

Problem Set 1: Due Monday, January 29

1. In this problem, we will construct Gabaix's granular residual.
2. The data necessary to construct the granular residual can be found on WRDS (<https://wrds-web.wharton.upenn.edu/wrds/>) If you haven't already, request an account to WRDS. It might take a day or two to get access to WRDS, so you submit your registration request as soon as possible. From the website, download the following variables from the CRSP/Compustat Database: emp, sale, eostat, idbflag, loc, sic for the 1957 to 2013 period. Save your file as gabaix.dta. On my website, I have some code (and a small dataset, aggregate.csv) that will be useful in producing the granular residual.
3. On line 9 of ps1.do, I have a statement that begins with `drop if sic==...` What is this line of code doing? Why does it make sense to drop firms within these industries? See Appendix C in the paper.
4. In which years was the granular residual smallest/largest? Which firms were responsible for these large deviations?
5. Were 2009 and 2010 "granular years"? Would the granular residual be much more negative, these years, with banks included in the sample?