**“pred\_qual\_sep\_wages.sas” Output Description**

**Datasets:**

The output that will be produced are two CSV files -- med\_baseline\_qual\_sep\_vars.csv and med\_expanded\_qual\_sep\_vars.csv. The files are structured as below, where XXX refers to “bln” for baseline, or “exp” for expanded. To get the values for firm 10’s 2005, high quality, low separation wages and workers, one would need to find the observation with payer\_tin\_w2\_max==10, tax\_yr==2005, abv\_med\_qual\_XXX==1, and abv\_med\_sep\_XXX==0.

* payer\_tin\_w2\_max
  + The EIN of the firm associated with the observation.
* tax\_yr
  + The year of the associated observation.
* abv\_med\_qual\_XXX
  + An indicator, =1 for the observations that are above median XXX quality
* abv\_med\_sep\_XXX
  + An indicator, =1 for the observations that are above median XXX predicted separation probability.
* w2\_XXX\_qual\_sep
  + The mean wages of workers in the given firm-year-quality-separation probability cell.
* num\_XXX\_qual\_sep
  + The number of workers in the given firm-year-quality-separation probability cell.

**Regressions:**

The regression output will be in the accompanying "pred\_qual\_sep\_wages.lst" file produced by running the program. It will have 4 regression models in the following order:

*Quality regressions:*

#1) proc reg: baseline quality model, level wages

equation:

wages = const. + age +age^2 + age^3 + age^4 + male + female + inventor + male\*age + male\*age^2 + male\*age^3 + male\*age^4 + female\*age + female\*age^2 + female\*age^3 + female\*age^4 + inventor\*age + inventor\*age^2 + inventor\*age^3 + inventor\*age^4 + male\*inventor + female\*inventor + tax year fixed effects

#2) proc reg: expanded quality model (i.e., with wage history), level wages

equation:

wages = const. + age +age^2 + age^3 + age^4 + male + female + inventor + male\*age + male\*age^2 + male\*age^3 + male\*age^4 + female\*age + female\*age^2 + female\*age^3 + female\*age^4 + inventor\*age + inventor\*age^2 + inventor\*age^3 + inventor\*age^4 + male\*inventor + female\*inventor + previous\_wages + previous\_wages^2 + first\_job + tax year fixed effects

*Separation regressions:*

#3) proc reg: baseline separation linear probability model

equation:

1(separation) = const. + age +age^2 + age^3 + age^4 + male + female + inventor + male\*age + male\*age^2 + male\*age^3 + male\*age^4 + female\*age + female\*age^2 + female\*age^3 + female\*age^4 + inventor\*age + inventor\*age^2 + inventor\*age^3 + inventor\*age^4 + male\*inventor + female\*inventor + tax year fixed effects

#4) proc reg: expanded separation linear probability model

equation:

1(separation) = const. + age +age^2 + age^3 + age^4 + male + female + inventor + male\*age + male\*age^2 + male\*age^3 + male\*age^4 + female\*age + female\*age^2 + female\*age^3 + female\*age^4 + inventor\*age + inventor\*age^2 + inventor\*age^3 + inventor\*age^4 + male\*inventor + female\*inventor + previous\_wages + previous\_wages^2 + first\_job + tax year fixed effects