This document provides variable descriptions for the 3 data sets with the “quartile\_wage\_measures\_20180428.sas” code file. The data sets are not necessarily balanced panels at the firm—tax\_yr—quartile level. A set of observations (e.g. firm 100's year 2000 quartiles) can be missing because there were fewer than 4 employees in that year at that firm.

mean\_quartile\_wages.csv

(unit of obs: payer\_tin\_w2\_max BY tax\_yr BY quartile)

* + **payer\_tin\_w2\_max**: Firm TIN
  + **tax\_yr**: Year of W2 observations, ranges from 1996 to 2014
  + **quartile**: The wage quartile of the firm, ranges from 1 to 4
  + **mean\_quartile\_wages**: The mean wage for the given firm in the given tax year in the given quartile.
  + **emp\_quartile**: The number of employees at the given firm in the given tax year in the given quartile.

mean\_quartile\_stayerwages.csv

(unit of obs: payer\_tin\_w2\_max BY tax\_yr BY quartile)

* + **payer\_tin\_w2\_max**: Firm TIN
  + **tax\_yr**: Year of W2 observations, ranges from 1996 to 2014
  + **quartile**: The wage quartile of the firm, ranges from 1 to 4
  + **cht\_quartile\_size**: The size of the tax year cohort at the given firm in the given quartile. So the associated value for the firm 100, tax year 2000, and quartile 3 observation would give the number of people employed at firm 100 in 2000 in quartile 3.
  + **mean\_quartile\_stayerwagesXXXX**: Here XXXX stands for the year of comparison. This variable gives the mean wages of stayers in year XXXX from the given tax year cohort at the firm within the same quartile. So if you are looking at the firm 100, tax year 2000, and quartile 3 observation, the variable mean\_quartile\_stayerwages2002 would give the mean 2002 wages of employees who were at firm 100 in 2000 in quartile 3 but and also work at firm 100 in 2002. It is important to note that due to Pat's request to keep cohorts fixed with regards to time and cohort that the stayers from quartile 3 need not be in quartile 3 in 2002 even though they were in quartile 3 in 2000.
  + **stayers\_quartile\_sizeXXXX**: Here XXXX stands for the year of comparison. This variable gives the number of stayers for the given firm, tax year, and quartile combination in the year XXXX. So the firm 100, tax year 2000, and quartile 3 observation's value for stayers\_quartile\_size2004 would contain the number of employees who were working at firm 100 in 2000 in the third wage quartile who continue to work at firm 100 in 2004. Again, it is important to note that the stayers from quartile 3 need not be in quartile 3 in 2004 even though they were in quartile 3 in 2000.

quartile\_sep\_rate\_cht.csv

(unit of obs: payer\_tin\_w2\_max BY tax\_yr BY quartile)

* + **payer\_tin\_w2\_max**: Firm TIN
  + **tax\_yr**: Year of W2 observations, ranges from 1996 to 2014
  + **quartile**: The wage quartile of the firm, ranges from 1 to 4
  + **cht\_quartile\_size**: The size of the tax year cohort at the given firm in the given quartile. So the associated value for the firm 100, tax year 2000, and quartile 3 observation would give the number of people employed at firm 100 in 2000 in quartile 3. This is the same variable as in “mean\_quartile\_stayerwages.csv”.
  + **stayers\_quartile\_sizeXXXX**: Here XXXX stands for the year of comparison. This variable gives the number of stayers for the given firm, tax year, and quartile combination in the year XXXX. So the firm 100, tax year 2000, and quartile 3 observation's value for stayers\_quartile\_size2004 would contain the number of employees who were working at firm 100 in 2000 in the third wage quartile who continue to work at firm 100 in 2004. Again, it is important to note that the stayers from quartile 3 need not be in quartile 3 in 2004 even though they were in quartile 3 in 2000. This is the same variable as in “mean\_quartile\_stayerwages.csv”.
  + **cht\_quartile\_sep\_rateXXXX**: Here XXXX stands for the year of comparison. This variable gives the fraction of people at the firm in the observation's tax year in the observation's quartile who are no longer employed at the firm in the year XXXX (they can be in a different quartile). So for our favorite firm 100, tax year 2000, and quartile 3 observation, the value for cht\_quartile\_sep\_rate2011 would give the fraction of firm 100's quartile 3 workers in 2000 who no longer work at firm 100 in 2011 (again they can be in a different quartile from 3).