |
| hgc | 0.062*** |
| | (0.005) |
| collegenot college grad | 0.146*** |
| | (0.035) |
| tenure | 0.023*** |
| | (0.002) |
| age | −0.001 |
| | (0.003) |
| marriedsingle | −0.024 |
| | (0.018) |
| Constant | 0.639*** |
| | (0.146) |
| Observations | 1,669 |
| $R^2$ | 0.195 |
| Adjusted $R^2$ | 0.192 |
| Residual Std. Error | 0.346 (df = 1663) |
| F Statistic | 80.508*** (df = 5; 1663) |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 |

Table 3:

| | *Dependent variable:* | | |
|---|---|---|---|
| | logwage | | logwa |
| | (1) | (2) | |
| hgc | 0.062*** | 0.049*** | 0 |
| | (0.005) | (0.004) | |
| collegenot college grad | 0.146*** | 0.160*** | 0 |
| | (0.035) | (0.026) | |
| tenure | 0.023*** | 0.015*** | 0 |
| | (0.002) | (0.001) | |
| I(tenureˆ2) | | | − |
| | | | |
| age | −0.001 | −0.001 | 0. |
| | (0.003) | (0.002) | |
| marriedsingle | −0.024 | −0.029** | − |
| | (0.018) | (0.014) | |
| Constant | 0.639*** | 0.833*** | 0 |
| | (0.146) | (0.115) | |
| Observations | 1,669 | 2,229 | |
| $R^2$ | 0.195 | 0.132 | |
| Adjusted $R^2$ | 0.192 | 0.130 | |
| Residual Std. Error | 0.346 (df = 1663) | 0.311 (df = 2223) | 0.000 |
| F Statistic | 80.508*** (df = 5; 1663) | 67.496*** (df = 5; 2223) | 238,951,667,917,992,846,426 |

*Note:*