**8-16-17 SAS code data sets**

* **cohort\_appyear\_q\_wages.csv*:*** This data set contains variables at the TIN by tax year by application year wage quartile level on the cohort earnings.
  + **payer\_tin\_w2\_max**: The employer TIN of the observation.
  + **tax\_yr:** The tax year of the observation.
  + **quartile**: The within-firm wage quartile of the observation in the given tax year.
  + **cht\_quartile\_size**: This variable reports the number of people working at the given firm in the given tax year within the given wage quartile. For example, this variable’s value for the firm 3, tax year 2000, quartile 2 observation would store the number of people who worked at firm 3 in the year 2000 who were in the second wage quartile.
  + **cht\_q\_wages\_YYYY**: This variable reports the mean year YYYY wages of people working at the given firm in the given tax year within the given wage quartile. For example, the cht\_q\_wages\_2006 value for the firm 3, tax year 2000, quartile 2 observation would store the mean 2006 wages of people who worked at firm 3 in the year 2000 who were in the second wage quartile. Note that if the worker drops out of the panel their wages are counted as zero for this variable.
* **entrant\_wage\_growth.csv**: This data set contains variables at the TIN by tax year level (i.e. not cut by quartile) on the wage growth for people who entered into the firm 3 years ago.
  + **payer\_tin\_w2\_max**: The employer TIN of the observation.
  + **tax\_yr:** The tax year of the observation.
  + **ents\_who\_stay\_3yrs:** This variable measures the number of people working at the given firm in the given tax year who entered 3 years ago. For example, for the firm 3, tax year 2006 observation this variable would store the number of employees at firm 3 who were not working for firm 3 in 2002 but were working for firm 3 in 2003 and are now working at firm 3 in 2006. Note that this variable is only defined for tax year observations within 2004 and 2014.
  + **ent\_wage\_growth\_in3**: This variable measures the mean wage growth of people working at the given firm in the given tax year who entered 3 years ago. For example, for the firm 3, tax year 2006 observation this variable would store the mean of the 2006 wages minus the 2003 wages of employees at firm 3 who were not working for firm 3 in 2002 but were working for firm 3 in 2003 and are now working at firm 3 in 2006. Note that this variable is only defined for tax year observations within 2004 and 2014.
  + **ent\_wage\_growth\_in3\_dhs**: This variable measures the mean David, Haltiwanger, and Schuh metric of wage growth of people working at the given firm in the given tax year who entered 3 years ago. This is calculated as 2\*(currwages-lastwages)/(currwages+lastwages). For example, for the firm 3, tax year 2006 observation this variable would store this measure for the 2006 wages relative to the 2003 wages of employees at firm 3 who were not working for firm 3 in 2002 but were working for firm 3 in 2003 and are now working at firm 3 in 2006. Note that this variable is only defined for tax year observations within 2004 and 2014.